Probabilistic Methods in Combinatorics

Joshua Erde

Department of Mathematics,
Universität Hamburg.

Contents

Preface

These are skeleton notes for the course "Probabilistic Methods in Combinatorics" for the summer semester of 2017/2018 at the University of Hamburg. These notes are intended to indicate the key points of the lectures. Students are highly encouraged to take their own notes to supplement this material, and to consider reviewing the material both before and after the lectures, if time permits.

1 Preliminaries

1.1 Probability Theory

This section is intended as a short introduction to the very basics of probability theory, covering only the basic facts about finite probability spaces that we will need to use in this course.

Definition (Probability space, events, elementary events)**.** A *probability space* is a triple $(\Omega, \Sigma, \mathbb{P})$, where Ω is a set, $\Sigma \subseteq 2^\Omega$ is a σ-algebra (A non-empty collection of subsets of Ω which is closed under taking complements and countable unions/intersections), and $\mathbb{P}$ is a measure on Σ with $\mathbb{P}(\Omega) = 1$. The elements of Σ are called *events* and the elements of Ω are called *elementary events*. For an event A, $\mathbb{P}(A)$ is called the *probability of* A.

Definition (Finite probability space, uniform distribution)**.** A *finite probability spaces* is one where Ω is finite and $\Sigma = 2^\Omega$. In this case the probability measure $\mathbb{P}$ is determined by the value it takes on elementary events. That is, given any function $p : \Omega \to [0, 1]$ that satisfies $\sum_{\omega\in\Omega} p(\omega) = 1$, then the function on Σ given by $\mathbb{P}(A) = \sum_{\omega\in A} p(\omega)$ is a probability measure.

The *uniform distribution* on Ω, is the probability measure given by

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} \text{ for all } A \subseteq \Omega.$$

Definition (Random graph)**.** The probability space of *random graphs* $\mathcal{G}(n, p)$ is a finite probability space whose elementary events are all graphs on a fixed set of n vertices, and where the probability of each graph with m edges is

$$p(G) = p^m(1 - p)^{\binom{n}{2}-m}$$

- We will often denote an arbitary event from this space by $G(n, p)$. Note that p can, and often will be, a function of n.

Definition (Almost surely/with high probability)**.** Given a property of graphs P we say that $G(n, p)$ satisfies P *almost surely* or *with high probability* if

$$\lim_{n\to\infty} \mathbb{P}(G(n, p) \text{ satisfies } P) = 1.$$

Lemma 1.1 (The union bound)**.** *For any $A_1, A_2, \ldots, A_n \in \Sigma$,*

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \Sigma_{i=1}^{n} \mathbb{P}(A_i)$$

Definition (Independence, mutual independence)**.** Two events $A, B \subseteq \Sigma$ are *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

More generally, a set of events $\{A_1, A_2, \ldots, A_n\}$ is *mutually independent* if, for any subset of indices $I \subseteq [n]$,

$$\mathbb{P}\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} \mathbb{P}(A_i).$$

Definition (Conditional probability)**.** Given two events $A, B \in \Sigma$ such that $\mathbb{P}(B) \neq 0$, we define the *conditional probability of A, given that B occurs*, as

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

- A and B are independent if and only if $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Definition (Random variable)**.** A *real random variable* on probability space $(\Omega, \Sigma, \mathbb{P})$ is a function $X : \Omega \to \mathbb{R}$ that is $\mathbb{P}$-measurable.

In a finite probability space, any function $X : \Omega \to \mathbb{R}$ defines a random variable. Given a measurable set $A \subseteq \mathbb{R}$ the probability that the value X takes lies in A is $\mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$ which we will write as $\mathbb{P}(X \in A)$.

Definition (Expectation)**.** The *expectation* of a random variable X is

$$\mathbb{E}(X) = \int_\Omega X(\omega)\, d\mathbb{P}(\omega).$$

In the case of a finite probability space this can be expressed more clearly as

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} p(\omega) X(\omega).$$

- The set of random variables forms an algebra over $\mathbb{R}$ with addition and multiplication defined pointwise. For example the random variable $X + Y$ is the function from Ω to $\mathbb{R}$ defined by $(X + Y)(\omega) = X(\omega) + Y(\omega)$.

Lemma 1.2 (Linearity of expectation)**.** *For any two random variables X and Y*

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

Definition (Independence of random variables)**.** Two random variable X, Y are *independent* if, for any two measurable sets $A, B \subseteq \mathbb{R}$ we have

$$\mathbb{P}(X \in A \text{ and } Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

Lemma 1.3. *For any two independent random variables, X and Y,*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

1.2 Useful Estimates

Many proofs using the probabilistic method will reduce to calculating certain probabilities, for example showing they are less than 1 or tend to 0. For this purpose we will often need to estimate some quite complicated combinatorial expressions. In this section we will note down some useful estimates to apply later, both weak and strong.

- A weak upper bound for $n!$ is $n! \leq n^n$.
- A stronger estimate is
$$\left(\frac{n}{e}\right)^n \leq n! \leq en\left(\frac{n}{e}\right)^n.$$
- Stirling's formula gives
$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$
- A weak upper bound for $\binom{n}{k}$ is $\binom{n}{k} \leq n^k$, or even worse $\binom{n}{k} \leq 2^n$.
- A stronger estimate is
$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k.$$
- For the middle binomial co-efficient we have
$$\frac{2^{2m}}{2\sqrt{m}} \leq \binom{2m}{m} \leq \frac{2^{2m}}{\sqrt{2m}}$$
- Since $1 + x \leq e^x$ for all real x,
$$(1-p)^m \leq e^{-mp}.$$
- If $0 \leq p \leq \frac{1}{2}$, then
$$1 - p \geq e^{-2p}.$$

Definition (O, o, Ω,ω)**.** Given two functions $f, g : \mathbb{N} \to \mathbb{R}$ we say that:

- $f = O(g)$ if there exists $C > 0$ such that for all sufficiently large n, $f(n) \leq Cg(n)$;
- $f = \Omega(g)$ if there exists $C > 0$ such that for all sufficiently large n, $f(n) \geq Cg(n)$;
- $f = o(g)$ if for sufficiently large n, $f(n) \leq Cg(n)$, for any fixed $C > 0$;
- $f = \omega(g)$ if for sufficiently large n, $f(n) \geq Cg(n)$, for any fixed $C > 0$;

2 The Probabilistic Method

In its most basic form the probabilistic method can be described as follows: In order to prove the existence of a combinatorial object with certain properties we pick a random object from a suitable probability space and calculate the probability that it satisfies these conditions. If we can prove that this probability is strictly positive, then we conclude that such an object must exist, since if none of the objects satisfied the conditions, the probability of a random object doing so would be zero.

The probabilistic method is useful in cases when an explicit construction of such an object does not seem feasible, and when we're more interested in the existence of such an object than in a specific example.

2.1 Ramsey Numbers

Theorem 2.1 (Ramsey's Theorem, **without proof**). *For any k, l there exists an n such that every $|G| \geq n$ either has $\omega(G) \geq k$ or $\alpha(G) \geq l$.*

Definition (Ramsey numbers). Let $k, l \in \mathbb{N}$. The *Ramsey number* $R(k, l)$ is the smallest n such that

$$R(k,l) := \min\{n \,:\, \text{any graph on } n \text{ vertices contains a clique of size } k \text{ or an independent set of size } l\}.$$

Theorem 2.2 (A lower bound for the diagonal Ramsey numbers). *For any $k \geq 3$*

$$R(k,k) > 2^{k/2-1}.$$

2.2 Set Systems

Definition. A family $\mathcal{F}$ of sets is *intersecting* if for all $A, B \in \mathcal{F}, A \cap B \neq \emptyset$

- The size of a largest intersecting family of subsets of $[n]$ is 2^{n-1}.

Theorem 2.3 (The Erdős-Ko-Rado Theorem). *For any n and $k < n/2$, if $\mathcal{F} \subset [n]^{(k)}$ is an intersecting family, then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1}$$

Definition. Let $n(k, l)$ be the largest n such that there exists two families of sets $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ and $\mathcal{B} = \{B_1, B_2, \ldots, B_n\}$ satisfying the following properties

- $|A_i| = k$, $|B_i| = l$ for all $1 \leq i \leq n$;
- $A_i \cap B_i = \emptyset$ for all $1 \leq i \leq n$;
- $A_i \cap B_j \neq \emptyset$ for all $i \neq j$, $1 \leq i, j \leq n$.

Theorem 2.4. *For any $k, l \geq 1$, $n(k, l) = \binom{k+l}{k}$.*

3 The First Moment Method

Lemma 3.1. *Let $X_1, \ldots, X_n$ be random variables, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, and $X = \sum_{i=1}^{n} \lambda_i X_i$. Then*

$$\mathbb{E}(X) = \sum_{i=1}^{n} \lambda_i \mathbb{E}(X_i).$$

Definition (Indicator random variables)**.** For an event A in a probability space the *indicator random variable* I_A is defined by

- $I_A(\omega) = 1$ if $\omega \in A$
- $I_A(\omega) = 0$ if $\omega \notin A$

- $\mathbb{E}(I_A) = \mathbb{P}(A)$.
- If $X = \sum_{i=1}^{n} I_{A_i}$, then $\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{P}(A_i)$.
- For any random variable X, there always exists a point ω in the probability space such that $X(\omega) \geq \mathbb{E}(X)$ and a point ω' such that $X(\omega') \leq \mathbb{E}(X)$.

The above is the key idea to this section. Showing that the expected value of a random variable is small or large will guarantee that there is an point in the probability space on which this variable is small or large.

Lemma 3.2. *[Markov's Inequality] Let X be a non-negative random variable and $a > 0$, then*

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

3.1 Hamiltonian Paths in a Tournament

Definition. A *tournament* is an orientation of a complete graph. A *Hamiltonian path* in a tournament is a directed path that meets every vertex.

Theorem 3.3. *There exists a tournament T on n vertices which has at least $n!/2^{n-1}$ Hamiltonian paths.*

3.2 Turán's Theorem

Lemma 3.4. *Let G be a graph and, for each $v \in V(G)$, let $d(v)$ be the degree of v. Then G contains an independent set of size at least*

$$\sum_{x \in V} \frac{1}{d(v) + 1}$$

Corollary 3.5. *Let G be a graph on n vertices. Then G contains a clique of size at least*

$$\sum_{x \in V} \frac{1}{n - d(v)}$$

It is simple to deduce (a weak form of) Turán's Theorem from this corollary.

Theorem 3.6 (Turán's Theorem)**.** *Let G be a graph on n vertices such that $K_r \not\subset G$, then*

$$e(G) \leq \frac{(r-2)n^2}{2(r-1)}.$$

3.3 Crossing Number of Graphs

Definition. Given a graph $G = (V, E)$ an *embedding* of G into the plane is a planar representation of G, where each vertex is represented by a point, and each edge from u to v is represented by a curve between the points represented by u and v. The *crossing number* of an embedding of a graph is the number of pairs of curves which intersect, which do not share endpoints. The *crossing number* of G, $\mathrm{cr}(G)$ is the minimal crossing number of a planar embedding of G.

Theorem 3.7. *Let G be a graph such that $|E| \geq 4|V|$, then*

$$cr(G) \geq \frac{|E|^3}{64|V|^2}.$$

4 Alterations

Sometimes it is the case that the first attempt to find a 'good' object via a random construction fails, but what we find is an object which *almost* satisfies our conditions. Often it is possible to then deterministically modify this object to get what we need.

4.1 Ramsey Numbers Again

Theorem 4.1. *For any integer n,*

$$R(k,k) > n - 2.\binom{n}{k}2^{-\binom{k}{2}}.$$

Theorem 4.2 (Without proof). *For all integers n and $p \in [0,1]$*

$$R(k,l) > n - \binom{n}{k}p^{\binom{k}{2}} - \binom{n}{l}(1-p)^{\binom{l}{2}}.$$

Theorem 4.3 (Without proof). *If there exists an integer n and $p \in [0,1]$ such that*

$$\binom{n}{k}p^{\binom{k}{2}} + \binom{n}{l}(1-p)^{\binom{l}{2}} < 1$$

Then $R(k,l) > n$.

4.2 Graphs of High Girth and High Chromatic Numbers

Definition (k-colouring, chromatic number and girth)**.** A *k-colouring* of a graph G is a function $f : V(G) \to [k]$ such that $(v,w) \in E(G) \Rightarrow f(v) \neq f(w)$. The *chromatic number* of G, $\chi(G)$, is the smallest k such that a k-colouring exists. The *girth* of a graph G, $g(G)$, is the size of the shortest cycle in G.

Theorem 4.4. *For any $k, l > 0$, there exists a graph G such that $\chi(G) > k$ and $g(G) > l$.*

5 Dependent random choice

Dependent random choice is a relatively new technique based on the alteration method which has been used in various contexts, normally to do with embedding sparse graphs. The basic idea can be summarised as follows: We would like to find, in a dense graph, a set of vertices U such that every small subset of U has many common neighbours. To do this, we first pick a small set of vertices T at random from the graph and let U' be the set of common neighbours of T. Intuitively, if we have some subset of G with not many common neighbours, then it is unlikely that all the members of T will lie in this set of common neighbours, and hence it is unlikely to be a subset of U'. Therefore the expected number of 'bad' subsets in U' will be small and so by removing a small number of vertices, one from each 'bad' set, we should find a set U with the desired properties. Such a set U will be useful for finding embeddings of bipartite graphs.

Lemma 5.1. *Let G be a graph with $|G| = n$ and let $d = 2|E(G)|/n$ be the average degree of G. If there exist positive integers t, a, m, r such that*

$$\frac{d^t}{n^{t-1}} - \binom{n}{r}\left(\frac{m}{n}\right)^t \geq a,$$

then G contains a subset U of at least a vertices such that every subset $R \subset U$ of size $|R| = r$ has at least m common neighbours.

Lemma 5.2. *Let G be a graph, a, m, r be positive integers and suppose there exists a subset $U \subset V(G)$ of at least a vertices such that every subset $R \subset U$ of size r has at least m common neighbours.*

If H is a bipartite graph on vertex sets A and B such that $|H| \leq m$, $|A| \leq a$ and every vertex in B has degree at most r, then H is a subgraph of G.

5.1 Turán Numbers of Bipartite Graphs

Theorem 5.3 (The Erdős-Stone Theorem, **without proof**)**.** *For any graph H with $\chi(H) \geq 3$*

$$ex(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right)\binom{n}{2}$$

- For graphs with $\chi(H) = 2$, the situation is more complicated.

Theorem 5.4. *Let H be a bipartite graph on vertex sets A and B such that all vertices in B have degree at most r. Then there exists some constant $c = C(H)$ such that*

$$ex(n, H) \leq cn^{2-\frac{1}{r}}.$$

5.2 The Ramsey Number of the Cube

Definition (Ramsey number of H and hypercube)**.** The *Ramsey number* of an arbitrary graph H is

$$r(H) = \min\{n : \text{ Every 2 colouring of } K_n \text{ contains a monochromatic copy of } H\}.$$

The *r-dimensional Hypercube*, $\mathcal{Q}_r$, is a graph with vertex set $\{0,1\}^r$ where two vertices are adjacent if and only if they differ in exactly one coordinate.

Theorem 5.5.

$$r(\mathcal{Q}_r) \leq 2^{3r}$$

5.3 Improvements

Lemma 5.6 (Without proof). *Let $\epsilon > 0$, $r \leq n$ be positive integers, and G a graph on $N > 4r\epsilon^{-r}n$ vertices with at least $\epsilon\frac{N^2}{2}$ edges. Then there is a subset $U \subset V(G)$ with $|U| > 2n$ such that number of subsets $S \subset U$ with $|S| = r$ and less than n common neighbours is at most*

$$\frac{1}{(2r)^r}\binom{|U|}{r}.$$

Theorem 5.7 (Without proof).

$$r(\mathcal{Q}_r) \leq r2^{2r+3} \leq 2^{2r+o(r)}.$$

6 The Second Moment Method

6.1 Variance and Chebyshev's Inequality

Markov's inequality tells us that, for a non-negative random variable X, if the expectation of X gets small, then it's very likely that X is small. In general, if the expectation of X is large, it is not true that it is very likely that X is large. However this will be true if the *variance* is small.

Definition (Variance)**.** The *variance* of a random variable X is

$$\mathrm{Var}(X) := \mathbb{E}\big((X - \mathbb{E}(X))^2 \big) = \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2,$$

where the first equality is the definition, and the second follows from linearity of expectation.

Definition (Covariance)**.** The *covariance* of two random variables X and Y is

$$\mathrm{Cov}(X, Y) = \mathbb{E}\Big(\big(X - \mathbb{E}(X)\big)\big(Y - \mathbb{E}(Y)\big) \Big) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Lemma 6.1. *Given a sequence of random variables $X_1, X_2, \ldots, X_n$, let $X = \sum_i X_i$. Then*

$$Var(X) = \sum_{i=1}^{n} Var(X_i) + \sum_{i \neq k} Cov(X_i, X_j).$$

- If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$.
- The converse is not true.

Lemma 6.2 (Chebyshev's Inequality)**.** *Let X be a random variable with $Var(X) < \infty$. Then for any $t > 0$*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{Var(X)}{t^2}.$$

A consequence of this lemma is the idea from the start of the section: If $\mathbb{E}(X)$ is large and $\mathrm{Var}(X)$ is small then X is very likely to be large.

6.2 Threshold Functions

Definition (Monotone properties)**.** Given a random graph $G(n, p)$ and an arbitrary graph H, a natural question to consider is what is the probability that H appears as a subgraph of G. For example let us consider the triangle, K_3. We note that the property of containing a triangle as a subgraph is a *increasing property* of graphs; that means, if it holds for a graph G and $G \subset G'$, it also holds for G'. We can similarly define a *decreasing property* of graphs, and we call a property that is either decreasing or increasing a *monotone property.*

- Let T be the random variable which counts the number of triangles in $G(n, p)$.
- $\mathbb{E}(T) = \binom{n}{3} p^3$.

- If $p(n) = o(1/n)$, then

$$\mathbb{P}(K_3 \subset G(n,p)) = \mathbb{P}(T \geq 1) \leq \mathbb{E}(T) \to 0.$$

Lemma 6.3. *Let $X_1, X_2, \ldots$ be a sequence of non-negative random variables such that*

$$\frac{Var(X_n)}{\mathbb{E}(X_n)^2} \to 0.$$

Then

$$\mathbb{P}(X_n > 0) \to 1.$$

- If $X = \sum_{A \in \mathcal{A}} I_A$, then

$$\mathrm{Var}(X) = \sum_{A \in A} \mathbb{P}(A)(1 - \mathbb{P}(A)) + \sum_{A \neq B} \mathbb{P}(A)(\mathbb{P}(B|A) - \mathbb{P}(B)) = \sum_{A \in \mathcal{A}} \mathbb{P}(A) \left(\sum_{B \in \mathcal{A}} \mathbb{P}(B|A) - \mathbb{P}(B) \right).$$

- $\mathrm{Var}(T) = \binom{n}{3} p^3 \left(3(n-3)(p^2 - p^3) + (1 - p^3)\right)$.
- If $p(n) = \omega(1/n)$, then $\mathrm{Var}(T)/(\mathbb{E}(T))^2 \to 0$ and so

$$\mathbb{P}(K_3 \subset G(n,p)) = \mathbb{P}(T \geq 1) \leq \mathbb{E}(T) \to 1.$$

Definition (Threshold function)**.** A function $r : \mathbb{N} \to \mathbb{R}$ is a *threshold function* for a monotone graph property A, if for any $p : \mathbb{N} \to [0, 1]$

- $p(n) = o(r(n)) \Rightarrow \mathbb{P}(A \text{ holds for } G(n,p)) \to 0$;
- $p(n) = \omega(r(n)) \Rightarrow \mathbb{P}(A \text{ holds for } G(n,p)) \to 1$.

6.3 Balanced Subgraphs

Definition (Density, balanced subgraph)**.** Let H be a graph with $|H| = v$ and $e(H) = e$. The *density* of H is defined to be

$$\rho(H) = \frac{e}{v}.$$

H is said to be *balanced* if no subgraph of H has strictly greater density than H itself.

Theorem 6.4. *Let H be a balanced graph with density ρ. Then*

$$r(n) = n^{-\frac{1}{\rho}}$$

is a threshold function for the property that H is a subgraph of $G(n,p)$.

Theorem 6.5 (Without proof). *Let H be a graph and $H' \subset H$ a subgraph of H with the maximum density. Then*

$$r(n) = n^{-\frac{1}{\rho(H')}}$$

is a threshold function for the property that H is a subgraph of $G(n,p)$.

7 The Hamiltonicity Threshold

7.1 The Connectivity Threshold

Theorem 7.1 (Threshold for having an isolated vertex)**.** *$r(n) = \log(n)/n$ is a threshold function for the event that $G(n,p)$ contains an isolated vertex*

- If $p(n) = o(\log(n)/n)$, then
$$\mathbb{P}(G(n,p) \text{ is connected}) \leq \mathbb{P}(X_1 = 0) \to 0.$$
- Let X_k be the random variable which counts the number of components of size k in $G(n,p)$.

Lemma 7.2. *If $p(n) = c\log(n)/n$, for $c \geq 1$, then*
$$\sum_{k=2}^{\frac{n}{2}} \mathbb{P}(X_k > 0) = o(1).$$

Theorem 7.3 (Threshold for connectivity)**.** *$r(n) = \log(n)/n$ is a threshold function for the event that $G(n,p)$ is connected.*

7.2 Posá's Rotation-Extension Technique

Definition (Rotation, transform, U, P, N, R)**.** Given a graph G and a vertex $x_0 \in V(G)$ suppose that $P = x_0x_1\dots x_k$ is a longest path in G starting at x_0. Given an edge $(x_k, x_i) \in E(G)$ a *rotation* of P is a new path $P' = x_0x_1\dots x_ix_kx_{k-1}\dots x_{i+1}$. We say that a path P' is a *transform* of P if it can be obatined from P by a sequence of rotations. Let U be the set of endvertices of all possible transforms of P and let
$$N = \{x_i : \{x_{i+1}, x_{i-1}\} \cap U \neq \emptyset\}$$
be the set of neighbours of this set in P. Finally let $R = V(P) \setminus (U \cup N)$ be the rest of the vertices in P.

Lemma 7.4. *Let G be a graph, $x_0 \in V(G)$ and P a longest path in G starting at x_0. If U, N and R are defined as above then there are no edges in G between U and R.*

Lemma 7.5. *Let G be a connected graph. Suppose that the longest path in G has length $k \geq 2$, G contains no cycles of length $k+1$ and for some $u \in \mathbb{N}$ we have that for every subset $U \subset V(G)$ with $|U| < u$*
$$|U \cup N(U)| \geq 3|U|.$$
Then there are at least $u^2/2$ non-edges in G whose addition forms a $k+1$ cycle in G.

Lemma 7.6. *Suppose c is sufficiently large and $p = c\log(n)/n$. Then almost surely in $G(n,p)$ every subset $U \subset V(G)$ with $|U| \leq n/4$ satisfies*
$$|U \cup N(U)| \geq 3|U|.$$
In particular the property that some set U fails to satisfy this property is less then n^{-2}.

Lemma 7.7. *Suppose c is sufficiently large and $p = c\log(n)/n$. Then almost surely in $G(n,p)$ contains a Hamiltonian path.*

7.3 Hamiltonicty Threshold

- If $p(n) = c\log(n)/n$ for large enough c, then $G(n,p)$ almost surely contains a Hamiltonian path and satisfies the condition of Lemma 7.5 for a large u, and so, if $G(n,p)$ does not contain a Hamiltonian cycle then there are a large number of non-edges whose addition would form a Hamiltonian cycle in $G(n,p)$

Definition (Sprinkling)**.** If we pick two random graphs $G(n,p_1)$ and $G(n,p_2)$ and let $H = G(n,p_1) \cup G(n,p_2)$ be the union of the two graphs, then it is a simple a exercise to show that H is distributed as $G(n,q)$ for $q = p_1 + p_2 - p_1p_2$. We can think of the process as a two step exposure of the edges in $G(n,q)$. This idea is sometimes called *sprinkling*, we can think of it as picking $G(n,p_1)$ and then 'sprinkling' some extra edges on top, with a fixed probability.

Theorem 7.8. *$r(n) = \log(n)/n$ is a threshold function for the event that $G(n,p)$ is Hamiltonian.*

8 Strong Concentration

8.1 Motivation

In many application of the probabilistic method, we want to bound probabilities of the for $\mathbb{P}(X \geq \mathbb{E}(X) + t)$ for some random variable X, or often more generally $\mathbb{P}(|X - \mathbb{E}(X)| \geq t)$. We call bounds for such probabilities *tail estimates*. If we can show that with high probability $|X - \mathbb{E}(X)| \leq t$, where t is considerably smaller than $\mathbb{E}(X)$, we say that X is *concentrated* about its expectation. Chebyshev's inequality is a very general result of this type, however it is quite weak.

8.2 The Chernoff Bound

Theorem 8.1 (The Chernoff bound)**.** *Let $X_1, X_2, \ldots X_n$ be independent random variables taking the values 1 and -1, each with probability $1/2$. Let $X = X_1 + X_2 + \ldots + X_n$. Then, for any $t \geq 0$,*

$$\mathbb{P}(X \geq t) < e^{-\frac{t^2}{2n}} \text{ and } \mathbb{P}(X \leq -t) < e^{-\frac{t^2}{2n}}.$$

Theorem 8.2. *The maximum degree of $G(n, 1/2)$ is almost surely $n/2 + O(\sqrt{n \log(n)})$.*

Theorem 8.3 (Without proof). *Let $X_1, X_2, \ldots, X_n$ be independent random variables, each taking values in $[0, 1]$, let $X = X_1 + X_2 + \ldots + X_n$ and let $\sigma^2 = Var(X) = \sum_i Var(X_i)$. Then for any $t \geq 0$*

$$\mathbb{P}(X \geq \mathbb{E}(X) + t) < e^{-\frac{t^2}{2(\sigma^2 + \frac{t}{3})}} \text{ and } \mathbb{P}(X \leq \mathbb{E}(X) - t) < e^{-\frac{t^2}{2(\sigma^2 + \frac{t}{3})}}.$$

Definition. For any two vertices x and y in a connected graph G we define the *distance* between x and y, $\text{dist}(x, y)$, to be the length of the shortest path between x and y. The *diameter* of a graph G is the maximum distance between any pair of vertices, and the *radius* is the minimum distance r such that there exists a vertex x such that $\text{dist}(x, y) \leq r$ for all $y \in G$.

Theorem 8.4. *Let $d \geq 2$ be fixed. Suppose $c > 0$ and $p^d n^{d-1} = \log(n^2/c)$. Then almost surely $G(n, p)$ has diameter at least d.*

Theorem 8.5 (Without proof). *Let $d \geq 2$ be fixed. Suppose $c > 0$ and $p^d n^{d-1} = \log(n^2/c)$. Then (as $n \to \infty$) with probability $e^{-c/2}$ the diameter of $G(n, p)$ is d and with probability $1 - e^{-c/2}$ the diameter is $d + 1$.*

8.3 Combinatorial Discrepancy

Definition (Discreprancy)**.** Let $(V, \mathcal{F})$ be a set system and consider a two colouring $\chi : V \to \{-1, +1\}$. For $F \in \mathcal{F}$ we define

$$\chi(F) = \sum_{i \in F} \chi(i).$$

The *discrepancy of $\mathcal{F}$ with respect to χ* is

$$\text{disc}_\chi(\mathcal{F}) = \max_{F \in \mathcal{F}} |\chi(F)|$$

and the *discrepancy* of $\mathcal{F}$ is

$$\operatorname{disc}(\mathcal{F}) = \min_{\chi} \operatorname{disc}_{\chi}(\mathcal{F}).$$

Theorem 8.6. *Let $(V, \mathcal{F})$ be a set system such that $|V| = n$ and $|\mathcal{F}| = m$. Then if the maximum size of a set $F \in \mathcal{F}$ is s*

$$disc(\mathcal{F}) \leq \sqrt{2s \log(2m)}.$$

In particular, for any $\mathcal{F}$

$$disc(\mathcal{F}) \leq \sqrt{2n \log(2m)}.$$

8.4 A Lower Bound for the Binomial Distribution

Theorem 8.7. ***[Without proof]*** *Let X be a sum of independent random variables, each taking values in $[0, 1]$, and let $\sigma = \sqrt{Var(X)} \geq 200$. Then for all $t \in [0, \sigma^2/100]$, we have*

$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \geq ce^{-\frac{t^2}{3\sigma^2}}$$

for a suitable constant $c > 0$.

Theorem 8.8. *For n even, let $X_1, X_2, \ldots, X_n$ be independent random variables taking the values 1 and -1, each with probability $1/2$. Let $X = X_1 + X_2 + \ldots + X_n$. Then we have, for any integer $t \in [0, n/8]$,*

$$\mathbb{P}(X \geq 2t) \geq \frac{1}{15} e^{-\frac{16t^2}{n}},$$

Theorem 8.9. *For any $n, m \in \mathbb{N}$ such that $15n \leq m \leq 2^{\frac{n}{8}}$ there exists a set system $(V, \mathcal{F})$ such that $|V| = n$ and $|\mathcal{F}| = m$ such that*

$$disc(\mathcal{F}) \geq \Omega\left(\sqrt{n \log\left(\frac{m}{15n}\right)}\right).$$

9 The Lovás Local Lemma

9.1 The Local Lemma

In a typical probabilistic proof of a combinatorial result, one has to show that the probability of a certain event is positive. If we have n mutually independent events A_i, each of which hold with probability $p > 0$, then the probability that they all hold simultaneously is at least p^n, which is positive, but may be exponentially small in n.

It is natural to expect that something similar will be true if the events are not entirely independent, but only 'mostly independent', for some sensible definition of 'mostly independent'.

Definition (Dependency digraph)**.** Let $A_1, A_2, \ldots, A_n$ be events in an arbitrary probability space. A directed graph $D = ([n], E)$is called a *dependency digraph* for the events $A_1, A_2, \ldots, A_n$ if for all i the event A_i is mutually independent of all the events $\{A_j : (i,j) \notin D\}$.

- For an event A_i let us write $\overline{A_i}$ for the negation of this event.

Lemma 9.1 (The Lovás Local Lemma)**.** *Let $A_1, A_2, \ldots, A_n$ be events in an arbitrary probability space. Suppose that $D = ([n], E)$ is a dependency digraph for the events $\{A_i\}_{i=1}^n$ and there exists $x_1, x_2, \ldots x_n \in [0,1)$ such that*

$$\mathbb{P}(A_i) \leq x_i \prod_{(i,j)\in E} (1 - x_j)$$

for all $i \in [n]$. Then

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \overline{A_i}\right) \geq \prod_{i=1}^{n} (1 - x_i).$$

In particular, with positive probability no event A_i holds.

Corollary 9.2. *[Symmetric Local Lemma] Let $A_1, A_2, \ldots, A_n$ be events in an arbitrary probability space. Suppose that each event A_i is mutually independent of a set of all but at most d of the other A_j (equivalently there is a dependency digraph with all outdegrees less than d), and that $\mathbb{P}(A_i) \leq p$ for all i. If $ep(d+1) \leq 1$ then*

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \overline{A_i}\right) > 0.$$

9.2 Ramsey Bounds for the last time

Theorem 9.3. *If*

$$e\binom{k}{2}\binom{n-2}{k-2}2^{1-\binom{k}{2}} < 1$$

then $R(k,k) > n$.

- Can use the local lemma to show

$$R(k,3) \geq c\frac{k^2}{(\log(k))^2}.$$

- A similar argument shows that

$$R(k,4) \geq k^{\frac{5}{2}+o(1)}.$$

9.3 Directed Cycles

Theorem 9.4. *Let $D = (V, E)$ be a directed graph with minimum outdegree δ and maximum indegree Δ. Then for any $k \in \mathbb{N}$ such that*

$$k \leq \frac{\delta}{1 + \log(1 + \delta\Delta)},$$

D contains a directed cycle of length divisible by k.

9.4 The Linear Aboricity of Graphs

Definition (Arboricity, linear forest, linear arboricity)**.** Given a graph G the *aboricity* of G, $\mathrm{a}(G)$, is the minimum number of forests into which the edge set $E(G)$ can be partitioned. A *linear forest* is a forest in which every component is a path, and the *linear aboricity* of a graph, $\mathrm{la}(G)$, is the minimum number of linear forests into which the edge set $E(G)$ can be partitioned.

Conjecture 9.5 (The Linear Aboricity Conjecture)**.** *Let G be a d-regular graph. Then*

$$la(G) = \left\lceil \frac{d+1}{2} \right\rceil.$$

Definition (Regular digraph, linear directed forest, dilinear arboricity)**.** A *d-regular digragph* is a directed graph in which the indegree and outdegree of every vertex is precisely d. A *linear directed forest* is a directed graph in which every connected component is a directed path and the *dilinear aboricity* of a directed graph D, which we denote by $\mathrm{dla}(D)$, is the minimum number of linear directed forests into which the edge set $E(G)$ can be partitioned.

Conjecture 9.6. *Let D be a d-regular digraph. Then*

$$dla(D) = d + 1.$$

Lemma 9.7. *Let $H = (V, E)$ be a graph with maximum degree Δ, and let $V = V_1 \cup V_2 \cup \ldots \cup V_r$ be a partition of V into r pairwise disjoint sets. Suppose that $|V_i| \geq 2e\Delta$ for each $i \in [r]$. Then there is an independent set $W \subset V$ that contains a vertex from each V_i.*

- The *directed girth* of a graph is the minimum length of a directed cycle in that graph.

Theorem 9.8. *Let $D = (V, E)$ be a d-regular directed graph with directed girth $g \geq 8ed$. Then*

$$dla(D) = d + 1.$$

Lemma 9.9. *Let $D = (V, E)$ be a d-regular directed graph, where d is sufficiently large, and let p be an integer such that $10\sqrt{d} \leq p \leq 20\sqrt{d}$. Then there is a p-colouring of V, $f : V \to [p]$, such that, for each $v \in V$ and each $i \in [p]$ the number*

$$N^+(v, i) = |\{u \in V \, : \, (v, u) \in E \text{ and } f(u) = i\}|$$

and

$$N^-(v, i) = |\{u \in V \, : \, (u, v) \in E \text{ and } f(u) = i\}|$$

satisfy

$$\left| N^+(v, i) - \frac{d}{p} \right|, \left| N^-(v, i) - \frac{d}{p} \right| \leq 3\sqrt{\frac{d}{p} \log(d)}.$$

Theorem 9.10. *There exists a constant $c > 0$ such that for every d-regular digraph D*

$$dla(D) \leq d + cd^{\frac{3}{4}} (\log(d))^{\frac{1}{2}}.$$

Corollary 9.11. *There exists a constant $c > 0$ such that for every d-regular graph G*

$$la(G) \leq \frac{d}{2} + cd^{\frac{3}{4}} (\log(d))^{\frac{1}{2}}.$$

10 Martingales and Strong Concentration

10.1 The Azuma-Hoeffding Inequality

The strong concentration bounds we obtained in Section 8 only applied to sums of random variables which were independent. This is quite a strong condition to ask for, and in this section we will prove strong concentration results of a similar nature that do not need to rely on independence. S

Definition (Martingale)**.** Let $Z_1, Z_2, \ldots, Z_n$ and $X_0, X_1, \ldots, X_n$ be sequences of random variables on the same probability space such that X_i is determined by $\{Z_1, Z_2, \ldots, Z_i\}$ and, for all i,

$$\mathbb{E}(X_i \,|\, Z_1, Z_2, \ldots, Z_{i-1}) = X_{i-1}.$$

Then (X_i) is called a *martingale* with respect to (Z_i).

Lemma 10.1 (Doob Martingale)**.** *Let A and (Z_i) be random variables on the same probability space. The $X_i = \mathbb{E}(A \,|\, Z_1, Z_2, \ldots, Z_i)$ is a martingale with respect to (Z_i).*

Proof. Note firstly that X_i is determined by $\{Z_1, Z_2, \ldots, Z_i\}$. Also, for all i we have that

$$\mathbb{E}(X_i \,|\, Z_1, Z_2, \ldots, Z_{i-1}) = \mathbb{E}(\mathbb{E}(A \,|\, Z_1, Z_2, \ldots, Z_i) \,|\, Z_1, Z_2, \ldots, Z_{i-1}).$$

However it is clear that the above expectation is, for given $Z_1, Z_2, \ldots, Z_{i-1}$, averaging over all possible values of Z_i the expected value of $(A|Z_1, Z_2, \ldots, Z_i)$. Hence

$$\mathbb{E}(X_i \,|\, Z_1, Z_2, \ldots, Z_{i-1}) = \mathbb{E}(A \,|\, Z_1, Z_2, \ldots, Z_{i-1}) = X_{i-1}.$$

□

Definition. Consider the probability space $\mathcal{G}(n,p)$ and order the set of potential edges $(i,j) \in [n]^{(2)}$ arbitrarily as $e_1, e_2, \ldots, e_m$, where $m = \binom{n}{2}$. Let Z_i to be the indicator function of the event that e_i is an edge in $G(n,p)$. Any graph theoretic function f is some function of the random variables (Z_i). Given such a function f, the *edge exposure martingale* is the martingale $X_i = \mathbb{E}(f \,|\, Z_1, Z_2, \ldots, Z_i)$ with respect to (Z_i). Note that $X_0 = \mathbb{E}(f)$ and $X_n = f$.

Let $Z_i' \in \{0,1\}^{i-1}$ be the vector of indicator functions of whether the edge between the vertex i and $j < i$ is in $G(n,p)$. Again for any graph theoretic function f, the *vertex exposure martingale* is the martingale $X_i = \mathbb{E}(f \,|\, Z_1', Z_2', \ldots, Z_i')$ with respect to (Z_i'). In this case $X_1 = \mathbb{E}(f)$ and $X_n = f$.

Theorem 10.2 (The Azuma-Hoeffding Inequality)**.** *Let $c_1, \ldots, c_n > 0$ and let $(X_i)_0^n$ be a martingale with respect to $(Z_i)_1^n$ such $|X_i - X_{i-1}| \leq c_i$ for all $1 \leq i \leq n$ then*

$$\mathbb{P}(X_n \geq X_0 + t) \leq e^{-\frac{t^2}{2\sigma^2}} \text{ and } \mathbb{P}(X_n \leq X_0 - t) \leq e^{-\frac{t^2}{2\sigma^2}}$$

where $\sigma^2 = \sum_{i=1}^n c_i^2$.

For the proof of the above we need the following technical lemma.

Lemma 10.3. *Let Y be a random variable which takes values in $[-1,+1]$ such the $\mathbb{E}(Y)=0$. Then for any $t \geq 0$*

$$\mathbb{E}(e^{tY}) \leq e^{\frac{t^2}{2}}.$$

Definition. A graph theoretic functions f is *edge Lipschitz* if whenever H and H' differ in only one edge then $|f(H)-f(H')| \leq 1$. Equivalently, if we consider f as a function of the variables $f(Z_1, Z_2, \ldots, Z_n)$ then we require that changing one coordinate does not change f by more than 1. Similarly it is *vertex Lipschitz* if whenever H and H' differ at only one vertex $|f(H)-f(H')| \leq 1$.

Lemma 10.4. *For any graph theoretic function f, if f is edge Lipschitz then the corresponding edge exposure martingale satisfies $|X_i - X_{i-1}| \leq 1$ and similarly if f is vertex Lipschtiz.*

10.2 The Chromatic Number of a Dense Random Graph

- The chromatic number $\chi(G)$ is a vertex Lipschitz function.

Theorem 10.5. *For any n and p and for all $t \geq 0$*

$$\mathbb{P}(|\chi(G(n,p)) - \mathbb{E}(\chi(G(n,p))| \geq t) \leq e^{-\frac{t^2}{2(n-1)}}.$$

- If we take $t = \omega(\sqrt{n})$ we see that $\chi(G(n,p))$ is 'tightly' concentrated about its expectation, although we still don't know what it's expectation is.
- Let us just consider the case $p = 1/2$ for ease of presentation, the arguments are similar for all fixed p.
- Note $\chi(G) \geq \frac{n}{\alpha(G)}$.
- **Without proof** Using the second moment method it is possible to show that, for any $\epsilon < 0$ almost surely $\alpha(G(n,1/2))$ is between $(2-\epsilon)\log_2(n)$ and $2\log_2(n)$.
- Hence almost surely
$$\chi(G(n,1/2)) \geq (1+o(1))\frac{n}{2\log_2(n)}.$$

Definition ($f(k)$ and Y)**.** Let $f(k)$

$$f(k) = \binom{n}{k} 2^{-\binom{k}{2}}$$

and let k_0 be such that $f(k_0-1) > 1 > f(k_0)$ and let $k_1 = k_0 - 4$. Let Y be the maximal size of a family of edge disjoint cliques of size k_1 in $G(n,1/2)$ (that is, the number of cliques it contains, not the union of their sizes). '

Note that G will have no k_1-clique if and only if $Y = 0$ and that Y is edge Lipschitz.

Lemma 10.6.

$$\mathbb{E}(Y) \geq (1+o(1))\frac{n^2}{4k_1^4}.$$

Lemma 10.7.

$$\mathbb{P}(\omega(G(n,1/2)) < k_1) = e^{-n^{2+o(1)}}.$$

Theorem 10.8. *Almost surely*

$$\chi(G(n,1/2)) \leq (1+o(1))\frac{n}{2\log_2(n)}.$$

10.3 The Chromatic Number of Sparse Random Graphs

Theorem 10.9. *Let $p = n^{-\alpha}$ for some fixed $\alpha > 5/6$ then there exists some $u = u(n,p)$ such that almost surely*

$$u \leq \chi(G(n,p)) \leq u+3.$$

We use the following result in the proof.

Lemma 10.10. *Let $p = n^{-\alpha}$ for some fixed $\alpha > 5/6$ and let $c > 0$. Then almost always it is true that, for every subset $S \subset [n]$ with $|S| = c\sqrt{n}$*

$$\chi(G(n,p)|_S) \leq 3.$$

11 Talagrand's Inequality

Our previous results on strong concentration gave us an exponentially small bound on the probability of deviations from the mean of suitably well behaved random variables. Our notion of suitably well behaved was basically that it was close, or bounded by, some random variable on a product space where in each co-ordinate the random variable was bounded. However these results required that the deviations we considered be at least as large as the square root of the dimension of this product space. Talagrands inequality gives us a much stronger result for a similar class of random variables.

Definition (Product space)**.** Let $\{(\Omega_i, 2^{\Omega_i}, \mathbb{P}_i) : i \in [n]\}$ be finite probability spaces. We let

$$\Omega = \prod_{i=1}^{n} \Omega_i = \{(\omega_1, \omega_2, \ldots, \omega_i) : \omega_i \in \Omega_i \text{ for all } i \in [n]\}$$

be the product of the sets Ω_i and define a probability measure $\mathbb{P}$ on 2^Ω by defining the probability of elementary events to be

$$\mathbb{P}((\omega_1, \omega_2, \ldots, \omega_i)) = \prod_{i=1}^{n} \mathbb{P}_i(\omega_i)$$

and extending it to 2^Ω in the obvious way. Then the *product space* (of $\{(\Omega_i, \Sigma_i, \mathbb{P}_i) : i \in [n]\}$) is the probability space $(\Omega, 2^\Omega, \mathbb{P})$.

- For example $\mathcal{G}(n,p)$ is the product of $\binom{n}{2}$ identical probability spaces, each of which corresponds to a possible edge of $G(n,p)$.

- Given a random variable on a product space we can consider the *co-ordinate exposure martingale.*

Definition. Given a unit vector $\alpha \in \mathbb{R}^n$ and two points $\omega, \omega' \in \Omega$ the *α-Hamming distance* between ω and ω' is

$$d_\alpha(\omega, \omega') = \sum_{\omega_i \neq \omega_i'} \alpha_i.$$

Given a set $A \subset \Omega$ and a point ω for any α we can consider the α-Hamming distance between ω and A.

$$d_\alpha(\omega, A) = \inf\{d(\omega, \omega') \,:\, \omega' \in A\}.$$

We will think of ω as being far from A if it's far in *some* α-Hamming distance, with α a unit vector. That is, we define

$$d(\omega, A) = \sup_{|\alpha|=1} d_\alpha(\omega, A).$$

Theorem 11.1 (Talagrand's Inequality)**.** *Let $\{(\Omega_i, \Sigma_i, \mathbb{P}_i) \,:\, i \in [n]\}$ be probability spaces and let $(\Omega, \Sigma, \mathbb{P})$ be their product. If $A, B \in \Sigma$ are such that $d(\omega, A) \geq \tau$ for all $\omega \in B$, then*

$$\mathbb{P}(A)\mathbb{P}(B) \leq e^{-\frac{\tau^2}{4}}.$$

Definition. A random variable $X : \Omega \to \mathbb{R}$ is *c-Lipschitz* if changing just one co-ordinate can change the value of X by at most c. Given some function $f : \mathbb{N} \to \mathbb{N}$ we say that X *is f-certifiable* if whenever $X(\omega_1, \omega_2, \ldots, \omega_n) \geq s$ there is a subset $I \subset [n]$ of size $|I| = f(s)$ such that X is greater than s on the entire subspace

$$\{(\omega_1', \omega_2', \ldots, \omega_n') \,:\, \omega_i' = \omega_i \text{ for all } i \in I\}.$$

Corollary 11.2. *Let X be a c-Lipschitz random variable which is f-certifiable and let m be the median of X (that is m is the unique real number such that $\mathbb{P}(X > m) \leq 1/2$ and $\mathbb{P}(X < m) \leq 1/2$). Then for any $t \geq 0$*

$$\mathbb{P}(X \leq m - t) \leq 2e^{-\frac{t^2}{4c^2 f(m)}} \text{ and } \mathbb{P}(X \geq m + t) \leq 2e^{-\frac{t^2}{4c^2 f(m+t)}}.$$

- Note that the two tail estimates are not necessarily symmetric.
- The important thing is that gives us strong concentration in a way that does not depend on the dimension of the product space Ω.
- This theorem talks about a variable being concentrated about its median rather than its mean, however in a lot of cases one can show the median must be close to the expectation.

Lemma 11.3. *Let $\{(\Omega_i, 2^{\Omega_i}, \mathbb{P}_i) \,:\, i \in [n]\}$ be probability spaces and let $(\Omega, 2^\Omega, \mathbb{P})$ be their product. Let X be a c-Lipschitz, $f(s) = rs$-certifiable random variable and let m be the median of X. Then*

$$|\mathbb{E}(X) - m| \leq 20c\sqrt{rm}.$$

11.1 Longest Increasing Subsequence

- Suppose we pick a permutation of $[n]$ uniformly at random and consider the random variable X which counts the longest increasing subsequence.

- If we let Y be the longest decreasing subsequence, then a theorem of Erdős and Szekeres says that $XY \geq n$.
- If we let m be the median of X then $m = \Omega(\sqrt{n})$.
- By applying Talagrand's inequality one can show that, with high probability

$$X \sim m \pm n^{1/4} \log(n).$$

- The Azuma-Hoeffding inequality would only give concentration about the mean for deviations $t >> \sqrt{n}$.
- It can be shown that $\mathbb{E}(X) = \Omega(\sqrt{n})$.

11.2 Chromatic Number of Graph Powers

Definition (Powers of a graph)**.** Given a graph G the *kth power of* G, G^k is defined to be the graph with

$$V(G^k) = V(G) \text{ and } E(G^k) = \{(x, y) : \text{dist}_G(x, y) \leq k\}.$$

Theorem 11.4 (**Without proof**)**.** *Let H be such that $g(H) \geq 5$, then $\chi(g) \leq (1+o(1))\Delta(H)/\log(\Delta(H))$*

Theorem 11.5 (**Without proof**)**.** *Let H be such that $g(H) \geq 4$ (that is, triangle-free), then $\chi(g) \leq O\big(\Delta(H)/\log(\Delta(H))\big)$*

- If G is d-regular then $\chi(G) \leq d + 1$, and $\chi(G^k) \leq d^k + 1$.
- If $g(G) >> k$ then the above theorems give a better upper bound for $\chi(G^k)$.

Theorem 11.6. *Let $g \geq 3$ and k be fixed. Then for large enough d there exist graphs G with $g(G) \geq g$ and $\Delta(G) \leq d$ such that*

$$\chi(G^k) \geq \Omega\left(\frac{d^k}{\log(d)}\right).$$

We use the following lemma in the proof

Lemma 11.7. *Let $G(n, p)$ be chosen with $p = \frac{d}{2n}$. Then for an appropriate choice of constant c_k the following holds: For every subset $U \subset V(G)$ of size*

$$|U| = c_k n \frac{\log(d)}{d^k} = x$$

let P be the random variable which counts the maximum size of a family of paths of length k which lie in $G(n, p)$ such that both endpoints lie in U, all the internal vertices of the paths lie outside of U, and no two paths share a vertex except in U. Then almost surely

$$P \geq \frac{c_k^2 n \log(d)^2}{2^{k+7} d^k}.$$

11.3 Exceptional outcomes

Our previous concentration results have all relied on our random variables being c-Lipschitz for some small enough c, when considered as functions on an underlying product probability space. Whilst they can't tell us anything if the function is not well behaved in this way, sometimes it will be the case that the random variable is still be tightly concentrated.

- Consider the random variable T which counts the number of triangles in $G(n,p)$.
- $r(n) = 1/n$ is a threshold function for the event that $G(n,p)$ contains a triangle.
- Suppose $p = n^{\beta-1}$ for some small $\beta > 0$. In this case the expected number of triangles will get arbitrarily large.
- Since T is not 'very Lipschitz', our previous concentration results cannot guarantee that T is large with high probability if $\beta \leq 2/3$.
- However, whilst changing an edge can change T a lot, the expected change in T from any particular edge is still quite low.
- As the following example shows, this is not quite enough to guarantee concentration.

Example. Let $m = 4k$, we will consider a probability space on $\{0,1\}^m$ where each event $\omega = (\omega_1, \omega_2, \ldots, \omega_m)$ is such that the probability that $\omega_i = 1$ is, independently, $p = m^{-\frac{1}{2}}$ for each i. Consider the following function

$$f(t_1, t_2, \ldots, t_m) = (t_1t_2 + t_2t_3 + \ldots + t_{2k-1}t_{2k})(t_{2k+1} + t_{2k+2} + \ldots + t_{4k}).$$

- $\mathbb{E}(f) = \frac{1}{8}m^{\frac{1}{2}}$.
- $|\mathbb{E}(f|t_i = 0) - \mathbb{E}(f|t_i = 1)| = \frac{1}{2}$ for $i \leq 2k$.
- $|\mathbb{E}(f|t_i = 0) - \mathbb{E}(f|t_i = 1)| = \frac{1}{4}$ for $i \geq 2k$.
- With high probability $f(t)$ is either 0, or larger than $4\mathbb{E}(f)$.

- A stronger notion than being 'quite Lipschitz on average' would be that the function is 'quite Lipschitz' 'almost always, that is, there is a set of events of small probability such that everywhere else f is c-Lipschitz for some small c.
- In order to make this precise we will change our notion of certifiability slightly.

Definition. Given an exceptional set $\Omega^* \subset \Omega$ and $s, c > 0$ we say that a random variable X has (s,c)-*certificates* if for every $t > 0$ and every $\omega \in \Omega \setminus \Omega^*$ there is an index set I of size at most s so that $X(\omega') > X(\omega) - t$ for any $\omega' \in \Omega \setminus \Omega^*$ for which the ω and ω' differ in less than t/c co-ordinates.

- Note that, if X is f-certifiable and c-Lipschitz and s is the maximum value of f over the range of X, then X has (s,c)-certificates.

Theorem 11.8. *Let $\{(\Omega_i, 2^{\Omega_i}, \mathbb{P}_i) : i \in [n]\}$ be probability spaces and let $(\Omega, 2^{\Omega}, \mathbb{P})$ be their product. Let $\Omega^* \subseteq \Omega$ be a set of exceptional events. Suppose X is a random variable which has (s, c)-certificates, let m be the median of X and let $t \geq 0$. Then*

$$\mathbb{P}(|X - m| \geq t) \leq 4e^{-\frac{t^2}{4c^2 s}} + 4\mathbb{P}(\Omega^*).$$

Lemma 11.9. *Let $\{(\Omega_i, 2^{\Omega_i}, \mathbb{P}_i) : i \in [n]\}$ be probability spaces and let $(\Omega, 2^{\Omega}, \mathbb{P})$ be their product. Let $\Omega^* \subseteq \Omega$ be a set of exceptional events. Let X be a random variable which has (s, c)-certificates, let m be the median of X and let $M = \max\{ \sup |X|, 1\}$. Then*

$$|\mathbb{E}(X) - m| \leq 20c\sqrt{s} + 20M^2\mathbb{P}(\Omega^*).$$

Corollary 11.10. *Let $\{(\Omega_i, 2^{\Omega_i}, \mathbb{P}_i) : i \in [n]\}$ be probability spaces and let $(\Omega, 2^{\Omega}, \mathbb{P})$ be their product. Let $\Omega^* \subseteq \Omega$ be a set of exceptional events. Let X be a random variable which has (s, c)-certificates, let m be the median of X and let $M = \max\{ \sup |X|, 1\}$. If $\mathbb{P}(\Omega^*) \leq M^{-2}$ then for $t > 50c\sqrt{s}$*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq 4e^{-\frac{t^2}{16c^2 s}} + 4\mathbb{P}(\Omega^*).$$

- One can use these results to deduce that T is concentrated about its mean for β arbitrarily small.
- If we let $\beta = 2/3 + \epsilon$ for some small ϵ then we can use the Azuma-Hoeffding inequality to show that the number of triangles is tightly concentrated about its mean, and hence the probability that there are more that $n^{3\beta}$ triangles is exponentially small.
- Similarly using Chernoff's bound we can say the probability that any (potential) edge extends to more than $\max\{2np^2, n^\delta\}$ triangles is exponentially small.
- Since both these random variables are montone, this also holds for smaller p.
- This allows use to apply Corollary 11.10 with $s = n^{3\beta}$ and $c = \max\{2np^2, n^\delta\}$ to $\beta = 4/9 + \epsilon'$ and deduce concentration.
- However, we can now bootstrap this result for smaller p, using the previous step to get a better bound on s and c.

Theorem 11.11. *Let $p = n^{-1+\beta}$ for $\beta > 0$ and let $\delta > 0$. Then with high probability*

$$|T - \mathbb{E}(T)| \leq n^\delta \sqrt{\mathbb{E}(T)}.$$

12 Entropy Methods

12.1 Basic Results

Definition. Given a discrete random variable X let us denote by $p(x) := \mathbb{P}(X = x)$ for each x in the range of X. The *entropy* of X is

$$H(X) = \sum_x p(x) \log\left(\frac{1}{p(x)}\right).$$

- $H(X) \geq 0$.

Lemma 12.1. *Let X be a discrete random variable and let R be the range of X.*

$$H(X) \leq \log(|R|).$$

Definition. Given two discrete random variables, X and Y, we define the *joint entropy* (X, Y) to be

$$H(X, Y) = \sum_x \sum_y p(x, y) \log\left(\frac{1}{p(x, y)}\right),$$

where, as before, $p(x, y) := \mathbb{P}(X = y, Y = y)$. We also define the *conditional entropy*, of Y given X, to be

$$H(Y|X) = \sum_x p(x) H(Y|X = x) = \mathbb{E}_x\big(H(Y|X)\big).$$

Lemma 12.2. *Let X and Y be discrete random variables. Then*

$$H(Y|X) \leq H(Y).$$

Lemma 12.3 (Chain rule)**.** *Let X and Y be discrete random variables. Then*

$$H(X, Y) = H(X) + H(Y|X).$$

- If we define the joint entropy of a sequence of discrete random variables $X_1, X_2, \ldots, X_n$ in a similar way then

$$H(X_1, X_2, \ldots, X_n) = H(X_1) + H(X_2|X_1) + \ldots H(X_n|X_1, X_2, \ldots, X_{n-1}).$$

- We call this the *chain rule*.
- If the random variable Y completely determines the random variable Z then $H(X|Y) = H(X|Y, Z)$.

12.2 Brégman's Theorem

- For a graph G let us write $\Phi(G)$ for the set of perfect matchings of G and $\phi(G) = |\Phi(G)|$.

Theorem 12.4 (Brégman's Theorem)**.** *Let G be a bipartite graph on vertex classes A and B such that $|A| = |B| = n$. Then*

$$\phi(G) \leq \prod_{v \in A} (d(v)!)^{\frac{1}{d(v)}}.$$

12.3 Shearer's lemma and the Box theorem

- Given a sequence of discrete random variables random variables $X_1, X_2, \ldots, X_n$ and some subset $A \subseteq [n]$ let us define $X_A := (X_i : i \in A)$.

Lemma 12.5 (Shearer's inequality)**.** *Let $X_1, X_2, \ldots, X_n$ be discrete random variables and $\mathcal{A}$ a collection (not necessarily distinct) of subsets of $[n]$, such that each $i \in [n]$ is in at least m members of $\mathcal{A}$. Then*

$$H(X_1, X_2, \ldots, X_n) \leq \frac{1}{m} \sum_{A \in \mathcal{A}} H(X_A).$$

- If we take a shape $S \subset \mathbb{Z}^d$ and pick a point uniformly at random inside of S this gives a vector $X = (X_1, \ldots, X_n)$ of random variables.
- $H(X) = \log |S|$ and for $A \subset [n]$ we have the range of X_1 is the 'volume' of the $(n - |A|)$-dimensional projection onto the subspace where $x_i = 0$ for $i \in A$.
- Let us call this projection S_A and we note that $H(X_A) \leq \log |S_A|$.

Theorem 12.6 (The Loomis-Whitney inequality)**.** *Let $S \subset \mathbb{Z}^n$ then,*

$$|S|^{n-1} \leq \prod_{i=1}^{n} |S_{[n] \setminus \{i\}}|$$

- This theorem is tight when $|S|$ is a 'box', that is, a set of the form $[1, m_1] \times [1, m_2] \times \ldots \times [1, m_n]$.

Definition. We say a collection of sets $\mathcal{C} = \{C_1, \ldots, C_m\} \subset 2^{[n]}$ is a *k-uniform cover* if each $i \in [n]$ belongs to exactly k many of the C_j.

Theorem 12.7 (Uniform covers theorem)**.** *Let $S \subset \mathbb{Z}^n$ and let $\mathcal{C} \subset 2^{[n]}$ be a k-uniform cover, then*

$$|S|^k \leq \prod_{C \in \mathcal{C}} |S_C|$$

- $\mathcal{C} = \{[n] \setminus \{i\} : i \in [n]\}$ is an $(n-1)$-uniform cover of $[n]$, and so Theorem 12.6 follows from Theorem 12.7.
- By approximating a shape in $\mathbb{R}^n$ with finer and finer grids one can show that Theorem 12.7 still holds for any 'reasonable' (say, measurable) shape $S \subset \mathbb{R}^n$, where $|.|$ now denotes the normal volume.

Theorem 12.8 (Bollobás-Thomason Box Theorem)**.** *Let $S \subset \mathbb{R}^n$. Then there is a box $B \subset \mathbb{R}^n$ such that $|B| = |S|$ and $|B_A| \leq |S_A|$ for all $A \subseteq [n]$.*

Definition. Let $\mathcal{C}$ be a uniform cover of $[n]$ we say $\mathcal{C}$ is *irreducible* if we cannot write it as the disjoint union $\mathcal{C} = \mathcal{C}' \cup \mathcal{C}''$ of two uniform covers.

Lemma 12.9. *There are only finitely many irreducible uniform covers of $[n]$*

12.4 Independent Sets in a Regular Bipartite Graph

- For a graph G let us write $\mathcal{I}(G)$ for the set of independent subsets of $V(G)$.

Theorem 12.10. *Let G be a d-regular bipartite graph on $2n$ vertices with vertex classes A and B, and let $\mathcal{I}(G)$ be the class of independent subsets of $V(G)$. Then*

$$|\mathcal{I}(G)| \leq (2^{d+1} - 1)^{\frac{n}{d}}$$

12.5 Bipartite Double Cover

Definition (Bipartite double cover)**.** Given any graph G we define the *bipartite double cover* of G to be the cartesian product, $G \times K_2$.

Definition. We say that A is *independent from* B if there are no edges between A and B and we define the *size* of a pair of subsets (A, B) to be $|A| + |B|$.

- For a graph G let us write $\mathcal{J}(G)$ for the set of (A, B) such that A is independent from B and $G|_{A \cup B}$ is bipartite.

Lemma 12.11. *For any graph G, there exists a size preserving bijection between $\mathcal{I}(G) \times \mathcal{I}(G)$ and $\mathcal{J}(G)$.*

Theorem 12.12. *Let G be a d-regular graph on n vertices, and let $\mathcal{I}(G)$ be the class of independent subsets of $V(G)$. Then*

$$|\mathcal{I}(G)| \leq (2^{d+1} - 1)^{\frac{n}{2d}}$$

13 Derandomization and Combinatorial Games

13.1 Maximum Cuts in Graphs

Theorem 13.1. *Any graph G, with $e(G) = m$, contains a bipartite subgraph with at least $m/2$ edges.*

While the theorem asserts the existence of a large cut, it gives no indication of how to find one deterministically. However in many cases, including this one in particular, we can "derandomize" such an argument to produce a fully deterministic algorithm.

Example (Derandomising MAXCUT)**.**

13.2 Ramsey graphs

Example (Building a Ramsey graph)**.**

13.3 Positional Games

Definition (Strong positional game, Red-win, Blue-win, draw)**.** A *strong positional game* consists of a pair $(X, \mathcal{F})$ where X is a set, called the *board*, and $\mathcal{F} \subset 2^X$ is a family of *winning lines*. The game is played by two players, sometimes referred to as Red and Blue, who take turns claiming points of the board (with Red going first), which we may think of as colouring some point $x \in X$ as either red or blue. Given a particular play of the game, that is a sequence of moves $(r_1, b_2, r_3, b_4 \ldots)$, the winner is the first player to claim all points in some winning set $F \in \mathcal{F}$. If at no point during the game either player achieves this, the game is a draw.

If Red has a strategy to win a game $(X, \mathcal{F})$ we call the game *Red-win*, and similarly for Blue. If both players have a drawing strategy we call the game a draw.

Lemma 13.2. *If X is finite then all strong positional games $(X, \mathcal{F})$ are either Red-win, Blue-win, or a draw.*

Theorem 13.3 (Strategy stealing)**.** *Let $(X, \mathcal{F})$ be a strong positional game with X finite. Then $(X, \mathcal{F})$ is either a Red-win or a draw.*

- Consider $(X, \mathcal{F})$ as a hypergraph. If *every* 2-colouring of X contains a monochromatic $F \in \mathcal{F}$, then $(X, \mathcal{F})$ must be Red-win.

Lemma 13.4. *Suppose $(X, \mathcal{F})$ is a hypergraph. If*

$$\sum_{F \in \mathcal{F}} 2^{-|F|} < 1/2$$

then there exists a 2-colouring of X containing no monochromatic $F \in \mathcal{F}$.

We can derandomize this argument to give a winning strategy for Blue.

Theorem 13.5. *[The Erdős-Selfridge Theorem] Suppose $(X, \mathcal{F})$ is a strong positional game. If*

$$\sum_{F \in \mathcal{F}} 2^{-|F|} < 1/2$$

then $(X, \mathcal{F})$ is a draw.

Definition. The *Ramsey Game* is the game $(E(K_n), \mathcal{F})$ where he board is the edge set of a complete graph K_n and $\mathcal{F}$ is the collection of edge sets of complete subgraphs K_k.

- If $n \geq R(k, k)$ the Ramsey game is Red-win.
- If
$$\binom{n}{k} 2^{-\binom{k}{2}} < 1/2$$
the game is a draw.
- It follows that we can find a *balanced* colouring of K_n with no monochromatic K_k.

Definition. The n-in-a-row game is the game $(\mathbb{Z}^2, \mathcal{F})$ where the board is $\mathbb{Z}^2$, and $\mathcal{F}$ is the collection of all consecutive lines of n points in a row, either horizontally, vertically or diagonally.

- For $n \leq 4$ one can check by hand that the game is Red-win.

Theorem 13.6. *For $n \geq 40$ the n-in-a-row game is a draw.*

- One can show with a *pairing strategy* that the 8-in-a-row game is a draw.
- It is unknown who wins for $n = 5, 6, 7$.

13.4 Weak Games

Definition. Given a strong positional game $(X, \mathcal{F})$ the corresponding *weak positional game*, or *Maker-Breaker game*, $MB(X, \mathcal{F})$ is played as follows. The two players, Maker and Breaker, take turns claiming points of the board X, with Maker going first. We will still sometimes think of Maker as colouring his points red, and Breaker blue. Maker wins if at some point in the game he can claim all the points in some winning set $F \in \mathcal{F}$, and Breaker wins otherwise. If Maker has a winning strategy we call the game *Maker-win*, and similarly if Breaker has a winning strategy.

- Every Maker-Breaker game is either Maker-win or Breaker-win.
- If $MB(X, \mathcal{F})$ is Breaker-win, then $(X, \mathcal{F})$ is a draw.
- If $(X, \mathcal{F})$ is Red-win, then $MB(X, \mathcal{F})$ is Maker-win.
- Neither converse is true.

- Given $MB(X, \mathcal{F})$ let us write

$$\Delta_2(\mathcal{F}) = \max_{x,y \in X} |\{F \in \mathcal{F} : x, y \in F\}|.$$

Theorem 13.7. *Suppose $MB(X, \mathcal{F})$ is a weak positional game, with $\mathcal{F}$ n-uniform. If*

$$|\mathcal{F}| > 2^{n-3}|X|\Delta_2(\mathcal{F})$$

then $MB(X, \mathcal{F})$ is Maker-win.

13.5 The Neighbourhood Conjecture

Lemma 13.8 (Pairing strategy)**.** *Let $(X, \mathcal{F})$ be a positional game. Suppose that for every $\mathcal{G} \subset \mathcal{F}$*

$$\left| \bigcup_{G \in \mathcal{G}} G \right| \geq 2|\mathcal{G}|$$

(That is, the total number of points in X contained in some member of $\mathcal{G}$ is at least twice as large as the number of sets in $\mathcal{G}$). Then $(X, \mathcal{F})$ is a draw.

Corollary 13.9. *Let $(X, \mathcal{F})$ be a positional game in which every winning set has size $\geq n$. If every point $x \in X$ is in at most $n/2$ winning sets, then the game is a draw.*

Lemma 13.10. *Suppose $H = (X, \mathcal{F})$ is an n-regular hypergraph in which every vertex is in at most $2^{n-2}/n$ edges, then there exits a 2-colouring of X with no monochromatic edge.*

- Let denote by $f(n)$ the smallest number such that the following is true: Every positional game $(X, \mathcal{F})$ in which every winning set has size n and each point in is at most $f(n)$ winning sets is a draw.
- Corollary 13.9 tell us that $f(n) \geq n/2$.

Conjecture 13.11 (The Neighbourhood Conjecture)**.**

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

- The best current bounds are that $n/2 \leq f(n) \leq 2^{n-1}/n$

14 The Algorithmic Local Lemma

Suppose we have a probability space Ω which comes with an underlying set of mutually independent random variable $Z_1, \ldots, Z_n$. We are considering some set of events $(A_i \in i \in I)$, as in the local lemma, where each A_i is determined by some subset $\text{vbl}(A_i) \subset \{Z_1, \ldots, Z_n\}$ of the Z_j.

In this case, a very natural dependency graph to take is that graph where $i \sim j$ if and only if $\text{vbl}(A_i) \cap \text{vbl}(A_j) \neq \emptyset$, when the events A_i and A_j depend on a common variable. Let use denote by D_i the neighbourhood of i in this graph.

Given an *assignment* of values to the variable Z_j, that is, some $\gamma = (\gamma_1, \ldots, \gamma_n)$ where $\gamma_j \in \text{range}(Z_j)$ for each j, we say that A_i is violated by this assignment if A_i is true under this assignment.

Our aim will be, given that the events $(A_i,\ i \in I)$ satisfy the conditions of the local lemma, to algorithmically find an assignment of values such that no event A_i is violated. The algorithm can be stated extremely simply:

- Pick a random assignment for each Z_j independently
- while there exists a violated A_i
  - Pick a violated A_i (according to some deterministic rule)
  - Re-sample the assignments for each $Z_j \in \text{vbl}(A_i)$
- return the values of the Z_j

It is clear that if this algorithm terminates, then we have found our desired requirement.

Theorem 14.1. *Let $Z_1, \ldots, Z_n$ and $(A_i : i \in I)$ be as above. If there exists real numbers $0 < x_i < 1$ such that*

$$\mathbb{P}(A_i) \leq x_i \prod_{j \in D_i} (1 - x_j)$$

for each i then the algorithm finds an assignment of values to the Z_j such that no event A_i is violated in expected time at most

$$\sum_{i \in I} \frac{x_i}{1 - x_i}$$

Definition. The *log* of the algorithm is a sequence $L = (L(1), L(2), L(3) \ldots)$ where $L(t)$ is the event A_i such that $\text{vbl}(A_i)$ was re-sample in the tth step of the algorithm. We note that the log may be an infinite sequence.

A *witness tree* is a rooted tree T whose vertices are labelled with events A_i such that if A_j is a child of A_i, then $j \in D_i \cup \{i\}$. We call T *proper* if at each vertex the set of labels on its children is distinct.

Given a log L we define a witness tree $T(t)$ for each step of the algorithm recursively. We first label the root with the event $L(t)$. Then, for each $i = t-1, t-2, \ldots$ we consider the event $L(i) = A_k$. If there exists a vertex in the tree which is labelled with an event in $D_k \cup \{k\}$ then we pick one furthest from the root (breaking ties arbitrarily) and we add a leaf to the tree behind this vertex labelled A_k. If no such vertex exists then we go on to the next $L(i-1)$.

We say a witness tree T *occurs* in L if there is some t such that $T = T(t)$.

Lemma 14.2. *Let T be a witness tree and L the log file of a random execution of the algorithm.*

- *If T occurs in L then T is proper;*
- $\mathbb{P}(T \text{ occurs in } L) \leq \prod_{v \in T} \mathbb{P}(A_{[v]})$.

- For an event A_i let N_i be the number of times that $\mathrm{vbl}(A_i)$ is re-sampled in the algorithm.
- Note, given L, N_i is precisely the number of times that the root of the witness tree $T(t)$ is labelled A_i.
- By counting over *all* proper witness trees with root A_i

$$\begin{aligned}\mathbb{E}(N_i) &= \sum_{T\,:\,\mathrm{root}=A_i} \mathbb{P}(T\text{occurs in L}) \\ &\leq \sum_{T\,:\,\mathrm{root}=A_i} \prod_{v\in V(T)} \mathbb{P}(A_{[v]}) \\ &\leq \sum_{T\,:\,\mathrm{root}=A_i} \prod_{v\in V(T)} x_{[v]} \prod_{j\in D_{[v]}} x_j\end{aligned}$$

We will evaluate this sum using another random process called a *Galton-Watson branching process.*

In this process we build a labelled tree by first picking a root with a label A_i and then, for each $j \in D_i \cup \{i\}$ we add with probability x_j a child of the root with label A_j. We then do the same for each child of the root according to the same rule, and so on. This process may die out eventually, or may produce an infinite tree.

Lemma 14.3. *Let T be a proper witness tree with root A_i. The probability that T is given by the above Galton-Watson branching process is*

$$p_T = \frac{1-x_i}{x_i} \prod_{v\in V(T)} x'_{[v]}$$

where $x'_i = x_i \prod_{j\in D_i}(1-x_j)$.

So putting this all together we see that

$$\mathbb{E}(N_i) \leq \sum_{T\,:\,\mathrm{root}=A_i} \prod_{v\in V(T)} x'_{[v]} \leq \sum_{T\,:\,\mathrm{root}=A_i} p_T \leq \frac{x_i}{1-x_i}.$$

- This gives us a randomised algorithm that has a low expected running time, but it can be changed into a deterministic algorithm using the method of conditional expectations as in the previous section.