# The Probabilistic Method in Combinatorics 80721

Based on lectures by Dr. Yuval Peled, and the book by Alon and Spencer - *The probabilistic method*

**Notes by Asaf Etgar**
Spring 2022

Last edit: June 9, 2022

For any corrections, requests for additions or notes - please email me at
`asafetgar@gmail.com`

*These notes have not been revised by the course staff, and some things may appear differently than in the lectures/ recitations.*

# Contents

# Chapter 1

# Introcuction

## 1.1 Ramsey Numbers

**Claim 1.1.1.** *For any graph $G = (V, E)$ there exists a partitioning of $V = A \sqcup B$ such that at least half of the edges are $A - B$ edges.*

*Proof.* Consider a random partition of $V$, $A, B$. That is, each vertex $v$ is in $A$ or in $B$ w.p $\frac{1}{2}$ independently. Then:

$$\mathbb{E}\left[e(A,B)\right] \overset{linearity}{=} |E| \cdot \Pr\left[e \text{ is an } A, B \text{ edge}\right] = \frac{|E|}{2}$$

Which implies that there exists. a partition with said property. $\square$

*Remark.* One can prove this claim without the use of probability.

There are questions that we do not know yet how to solve without the use of probability:

**Definition 1.1** (Ramsey Number)**.** The number $R(k,l)$ is the minimal $n$ such that every graph $G$ over $n$ vertices contains a $k$-clique or an $l$ -anti-clique.

**Theorem 1.1** (Ramsey)**.** *$R(k,l) \leq \binom{k-l-2}{k-1}$. In particular, $R(k,k) \leq \binom{2k-2}{k-1} \approx \frac{4^{k-1}}{\sqrt{\pi k}}$*

**Theorem 1.2.** *If $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$*

*Proof.* Consider now a random graph $G \sim G(n, \frac{1}{2})$. For any $A \in \binom{[n]}{k}$, denote by $M_A$ the event that $A$ is a clique or anti-clique in $G$. Then:

$$\Pr\left[M_A\right] = \Pr\left[A \text{ is a clique}\right] + \Pr\left[A \text{ is an anti-clique}\right] = 2^{1-\binom{k}{2}}$$

And therefore

$$\Pr\left[\exists \text{ a clique or anti-clique of size } k\right] = \Pr\left[\bigcup_{A \in \binom{[n]}{k}} M_A\right] \leq \sum_{A \in \binom{[n]}{k}} \Pr\left[M_A\right] = \binom{n}{k} 2^{1-\binom{k}{2}} < 1$$

Hence there exists a graph over $n$ vertices without a clique or anti-clique of size $k$. $\square$

*Remark.* Note that

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

And also, $\binom{n}{k} \leq \frac{n^k}{k!}$, and by Stirling's approximation - $k! \geq \left(\frac{k}{e}\right)^k$. Pluggin in the inequality:

$$\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$$

Comparing this to the formula in 1.1, this is a very loose bound (at least for $k, k$).

*Remark.* If $n = 2^{k/2}$, then:

$$\Pr\left[\text{A random graph contains a clique or anti-clique}\right] \leq \binom{n}{k} 2^{1-\binom{k}{2}} \xrightarrow{n,k} 0$$

Which means "almost all graphs are Ramsey graphs", but we do not yet have any explicit construction.

This theorem implies that we know the existence of a graph of order $n$ and without clique or anti-clique of size $k \approx 2\log n$. The best construction known without the use of probability is for $k = \log^{C\cdot\log\log\log n} n$.

# Chapter 2

# Linearity of Expectation

## 2.1 Sum-Free Sets

**Theorem 2.1.** *For any $B \in \binom{\mathbb{N}}{n}$ (with repetitions), there exists $A \in \binom{B}{n/3}$ such that there are no $a, b, c \in A$ with $a + b = c$*

*Proof (by Erdős).* Denote $[x] = x - \lfloor x \rfloor$, and for any $t \in [0, 1]$ let $A_t = \left\{ b \in B \mid [tb] \in \left(\frac{1}{3}, \frac{2}{3}\right) \right\}$. For any $t$, $A_t$ is sum-free: If $a, b \in A_t$ and $[ta], [tb] \in \left(\frac{1}{3}, \frac{2}{3}\right)$, then $[a + b] \notin \left(\frac{1}{3}, \frac{2}{3}\right)$. We consider the probability space of the coin tosses of $t$. Denote $X_i = \mathbb{1}_{b_i \in A_t}$, then $\Pr[X_i = 1] = \frac{1}{3}$. Hence consider the expectation of a size of a random $A_t$:

$$\mathbb{E}[|A_t|] = \sum_{i \in [n]} \mathbb{E}[X_i] = \frac{n}{3}$$

□

*Remark.* The general idea of probabilistic methods is to find an object in which the property always holds, and then average over these objects

## 2.2 Tournaments

**Definition 2.1.** A tournament is an orientation of $K_n$.

**Definition 2.2.** We say a vertex $v$ *overcomes* some $A \subset V \setminus x$ if $v \to x$ for any $x \in A$ (that is, the orientation of $vx \in E(K_n)$ is $v \to x$).

**Theorem 2.2.** *If $\binom{n}{k}\left(1 - 2^{-k}\right)^{n-k}$, then there exists a tournament such that for any $A \in \binom{V}{k}$ there exists $v$ that overcomes $A$.*

*Proof.* Denote by $S_k$ the event that for any $A$ of size $k$ there exists an overcoming $v$. Consider a random tournament, and let $A \in \binom{V}{k}$, what is the probability that no $v$ overcomes $A$?

$$\Pr[\text{No } v \text{ overcomes } A] = \left(1 - 2^{-k}\right)^{n-k}$$

(some $v$ overcomes $A$ w.p $2^k$, and they are independent) Then:

$$\Pr[S_k^c] \leq \binom{n}{k}\left(1 - 2^{-k}\right)^{n-k} < 1$$

□

*Remark.* The union bound is quite similar to linearity of expectation.

**Theorem 2.3.** *There exists a tournament with at least $n! \cdot 2^{-(n-1)}$ Hamiltonian cycles.*

*Proof.* Consider a random tournament. Then:

$$\mathbb{E}\left[\# \text{ of Hamiltonian cycles}\right] = \sum_{\pi \in S_n} \Pr\left[\pi(V) \text{is a cycle}\right] = n! 2^{-(n-1)}$$

(the last equation is the probability of this permutation defininig a cycle) Then there must exist a tournament with at least this number of cycles. ☐

## 2.3 ??? If you have a suggestion for a name, let me know!

**Theorem 2.4.** *Let $v_1, \ldots v_n \in \mathbb{R}^d$ be unit vectors. Then there exists $\varepsilon_1, \ldots \varepsilon_n \in \{\pm 1\}$ such that*

$$\left\| \sum_{i \in [n]} \varepsilon_i v_i \right\| \leq \sqrt{n}$$

*and there exists such $\varepsilon_i$ for the opposite inequality.*

*Proof.* Consider a random choice of $\varepsilon_i$. Denote $X = \left\| \sum_{i \in [n]} \varepsilon_i v_i \right\|^2$. Then:

$$X = \left\| \sum_{i \in [n]} \varepsilon_i v_i^2 \right\| = \sum_{i \in [n]} \varepsilon_i^2 v_i + 2 \cdot \sum_{i < j} \varepsilon_i \varepsilon_j \left\langle v_i, v_j \right\rangle$$

then

$$\mathbb{E}\left[X\right] = n + 2 \cdot \sum_{i < j} \left\langle v_i, v_j \right\rangle \mathbb{E}\left[\varepsilon_i \varepsilon_j\right] = n$$

Since $\mathbb{E}\left[\varepsilon_i \varepsilon_j\right] = \mathbb{E}\left[\varepsilon_j\right] \mathbb{E}\left[\varepsilon_j\right] = 0 \cdot 0 = 0$, and the claim follows as usual. ☐

### 2.3.1 Derandomization

We would like to de-randomize the process and find an efficient algorithm of finding these $\varepsilon_i$. By the law of total expectation, $\mathbb{E}\left[X\right] = \frac{1}{2}\mathbb{E}\left[X \mid \varepsilon_1 = 1\right] + \frac{1}{2}\mathbb{E}\left[X \mid \varepsilon_1 = -1\right]$.

**Claim 2.3.1.** *If we've fixed $\varepsilon_1 \ldots \varepsilon_{i-1}$ such that $\mathbb{E}\left[X \mid \varepsilon_1 \ldots \varepsilon_{i-1}\right] \leq n$, then we can efficiently find $\varepsilon_i$ such that $\mathbb{E}\left[X \mid \varepsilon_1 \ldots \varepsilon_{i-1}, \varepsilon_i\right] \leq n$*

*Proof.*

$$\mathbb{E}\left[X \mid \varepsilon_1 \ldots \varepsilon_{i-1}\right] \stackrel{\star}{=} \frac{1}{2}\mathbb{E}\left[X \mid \varepsilon_1 \ldots \varepsilon_{i-1}, \varepsilon_i = 1\right] + \frac{1}{2}\mathbb{E}\left[X \mid \varepsilon_1 \ldots \varepsilon_{i-1}, \varepsilon_i = -1\right]$$

with $\star$ by law of total expectation (w.r.t the random variable $\varepsilon_i$). But

$$\mathbb{E}\left[X \mid \varepsilon_1 \ldots \varepsilon_{i-1}, \varepsilon_i = 1\right] = n + 2 \sum_{j' < j \leq i} \varepsilon_j \varepsilon_{j'} \left\langle v_j, v_i \right\rangle + 0$$

And we know the values of $\varepsilon_{j'}, \varepsilon_j$, so we can compute it efficiently. Then we choose the epsilon TodoCompete from the book ☐

## 2.4 Turan's theorem

**Theorem 2.5.** *In any graph $(V, E)$, there exists an independent set of size at least $\sum_{v\in V} \frac{1}{deg(v)+1}$*

*Proof.* Consider a random ordering of $V$. We choose a vertex to add to the set $I$ ("independent") if he appears before all of his neighbors. Clearly $I$ is independent. And:

$$\mathbb{E}\left[|I|\right] = \sum_{v\in V} \Pr\left[v \in I\right] = \sum_{v\in V} \Pr\left[v \text{ is the first of hie neighbors in the ordering}\right] = \sum_{v\in V} \frac{1}{deg(v)+1}$$

□

**Corollary 2.6.** *In $G$ there exists a clique of size $\geq \sum_{v\in V} \frac{1}{n-deg(v)}$*

**Theorem 2.7** (Turán)**.** *If the maximal clique is of size $r$, then*

$$r \geq \sum_{v\in V} \frac{1}{n-\deg(v)} \geq \frac{n^2}{n^2 - 2|E|}$$

*Therefore $|E| \leq \left(1 - \frac{1}{n}\right) \cdot \frac{n^2}{2}$*

## 2.5 Unbalancing Lights

Let $A$ be an $n \times n$ matrix over $\{\pm 1\}$. There is a switch for every row and every column, which flips all bits corresponding to it.

**Theorem 2.8.** *There exists $x, y \in \{\pm 1\}^n$* [I] *such that $x^\top Ay \geq \left(\sqrt{\frac{2}{\pi}} + o(1)\right) n^{\frac{3}{2}}$*

*Proof.* Choose a random $y$ (that is, $y_i \sim U(\pm 1)$ iid. Let $R_i = \sum_{j=1}^{n} A_j^i y_j$. Since $y_i$ are iid, $A_j^i y_j \sim$ a sum of $n$ signs$\pm 1$ iid. Then by CLT:

$$\frac{1}{\sqrt{n}} R_i \overset{distribution}{\longrightarrow} \mathcal{N}(0,1)$$

And therefore $\mathbb{E}\left[\frac{1}{\sqrt{n}}|R_i|\right] \overset{n\to\infty}{\longrightarrow} \mathbb{E}\left[|z|\right] = \sqrt{\frac{2}{\pi}}$. Hence:

$$\mathbb{E}\left[\sum_{i=1}^{n} |R_i|\right] = \sum_{i=1}^{n} \mathbb{E}\left[|R_i|\right] = \left(\sqrt{\frac{2}{\pi}} + o(1)\right) n^{\frac{3}{2}}$$

Then there exists $y$ such that $\sum_{i=1}^{n} |R_i| \geq \left(\sqrt{\frac{2}{\pi}} + o(1)\right) n^{\frac{3}{2}}$. As for $x$, note that

$$x^\top Ay = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i A_j^i y_j = \sum_{i=1}^{n} x_i R_i \overset{\star}{=} \sum_{i=1}^{n} |R_i| \geq \left(\sqrt{\frac{2}{\pi}} + o(1)\right) n^{\frac{3}{2}}$$

with $\star$ since we can take $x_i = sign(R_i)$.[II] □

## 2.6 2-colorings of hypergraphs

[I] think of $x$ as responsible of rows, and $y$ of columns

[II] In some sense, this is "the smartest move" in order to get as many light bulbs lit as possible.

**Definition 2.3** ($k$-uniform Hypergraph)**.** $H = (V, E)$ is a $k$-uniform Hypergraph with $V(H)$ its vertices and $E(H) \subset \binom{V(H)}{k}$. In particular, a 2-uniform hypergraph is just a graph.

**Definition 2.4** (2-coloring of hypergraph)**.** Let $H$ be a $k$-graph. A *2 coloring* of $H$ is a function $f : V(H) \to \{0, 1\}$ such that there is no monochromatic edge, that is $\forall e \in E(H) \quad \exists x, y \in e \quad f(x) \neq f(y)$.

**Definition 2.5.** Denote $m(k)$ the minimal number of edges in a $k$-graph that is not 2 colorable.

**Example 2.1.** $m(2) = 3$, consider a triangle.

**Example 2.2.** $m(3) = 7$, consider the Fano Plane.

**Theorem 2.9.** $m(k) \geq 2^{k-1}$

*Proof.* Let $H$ be a hypergraph with less than $m2^{k-1}$ edges. Let $f$ be a uniformly random coloring of $H$ Then

$$\Pr[e \text{ is monochromatic}] = 2^{1-k}$$

Therefore

$$\mathbb{E}[\#\text{monochromatic edges}] < m \cdot 2^{1-k} = 1$$

□

**Theorem 2.10.** $m(k) = O(k^2 2^k)$.

*Proof.* Let $n = k^2$, and choose $c \cdot k^2 2^k$ (we specify $c$ later) edges uniformly IID. We show that $\mathbb{E}[\#\text{colorings without monochromatic edges}] < 1$. Fix a coloring $\varphi : [n] \to \{0, 1\}$. Let $a = |\varphi^{-1}(0)|$, then

$$\Pr[e \text{ is monochromatic under } \varphi] =$$

$$\frac{\binom{a}{k} + \binom{n-a}{k}}{\binom{n}{k}} \geq \frac{2 \cdot \binom{n/2}{k}}{\binom{n}{k}} \geq \frac{2 \cdot \frac{(\frac{n}{2} - k)^k}{k!}}{\frac{n^k}{k!}} = 2\left(\frac{1}{2} - \frac{k}{n}\right)^k = 2 \cdot \left(\frac{1}{2}\right)^k \left(1 - \frac{2}{k}\right)^k \geq c\left(\frac{1}{2}\right)^k$$

for some $c$. Now we have

$$\mathbb{E}[\#\text{colorings without monochromatic edges}] = \sum_{\varphi} \Pr[\text{no edge is monochromatic under } \varphi] \leq$$

$$\leq 2^n \left(1 - \frac{c}{2^k}\right)^m \leq e^{\log(2) \cdot n - c \cdot 2^{-k} \cdot m} \overset{\star}{<} 1$$

With $\star$ by choice of $m = \frac{2\log(2)}{c} k^2 \cdot 2^k$. □

**Theorem 2.11** (Improvement on the bound from 2.9)**.** $m(k) \geq t2^k \sqrt{\frac{k}{\log(k)}}$.

*Proof.* Or proof is algorithmic: Let $H$ be with $t$ edges, color all $V(H)$ in blue. Traverse $V(H)$ in a random order, let $v$ the current vertex. If $v$ is the last-visited vertex in a monochromatic blue edge - alter its color to blue. The algorithm fails only if there is a monochromatic red edge. This happens only if the vertex $v = e \cap f$, $f$ is red, $e$ is blue and $v$ is the first in $f$ and last in blue (this is a bad configuration). What is the probability of such configuration to occure? We consider the probability over the coin tosses of $\pi \sim U(S_n)$, and claim

$$\Pr[\text{there exists a bad configuration}] < 1$$

We note that:

$$\Pr\left[\text{there exists a bad configuration}\right] \leq \mathbb{E}\left[\#(e,f) \text{ are bad edges}\right] \overset{\star}{\leq} m^2 \cdot \frac{((k-1)!)^2}{(2k-1)!} = \frac{m^2}{(2k-1)\binom{2k-2}{k-1}} \overset{\star\star}{=} \frac{m^2 \cdot (c+o(1))}{\sqrt{k \cdot 4^k}} \overset{?}{<} 1$$

with $\star$ a bound on the number of edges that intersect in a unique vertex, times the probability of having a bad configuration, and $\star\star$ since $\binom{2n}{n} = \frac{c+o(1)}{\sqrt{n}} 2^{2n}$. In order to have ?, take $m < c' \cdot k^{\frac{1}{4}} \cdot 2^k$. This is not the bound we want - as the power of $k$ is $\frac{1}{4}$. The problem is the expectation sometimes lies - that is, the expectation can be much larger than the probability we want to bound (see the remark below).

We traverse the vertices differently: For a vertex $v$, choose $r_v \sim U([0,1])$ i.i.d, and traverse $V(H)$ according to $r_v$ from the smallest to largest. Let $p$ be some probability chosen later, and denote

$$L = \left[0, \frac{1-p}{2}\right] \quad M = \left[\frac{1-p}{2}, \frac{1+p}{2}\right] \quad R = \left[\frac{1+p}{2}, 0\right]$$

Now:

$$\begin{aligned}
&\Pr\left[\text{there exists a bad configuration}\right] \leq \\
&\overbrace{\Pr\left[\exists e \in L \cup R\right]}^{1} + \overbrace{\Pr\left[\exists \text{bad configuration whose intersection is in } M\right]}^{2} \leq \\
&\overbrace{m \cdot 2(|L|)^k}^{1} + \overbrace{m^2 \int_{\frac{1-p}{2}}^{\frac{1+p}{2}} r_v^{k-1}(1-r_v)^{k-1} dr_v}^{2} = \\
&m \cdot rbk \frac{1-p^k}{2} + m^2 \int_{\frac{1-p}{2}}^{\frac{1+p}{2}} r_v^{k-1}(1-r_v)^{k-1} dr_v \leq \\
&2m \frac{e^{-pk}}{2^k} + m^2 \cdot p \left(\frac{1}{4}\right)^{k-1} \overset{?}{<} 1
\end{aligned}$$

Choosing $p = \frac{\log k}{k}$ and $m < \frac{1}{4} 2^k \sqrt{\frac{k}{\log k}}$ yields the result. $\square$

*Remark.* Let $X_n = n^2$ with probability $1/n$ and 0 otherwise. Note that $\Pr\left[X_n > 0\right] = 1/n$, while $\mathbb{E}\left[X_n\right] = n$.

# Chapter 3

# Alterations Method

Up to this point, we made a random choice of object and use it. We now deal with the setting where a naïve random choice is not good enough - but we cat alter it a little bit so it would be good. The idea here is to bound the expectations of alterations needed to the random object.

## 3.1 Dominating Sets

**Definition 3.1.** Let $G$ be a graph. $A \subset V$ is *Dominating* if any $v \in V$ has a neighbor in $A$.

**Theorem 3.1.** *Let $G$ be of minimal degree $\delta$, then there exists a dominating set of size $n \cdot \frac{\ln(1+\delta}{1+\delta}$.*

*Proof.* Let $B \subset V$ such that any $v \in B$ with probability $p$ (will be chosen later) independently. Let $C_B$ be the collection of vertices that all of their neighbors are not in $B$, that is $C_B = \{x \notin B \mid \forall vx \in E \quad v \notin B\}$. Clearly $A = B \cup C_B$ is dominating. Then

$$\mathbb{E}[|A|] = \mathbb{E}[|B|] + \mathbb{E}[|C_B|] = np + n\Pr[v \in C_B] \overset{\star}{\leq} npne^{-p(1+\delta)}$$

With $\star$ since $\Pr[v \in C_B] = (1-p)^{1+deg(x)} \leq (1-p)^{1+\delta)} \leq e^{-p(1+\delta)}$. Find the optimal $p$ by differentiating w.r.t $p$, and get $p = \frac{\ln(1+\delta)}{1+\delta}$, then $\mathbb{E}[|A|] \leq n\left(\frac{\ln(1+\delta)+1}{1+\delta}\right)$ $\square$

## 3.2 Ramsey Numbers - Revisited

Recall that 1.2 gives us a lower bound on Ramsey numbers. We will use alterations to improve this lower bound.

**Theorem 3.2.** *For any $n, k$, $R(k,k) \geq n - \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$*

*Proof.* Consider $G \sim \mathcal{G}(n, \frac{1}{2})$. Note that $\mathbb{E}[\#\text{monochromatic sets of size } k] = \binom{n}{k}2^{1-\binom{k}{2}}$ as we've seen, therefore there exists a graph with at most this amount of monochromatic sets of size $k$, denote it $G$. Let $G'$ be the graph obtained from $G$ by removing a single vertex of any monochromatic set of size $k$. Then $|V(G)|$ is at least $n - \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$, and clearly in $G'$ there is no monochromatic set of size $k$. $\square$

**Corollary 3.3.** *$R(k,k) \geq n - \frac{e^n}{k}2^{1-\binom{k}{2}}$ by the Stirling-esque estimation done in chapter 1. The optimal $n$ is $\frac{2^{k/2} \cdot k}{e}$ which yields $R(k,k) \geq 2^{k/2}k \cdot \left(\frac{1+o(1)}{e}\right)$.*

## 3.3 Girth and coloring

Let $G = (V, E)$ be a graph.

**Definition 3.2** (Girth)**.** The *girth* of $G$ is the length of a minimal cycle in $G$.

*Remark.* In particular, if the girth is $\geq g$, then for any $v \in V$, its $g$-neighborhood looks like a tree.

**Definition 3.3** (Chromatic Number)**.** The *chromatic number* of $G$, denoted $\chi(G)$ is the minimal $k$ such that there exists a proper coloring $c : V \to [k]$ of $G$.

*Remark.* It is difficult to know what $\chi(G)$ is - it is NP-hard

**Definition 3.4** (Independence number)**.** The *Independence Number* of a graph $G$, denoted $\alpha(G)$, is the size of a largest independent set in $G$.

**Claim 3.3.1.** *If $T$ is a tree, then $\chi(T) = 2$*

*Proof.* It is bipartite - use BFS. □

**Theorem 3.4** (Erdős)**.** *For any $k, g$ there exists a graph $G$ with $\chi(G) \geq k$ and girth $\geq g$.*

*Remark.* This is surprising! Any neighborhood seems like $\chi$ should be small (as neighborhoods look like trees) - but it turns out it cannot be considered locally; $\chi$ is a *global* property of $G$.

For ease - we write $\alpha(G) = \alpha$, same for $\chi$.

**Lemma 3.4.1.** *$V(G) \leq \alpha \cdot \chi$.*

*Proof.* If $c : V \to [\chi]$ is a proper coloring, any $c^{-1}(i)$ is independent. □

**Lemma 3.4.2.** *There exists a graph $G$ over $n$ vertices (for a large enough $n = n(k, g)$) with the following properties:*

1. *The number of cycles of length $\leq g$ is smaller than $\frac{n}{2}$*
2. *$\alpha(G) \leq 3 \log n \cdot n^{1-\frac{1}{2g}}$*

*Proof.* Let $G \sim \mathcal{G}(n, p)$ with $p = n^{\frac{1}{2g}-1}$. Let $X$ be the number of cycles of length $\leq g$. Then:

$$\mathbb{E}[X] \stackrel{1}{=} \sum_{r=3}^{g} \binom{n}{r} \cdot \frac{(r-1)!}{2} \cdot p^r \stackrel{2}{\leq} \sum_{r=3}^{g} (n \cdot p)^r \stackrel{3}{\leq} g \cdot (n \cdot p)^g = g\sqrt{n}$$

Justifications:

1. Choose which vertices are in a cycle of length $r$ ($\binom{n}{r}$) and order them in a cycle ($(r-1)!/2$ options) and multiply by the probability of such cycle to exists.
2. Bound $\binom{n}{r} \cdot \frac{(r-1)!}{2}$ from above naturally.
3. Bound the sum with the largest element in the summation.

Hence by Markov:

$$\Pr[X > n/2] \leq \frac{g\sqrt{n}}{n/2} \xrightarrow{n \to \infty} 0$$

Which implies the first property. For the second property, let $t = 3 \log n \cdot n^{1-\frac{1}{2g}}$. Now:

$$\Pr[\alpha(G) \geq t] \leq \binom{n}{t} (1-p)^{\binom{t}{2}} \leq n^t \left(e^{-p}\right)^{\binom{t}{2}} \leq n^t e^{-p \cdot \binom{t}{2}} = e^{t(\log n - \frac{p \cdot t}{2} + 1)} = e^{t(-\frac{1}{2}\log n + 1)} \xrightarrow{n \to \infty} 0$$

□

*Proof (of 3.4).* Let $G'$ be a graph obtained from $G$ by removing a single vertex from any cycle of length smaller than $g$. Then $G'$'s girth is at least $g$. And $\alpha(G') \le \alpha(G) \le 3\log n \cdot n^{1-\frac{1}{2g}}$, and note that

$$\chi(G') \ge \frac{|V(G')|}{\alpha(G')} \ge \frac{\frac{n}{2}}{3\log n \cdot n^{1-\frac{1}{2g}}} \ge \frac{n^{\frac{1}{2g}}}{6\cdot\log(n)} \overset{n\to\infty}{\longrightarrow} \infty$$

$\square$

## 3.4 Heilbronn triangle problem

Let $P \subset [0,1]^2$, $|P| = n$, denote $T(P) = \min\limits_{x,y,z\in P} Area(xyz)$, and let $T(n) = \max\limits_{|P|=n} T(P)$. Heilbornn conjectured[I] that $T(n) = \Theta(\frac{1}{n^2})$.

**Theorem 3.5** (KPS, no proof). $T(n) = \Omega\left(\frac{\log(n)}{n^2}\right)$

*Remark.* We still do not know some $f$ for which $T(n) = \Theta(f(n))$, the best upper bound is still not tight.

**Theorem 3.6.** $T(n) \ge \frac{1}{70n^2}$

*Proof.* Let $\varepsilon = \frac{1}{70n^2}$. Generate $2n$ points $\sim U([0,1]^2)$ IID and remove a point from any triangle of area less than $\varepsilon$. Given a tringle $xyz$, Let $t$ be the distance $xy$. then $t$ has some density $f_{dist}(t)$. Then:

$$\Pr\left[area(xyz) \le \varepsilon\right] l \int_0^{\sqrt{2}} \sqrt{2}\cdot 4\frac{\varepsilon}{t} f_{dist}(t)dt = (\star)$$

Note that $f_{dist}(t) = \lim_{h\to 0}\frac{1}{h}\Pr\left[t \le dist(x,y)) \le t+h\right]$ by the definition of density. Hence $f_{dist(x,y)}(t) \le \lim_{h\to 0}\frac{1}{h}\pi((t+h)^2 - t^2)) = 2\pi t$, then:

$$(\star) \le \int_0^{\sqrt{2}} \sqrt{2}\frac{4\varepsilon}{t} 2\pi t dt = 16\pi\varepsilon$$

Which implies

$$\mathbb{E}\left[\text{number of triangles with area smaller than } \varepsilon\right] \le \binom{2n}{3}\frac{16\pi}{70n^2} < n$$

$\square$

*Remark.* Erdős has a non-combinatorial construction. Let $n$ be some prime, and consider the grid $[n-1]\times[n-1]$[II] and take $\left\{(k, k^2 \mod n)\right\}_{k\in[n-1]}$. Note that the smallest triangle of 3 points in $\mathbb{Z}^2$ is of area $1/2$, unless the three points are on the same diagonal. If they are on the diagonal $ax+b$, this means that there exists three values of $k$ such that $(ak+b) = k^2 \mod n$, but this is a quadratic polynomial in $\mathbb{F}_n[x]$, therefore it cannot have more than 2 solutions. Hence by scaling, $T(n) \ge \frac{1}{2(n-1)^2}$

[I] falsely

[II] Can rescale for the unit cube later...

# Chapter 4

# Second Moment Method

Up until now we discussed *first moment methods*. More formally, if $X = X_n \geq 0$ is an integer valued random variable, then the first moment method tells us that if $\mathbb{E}[X_n] \overset{n\to\infty}{\longrightarrow} 0$, then $\Pr[X_n > 0] \overset{n\to\infty}{\longrightarrow} 0$.

**Example 4.1** (First Moment Method)**.** When $G \sim \mathcal{G}(n,p)$ is triangle-free? Denote $X$ the number of traiangels in $G$. Then

$$\mathbb{E}[X] = \binom{n}{3} p^3 \leq (np^3)$$

Then taking $p = o\left(\frac{1}{n}\right)$ results in $\Pr[X > 0] \overset{n\to 0}{\longrightarrow} 0$. Is this bound *tight*? We saw that the expectation does not always give us a good bound - we need a way to reason about when is $X$ concentrated about its expectation - that is the variance.

**Definition 4.1** (Variance)**.** The variance of $X$ is $VarX = \mathbb{E}\left[(X - \mathbb{E}[x])^2\right]$

**Definition 4.2** (Covariance)**.** The covariance of $X, Y$ is

$$covX, Y = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

**Theorem 4.1** (Chebyshev)**.** $Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{VarX}{t^2}$

**Corollary 4.2.** $Pr[X = 0] \leq \frac{Var(X)}{\mathbb{E}[X]^2}$

**Corollary 4.3.** *If* $VarX = o\left(\mathbb{E}[X]^2\right)$ *then* $Pr[X = 0] \longrightarrow 0$

This results in the *Second moment method*: If $VarX = o\left(\mathbb{E}[X]^2\right)$ then $\Pr[X \geq 0] \overset{n\to\infty}{\longrightarrow} 0$. An equivalent condition is $\mathbb{E}[X^2] = \mathbb{E}[X]^2(1 + o(1))$[I] An important case is when $X = \sum_{i=1}^m X_i$, in that case

$$VarX = \sum_{i=1}^m VarX_i + \sum_{i=1}^m \sum_{j\neq i} covX_i, X_j$$

If we denote $i \sim j$ when $X_i, X_j$ are dependent, then

$$VarX = \sum_{i=1}^m VarX_i + \sum_{i=1}^m \sum_{j\sim i} covX_i, X_j$$

[I]Since $VarX = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$

**Assumptions:**

1. $X_i = \mathbb{1}_{A_i}$, then $VarX_i = \Pr[A_i] \cdot (1 - \Pr[A_i]) \le \Pr[A_i]$ and $CovX_i, X_j = \Pr[A_i \cap A_j] - \Pr[A_i] \cdot \Pr[A_j] \le \Pr[A_i \cap A_j] = \Pr[A_i] \cdot \Pr[A_j \mid A_i]$. Under this assumption, we get
$$VarX \le \mathbb{E}[X] + \sum_i \Pr[A_i] \cdot \sum_{i \sim j} \Pr[A_j \mid A_i]$$
2. A symmetry assumption is $\sum_{i\sim j} \Pr[A_j \mid A_i]$ is independent of $i$. This is usually true in many cases. We denote $\sum_{i\sim j} \Pr[A_j \mid A_i] = \Delta^*$[II]. With this notation, $VarX \le \mathbb{E}[\cdot](1 + \Delta^*)$.

   **Corollary 4.4.** *If* $\mathbb{E}[X] \longrightarrow \infty$, $\Delta^* = o(\mathbb{E}[X])$, *then* $Pr[X = 0] \longrightarrow 0$.

## 4.1 $H$-free graphs

The general question we deal with is

**Question 4.1.1.** *Given a small graph $H$, what is the threshold function of $G \sim \mathcal{G}(n, p)$ containing $H$?*

More formally:

**Definition 4.3** (Threshold Function)**.** Let $G \sim \mathcal{G}(n, p)$ and $H$ some fixed small graph. We say $f(n)$ is a *threshold function* for finding $H$ in $G$ if

$$p \ll f(n) \Rightarrow \Pr[G \text{ contains a copy of } H] \overset{n\to\infty}{\longrightarrow} 0$$
and
$$p \gg f(n) \Rightarrow \Pr[G \text{ contains a copy of } H] \overset{n\to\infty}{\longrightarrow} 1$$

We now explore some threshold functions

### 4.1.1 Triangles in $\mathcal{G}(n, p)$

We've shown that if $G \sim \mathcal{G}(n, p)$ and $p \ll \frac{1}{n}$ then $\Pr[G \text{ contains a triangle}] \to 0$.

**Claim 4.1.1.** *If $p \gg \frac{1}{n}$ then $Pr[G$ contains a triangle$] \to 1$*

*Remark.* In a case like this, we say that $\frac{1}{n}$ is the *threshold function* for triangle existence in $\mathcal{G}(n, p)$.

*Proof.* Denote by $X$ the number of triangles in $G$, then:

$$\mathbb{E}[X] = \binom{n}{3} p^3 = (1_o(1)) \frac{n^3 p^3}{6} \overset{p \gg \frac{1}{n}}{\longrightarrow} \infty$$

Denote by $T$ a triangle in $G$.

$$\Delta^* = \sum_{T' \sim T} \Pr[A_{T'} \mid A_T] \overset{\star}{=} 3 \cdot (n-3) p^2 \le 3np^2$$

[II]Usually $\sum_i \Pr[A_i] \cdot \sum_{i\sim j} \Pr[A_j \mid A_i] := \Delta$

With $\star$ since any $T'$ dependent on $T$ is taken by choosing 2 vertices in $T$ and a vertex not in $T$. Now we have

$$\frac{\Delta^*}{\mathbb{E}[X]} \leq \frac{18}{n^2 p} \longrightarrow 0$$

And we are done. $\square$

*Remark.* In fact we've shown that $\frac{X}{\mathbb{E}[X]} \overset{\text{in probability}}{\longrightarrow} 1$. This is some kind of law of large numbers. This can be seen by 4.1:

$$\Pr[|X - \mathbb{E}[X] > \varepsilon \mathbb{E}[X]] \leq \frac{VarX}{\varepsilon^2 \mathbb{E}[X]^2} \longrightarrow 0$$

### 4.1.2 $K_4$ in $\mathcal{G}(n,p)$

Denote $X$ the number of $K_4$ copies in $G$. Then

$$\mathbb{E}[X] = \binom{n}{4} p^6 = \frac{1+o(1)}{24} n^4 p^6$$

Then if $p \ll n^{-2/3}$, $\Pr[X=0] \longrightarrow 0$, and otherwise $\mathbb{E}[X] \to \infty$. Now denote by $S$ a fixed copy of $K_4$ in $G$. Then:

$$\Delta^* = \sum_{S' \sim S} \Pr[A_{S'} \mid A_S] \leq \overbrace{6n^2p^5}^{\text{Share 2 vertices}} + \overbrace{4np^3}^{\text{Share 3 vertices}}$$

Then

$$\frac{\Delta^*}{\mathbb{E}[X]} \leq O\left(\frac{1}{n^2 p} + \frac{1}{n^3 p^3}\right) \overset{p \gg n^{-2/3}}{\longrightarrow} 0$$

### 4.1.3 $K_4 * e$ in $\mathcal{G}(n,p)$

In this case

$$\mathbb{E}[X] = 5\binom{n}{5} p^7 = \frac{1+o(1)}{24} n^5 p^7$$

Then $p \ll n^{-5/7}$ implies $\Pr[X>0] \longrightarrow 0$. Is it true that $p \gg n^{-5/7}$ implies $\Pr[X=0] \longrightarrow 1$? Note that $n^{-5/7} \ll n^{-2/3}$, and since existence of $K_4 * e$ implies existence of $K_4$, had $n^{-5/7}$ been the threshold, it would contradict our previous proof.

**Definition 4.4** (Maximal Subgraph Density)**.** Given $H$, we define its maximal density by

$$m(H) := \max_{\emptyset \neq A \subset V(H)} \frac{e(A)}{|A|}$$

*Remark.* First moment argument shows that $p \ll n^{-\frac{1}{m(H)}}$.

**Theorem 4.5** (Threshold characterization)**.** *The threshold function of $H$ is $n^{-\frac{1}{m(H)}}$*

## 4.2 Cliques in $\mathcal{G}(n, 1/2)$

**Theorem 4.6.** *Let $G \sim \mathcal{G}(n, 1/2)$. Denote $X$ = size of maximal clique in $G$. There exists $k = k(n)$ and $k = \Theta(2\log_2(n))$ such that*

$$Pr[X \in \{k, k+1\}] \overset{n\to\infty}{\longrightarrow} 1$$

*Proof (Sketch).* Define $f(k) = \mathbb{E}[\#\text{of } k\text{-cliques in } G] = \binom{n}{k}2^{-\binom{k}{2}}$. We claim that if $f(k) \to \infty$ then there exists a clique of size $k$.

$$\Delta^* = \sum_{i=2}^{k-1} \binom{k}{i} \cdot \binom{n-k}{k-i} \left(\frac{1}{2}\right)^{\binom{k}{2}-\binom{i}{2}}$$

Now it can be shown that $k \sim 2\log_2(n$, if $f(k) \to \infty$ then $\frac{\Delta^*}{f(k)} \to 0$ by case analysis. We note that

$$\frac{f(k+1)}{f(k)} = \frac{\binom{n}{k+1}2^{-\frac{k(k+1)}{2}}}{\binom{n}{k}2^{-\frac{k(k-1)}{2}}} = \frac{n-k}{(k+1)}2^{-k} \approx 2^{-2\log_2(n)} < \frac{1}{n}$$

We now consider a $k_0$ such that $f(k_0) \geq 1$ and $f(k_0+1) < 1$. If $f(k_0(n))$ tends to $\infty$ with $n$ and $f(k_0) = o(n)$, then $f((k_0+1)(n)) \overset{n\to\infty}{\longrightarrow} 0$. In this case, there exists a maximal $k_0$-clique, and no $(k_0+1)$-clique, so the maximal clique of $k_0$. A similar argument may show the result, but it's quite annoying. The book has the complete proof □

## 4.3 Distinct Sums Problems

The *Distinct Sums Problems* is a problem suggested by Erdős[III].

**Problem.** *What is the maximal size of $S \subset [n]$ with distinct partial sums?*

**Solution** (lower bound)**.** Take $S = \{1, 2, 4, 8, \ldots\}$, then $|S| = \log_2(n)$.

**Question 4.3.1.** *Is $|S| \leq \log_2(n) + o(1)$?*

**Solution** (Upper bound)**.** $2^{|S|} \leq n \cdot |S|$

**Corollary 4.7.** $|S| \leq \log(n) + \log\log(n) + o(1)$

**Claim 4.3.1.** [IV] $|S| \leq \log(n) + 0.5\log\log(n) + o(1)$

*Proof.* Let $S = \{s_1 \ldots s_m\}$ and consider a random partial sum $X = \sum_{i=1}^m b_i s_i$ with $b_i \sim U(\{0,1\}$.

$$\mu := \mathbb{E}[X] = \sum_{i=1}^m \frac{s_i}{2} \quad \text{and} \quad VarX \sum_{i=1}^m \frac{s_i^2}{4}$$

Now try to bound the variation:

$$VarX = \frac{1}{4}\sum_{i=1}^m s_i^2 \overset{s_i \leq n}{\leq} \frac{mn^2}{4}$$

Then

$$\Pr[|x-\mu| \leq \lambda] \overset{Chebyshev}{\geq} 1 - \frac{mn^2}{4\lambda^2}$$
$$\Pr[|x-\mu| \leq \lambda] \overset{\text{distinct sums}}{\leq} (2\lambda+2)2^{-m}$$

[III] And it is still open with a \$300 prize awaiting to the solver!
[IV] This is worth \$150!

So

$$(2\lambda + 2)2^{-m} \geq 1 - \frac{mn^2}{4\lambda^2} \qquad \overset{\text{Take } \lambda=\sqrt{3}\sqrt{m}n}{\Longleftrightarrow}$$
$$c\sqrt{m}n2^{-m} \geq \frac{4}{c'mn^2} \geq \frac{11}{12}$$

Therefore $2^m \leq \tilde{C}\sqrt{m}n$ And some computations result in the bound. ∎

## 4.4 Hardy- Ramanujan Thoerem

The question we deal with is "how many prime numbers divide a random number $\sim U([n])$?

**Theorem 4.8** (Hardy- Ramanujan)**.** *For $x \in \mathbb{N}$ let $\nu(x)$ be the number of prime divisors of $x$ (without multiplicity). If $x \in U([n])$, then for a large enough $n$,*

$$\forall \varepsilon > 0 \quad \exists A > 0 \quad Pr\left[|\nu(x) - \log\log n| > A\sqrt{\log\log n}\right] < \varepsilon$$

**Theorem 4.9** (Erdős-Kac)**.** *With the same notations,*

$$\frac{\nu(x) - \log\log n}{\sqrt{\log\log n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

**Theorem 4.10** (Merten)**.** $\sum_{p\leq m} \frac{1}{p} = \log\log M = O(1)$

*Proof(of 4.8).* Denote by $X(x)$ the amount of prime divisors of $x$ not greater then $n^{1/10}$. Clearly $X \leq \nu$. Since there cannot be more than 10 divisors larger than $n^{1/10}$[V], we have $\nu - 10 \leq X$. Denote $\chi_p \mathbb{1}_{p|x}$. Then:

$$X = \sum_{p\leq n^{1/10}} \chi_p \quad \text{and} \quad \mathbb{E}\left[\chi_p\right] = \frac{\left\lfloor \frac{n}{p} \right\rfloor}{n} = \frac{1}{p} + O(1/n) \Rightarrow$$
$$\mathbb{E}\left[X\right] = \sum_{p\leq n^{1/10}} \frac{1}{p} + O(1) \overset{4.10}{=} \log\log n + O(1)$$

Now we calculate the $\Delta^*$ part. For any $p, q$ primes, we have $\Pr\left[\chi_p \cdot \chi_q\right] = \frac{\left\lfloor \frac{n}{pq} \right\rfloor}{n}$ (by the Chinese Reminder Theorem), therefore:

$$cov(\chi_p, \chi_q) = \mathbb{E}\left[\chi_p\chi_q\right] - \mathbb{E}\left[\chi_p\right]\mathbb{E}\left[\chi_p\right] = \frac{\left\lfloor \frac{n}{pq} \right\rfloor}{n} - \frac{\left\lfloor \frac{n}{p} \right\rfloor}{n}\frac{\left\lfloor \frac{n}{q} \right\rfloor}{n} = \frac{1}{pq} - \left(\frac{1}{p} - \frac{1}{n}\right)\left(\frac{1}{q} - \frac{1}{n}\right) \leq \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right)$$

So

$$VarX = \sum_p Var\chi_p + \sum_{p\neq q} cov\chi_p, \chi_q \leq \log\log n + O(1) + \sum_{p\neq q} \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right) \leq$$
$$\log\log n + O(1) + (n^{1/10})^2\frac{1}{n} = \log\log n + o(1)$$

So the result follows from Chebyshev. ∎

## 4.5 The Röde Nibble

The question deals with the existence of designs.

[V] Let $m$ be the number of divisors larger than $n^{1/10}$, then $x$ is at least $n^{m/10} \leq x \leq n$.

**Definition 4.5** ($(n, k, r)$ - design)**.** : An $(n, k, r)$ - design is a $k$-graph over $n$ vertices such that any $r$-set of vertices is contained exactly in one edge.

**Example 4.2.** $k = 2, r = 1$ means a perfect matching.

**Question 4.5.1.** *Is is true that for any $r < k \ll n$ there exists a corresponding design?*

Of course not! Take an odd $n$, one cannot find a perfect matching in a graph over an odd number of vertices. We get division requirements: Denote by $e$ the number of edges in a design. Then

$$e \cdot \binom{k}{r} = \binom{n}{r} \Rightarrow \binom{k}{r} \mid \binom{n}{r}$$

**Definition 4.6.** The Complementary design with respect to $A \subset [n]$ with $|A| < r$ is with the edges

$$E_A = \{e \setminus A \mid A \subset e \in E\}$$

### 4.5.1 Approximations

**Definition 4.7** (Covering)**.** a covering is a relaxation of designs, when we demand that any $r$-tuple is contained in *at least* one edge

**Definition 4.8** (Packing)**.** a covering is a relaxation of designs, when we demand that any $r$-tuple is contained in *at most* one edge

*Remark.* In theres cases, clearly $|E| \geq \frac{\binom{n}{r}}{\binom{k}{r}}$ and $|E| \leq \frac{\binom{n}{r}}{\binom{k}{r}}$ respectively.

The *Erdős Hananni conjecture* is that for any $k, r$, when $n \to \infty$ there exists a covering of size[VI] $(1 + o(1))\frac{\binom{n}{r}}{\binom{k}{r}}$. This is equivalent of having a packing of size $(1 - o(1))\frac{\binom{n}{r}}{\binom{k}{r}}$. This conjecture was proved by using the *Röde Nibble*.

*Proof (for the case $r = 2, k = 3$).* [VII] We look for a collection of $(1+o(1))\frac{\binom{n}{2}}{3}$ triplets that cover all edges. Had we tried to choose any triangle independently with probability $1/n$, we would have failed miserably:

$$\Pr\left[\text{A specific edge is not covered}\right] = (1 - 1/n)^{n-2} \approx \frac{1}{e}$$

Which means that this method "misses" a constant amount of edges!
We try to choose any trianegl with probability $\frac{\varepsilon}{n}$, which results in approximately $\frac{\varepsilon n^2}{6}$ triangles. Then:

$$\Pr\left[\text{A specific edge is not covered}\right] = (1 - 1/n)^{n-2} \approx \frac{1}{e^{\varepsilon}}$$

So

$$\Pr\left[\text{a specific edge is covered}\right] \approx 1 - e^{-\varepsilon} \approx \varepsilon - \frac{\varepsilon^2}{2} \dots$$

$\square$

[VI] In the sense of $|E|$
[VII] Steiner Triplet systems

**Definition 4.9** (Typical Graph)**.** A graph with $m$ edges is called $(D, \delta, k)$-*Typical* if:

1. Aside from $\delta \cdot m$ edges, all edge is contained in $D(1 \pm \delta)$ triangles.
2. Any edge is contained in at most $kD$ triangles.

**Lemma 4.10.1.** *For any $\varepsilon > 0$, large enough $D$, $k$ and $\delta > 0$, there exsists $\gamma > 0$ such that in any $(D, \delta, k)$ typical graph there is a collection of $\frac{\varepsilon}{3}(m \pm \gamma)$ triangles, denoted $T$ such that $G \setminus T$ is a graph with $m \cdot e^{-\varepsilon}(1 \pm \gamma)$ edges, and is $(De^{-2\varepsilon}, \gamma, ke^{2\varepsilon})$-typical*

*Proof.* We sample each triangle i.i.d with probability $\frac{\varepsilon}{D}$. The number of triangles in the graph is at least $\frac{(1-\delta)mD(1-\delta)}{3} = \frac{mD}{3}(1 - \delta_1)$, and at most $\frac{\delta mkD+(1-\delta)mD(1+\delta)}{3} = \frac{mD}{3}(1 + \delta_1)$. Let $T$ be the number of triangles. Then $T \sim \text{Bin}\left[\frac{mD}{3}(1 \pm \delta_1), \frac{\varepsilon}{D}\right]$ and the first item in the definition is gained by first moment argument.
Let $X_e$ be the indicator of the event "e is not covered". Then if $d_e = D(1 \pm \delta)$[VIII], we get

$$\mathbb{E}\left[X_e\right] = \left(1 - \frac{\varepsilon}{D}\right)^{D(1 \pm \delta)} = e^{-\varepsilon}(1 + \delta_1)$$

Let $X = \sum_e X_e$, then $\mathbb{E}\left[X\right] = me^{-\varepsilon}(1 - \delta_1)$ and

$$cov X_e, X_{e'} = \Pr\left[\text{both not covered}\right] - \Pr\left[e \text{ is not covered}\right] \Pr\left[e' \text{ is not covered}\right]$$
$$\left(1 - \frac{\varepsilon}{D}\right)^{d_e + d_{e'} - 1} - \left(1 - \frac{\varepsilon}{D}\right)^{d_e} \left(1 - \frac{\varepsilon}{D}\right)^{d_{e'}} \leq \frac{\varepsilon}{D}$$

Then

$$VarX \leq me^{-\varepsilon}(1 \pm \delta_1) + mD(1 + \delta)2\frac{\varepsilon}{D} = O(m)$$

Then by Chebyshev, $\Pr\left[\text{The number of edges} \notin me^{-\varepsilon}(1 \pm \delta_2) < 0.01\right] \to 0$. It is left to show that $G \setminus T$ is typical.

**Claim 4.5.1.** *Other than $\delta_1 m$ edges, all edges are both good and contained in $(1 \pm \delta_1)D$ triangles whose edges are good.*

*Proof.*

$$\mathbb{E}\left[d_e(G \setminus T)\right] = (1 \pm \delta_1)De^{-2\varepsilon}(1 \pm \delta_1)^2$$
$$Var d_e(G \setminus T) \leq \mathbb{E}\left[d_e(G \setminus T)\right] + D^2\frac{\varepsilon}{D} = O(D)$$

Wo once again, Chebyshev we are done. ☐

☐

*Proof (Röde's Nibble - general case).* Denote $p = e^{-\varepsilon}$Let $G_0 = K_n$. Let $G_{i+1}$ be obtained from $G_i$ by removing each triangle with probability $\frac{\varepsilon}{p^{2i}D}$ .Then $|E(G_i)| \approx p^i\binom{n}{2}$: With this step we've chosen

$$\approx \frac{\varepsilon}{p^{2i}D} \cdot \overbrace{p^i\binom{n}{2}}^{|E(G_i)|} \cdot \qquad \frac{\overbrace{p^{2i}n}^{\#\,\triangle \text{ in typical edge}}}{3} \qquad = \varepsilon p^i \frac{n^2}{6}$$

Hence the number of triangles in the cover is

$$p^t\binom{n}{2} + \sum_{i=0}^{t-1} \varepsilon p^i \frac{n^2}{6} \leq \frac{n^2}{6}\left(3e^{-\varepsilon t} + \varepsilon\frac{1}{1 - e^{-\varepsilon}}\right) = \star$$

So when $\varepsilon \to 0$, we can choose a large enough $t$, we may have $\star \leq (1 + \delta)\frac{n^2}{6}$. The thing is - we've hidden all the error terms, but this can be dealt with. Super annoyingly. ☐

[VIII] The number of triangles containing $e$

# Chapter 5

# Lovász's Local Lemma

Up to this point we used probability to find an object of interest with high probability. The *Local Lemma* is a tool to prove an object's existence even if the probability of finding them is small - even exponentially. In fact, this is an *algorithmic* approach.

**Theorem 5.1** (The local Lemma, Symmetric). *Let $(A_i)_{i\in[n]}$ be events such that:*

1. *$A_i$ is independent in all $A_j$, except for at most $d$ of them*[a]
2. *$Pr[A_i] \leq p$*
3. *$e \cdot p \cdot (d+1) \leq 1$*

*Then $Pr\left[\cap_{i\in[n]}\overline{A_i}\right] > 0$*

[a]Not in pairs! $A_i$ is dependent on the event $\cup_{j\in K} A_j$ for some $K \in \binom{[n]\setminus i}{d}$

## 5.1 Results from the lemma

**Theorem 5.2** (Improvement on 1.1). *If $e \cdot 2^{1-\binom{k}{2}}\left(\binom{k}{2}\binom{n}{k-2}+1\right) < 1$, then $R(k,k) > n$*

*Proof.* Let $G \sim \mathcal{G}(n, 1/2)$. For any $S \in \binom{[n]}{k}$, denote $A_S$ the even that $S$ is a clique or an anti-clique. We know that $\Pr[A_S] = 2^{1-\binom{k}{2}}$. Note that $A_S$ is independent in all $A_M$ other than at most $\binom{k}{2}\binom{n}{k-2}$. Then by the 5.1 - there exists a graph in which no $A_S$ occures. $\square$

**Exercise.** Consider a $k$-SAT in which any variable appears in at most $r$ clauses $(k > 3r)$. Show a polynomial algorithm to decide satisfyability.

**Theorem 5.3.** *Any $k$-graph (a $k$-uniform hypergraph) in which any edge intersects at most $\frac{2^{k-1}}{e} - 1$ other edges is 2-colorable.*

*Proof.* Consider a random 2 coloring $c : X \to \{0, 1\}$. Assume $A_i$ is the event "the $i$'th edge is monochromatic", then $\Pr[A_i] = 2^{1-k}$, and $d = \frac{2^{k-1}}{e} - 1$, and the result follows from LLL[I]. $\square$

### 5.1.1 Colorings of $\mathbb{R}$

Consider a coloring $c : \mathbb{R} \to [k]$. We say $T$ is *Colorful* if $c[T] = [k]$. The question Lovás and ??? asked is given a finite $S$, can we color $\mathbb{R}$ such that *$S$ and all of its translations* are colorful.

[I]Lováz's Local Lemma

**Theorem 5.4.** *For any $k$ and for any $S$ of cardinality $m$ such that*

$$e \cdot k \left(1 - \frac{1}{k}\right)^m (m(m-1)+1) \leq 1$$

*there exists a coloring $c : \mathbb{R} \to [k]$ such that all of $S$'s translations are colorful.*[II]

*Proof.* Denote $c_x = \{c \mid x + S \text{ is colorful}\}$. We want to show that $\bigcap_{x \in \mathbb{R}} c_x \neq \emptyset$. By compactness arguments, it is sufficient to show that for any finite $X$, $\bigcap_{x \in X} c_x \neq \emptyset$. Consider a random coloring $c : \mathbb{R} \to [k]$. For any $x \in X$, the event $A_x =$ "$x + S$ is not colorful", hence

$$\Pr[A_x] \leq k \left(1 - \frac{1}{k}\right)^m$$

And we note that $A_x, A_y$ are independent unless $(x+S) \cap (y+S) \neq \emptyset$, hence there are at most $m(m-1)$ such $y$'s for which it happens. From LLL we are done. □

### 5.1.2 Coverings of $\mathbb{R}^3$

**Definition 5.1** ($k$-covering)**.** A $k$ covering of a metric space $X$ is a covering in which any element is in at least $k$ covering sets.

**Definition 5.2** (Reducible)**.** We say a covering $\mathcal{U}$ is *reducible* if it can be partitioned into two coverings $\mathcal{U}_1, \mathcal{U}_2$ that are disjoint in their open sets.

**Question 5.1.1.** *Is there a $k$-covering in which any point is covered exactly $k$ times? $O(k$ times?*

**Theorem 5.5** (Nani-Levicko, Pach. No Proof)**.** *For any $k$ there exists an irreducible $k$-covering of $\mathbb{R}^3$ by unit balls.*

**Theorem 5.6.** *Any $k$ covering in which any point is covered at most $t := c \cdot 2^{\frac{k}{3}}$ times is reducible.*

*Proof.* Given a covering $\mathcal{U} = \{B_i\}_{i \in I}$, define a hypergraph $H$ with vertex set $\mathcal{U}$ and edges indexed by $\mathbb{R}^3$: for any $x \in \mathbb{R}^3$, $e_x = \{B_i \mid x \in B_i\}$ and delete multiple edges. That is, edges correspond to "cells" in $\mathcal{U}$. for any $x$, $k \leq |e_x| \leq t$. We need to show that $H$ is 2 colorable, which will correspond to two subcoverings. It suffices to show that any finite subgraph of $H$ is 2-colorable (by compactness). Consider a random 2-coloring $c$, denote by $A_x$ the event $e_x$ is monochromatic, then $\Pr[A_x] \leq 2^{1-k}$. If $A_x, A_y$ are dependent, then $d(x,y) \leq 4$: If $d(x,y) > 4$, any two balls containing $x, y$ do not intersect. We now claim that any edge $e_x$ intersects at most $c \cdot t^3$ other edges. By the previous claim - any ball intersecting some ball containing $x$ is contained in $B_4(x)$, and any point is covered at most $t$ times - thus the sum of volumes of balls intersecting some ball with $x$ is $N \leq 4^3 \cdot B_1 \approx 4^3$. The number of sells is hence at most $N^3 \overset{\text{exercise}}{\leq} 4^9 t^3$, thus by LLL $H$ is 2-colorable if $e \cdot 2^{1-k}(4^9 t^3 + 1) \leq 1$: choosing $c$ appropriately guarantees this. □

*Remark.* If we start with a $2k$ cover and we want to partition into two $k$-coverings, this happens w.h.p polinomilally. From $3k$ to two $k$-coverings, we need to bound

$$\Pr[\text{Less than } k \text{ blue or less than } k \text{ red}] \geq 2 \cdot \lambda^{-k}$$

, and we get exponential h.p.

[II] Doing the calculations, we get $m \approx (3 + o(1)) k \log k$

## 5.2 Proof of the Local Lemma

**Definition 5.3** (Dependencies Graph)**.** Let $\mathcal{A} = \{A_1 \dots A_n\}$ be events in some probability space. The *Dependencies Graph* of $\mathcal{A}$ is the DiGraph with $V = \mathcal{A}$ and $A_i$ is independent of all $A_j$ such that $A_i A_j \notin E$. We identify $A_i$ with $i$.

**Theorem 5.7** (The True Local Lemma)**.** *Let $\mathcal{A}$ be some events with dependencies graph $\mathcal{D}$. If there exists $0 < x_i < 1$ with*

$$Pr[A_i] \leq x_i \cdot Pr[i \to_{\mathcal{D}} j]\,(1 - x_j)$$

*then*

$$Pr\left[\bigcap_{i\in[n]} \overline{A_i}\right] \geq \prod_{i\in[n]} (1 - x_i)$$

*In particular, if $\mathcal{A}$ are pairwise independent, then $Pr\left[\bigcap_i \overline{A_i}\right] = \prod_i (1 - x_i)$. In the symmetric case, taking $x_i = \frac{1}{d+1}$, by assumption $e \cdot p \cdot (d + 1) \leq 1$ we have $Pr[A_i] \leq p \leq \frac{1}{e\cdot(d+1)} \leq x_i \left(1 - \frac{1}{d+1}\right)^d$.*

*Remark.* This implies 5.1

*Proof.* For any $i$, for any $i \notin S \subset [n]$, $\Pr\left[A_i \mid \cap_{j\in S}\overline{A_j}\right] \leq x_i$. This implies the lemma, since

$$\Pr\left[\bigcap \overline{A_i}\right] = \prod_{i=1}^{n} \Pr\left[\overline{A_i} \mid \bigcap_{j=1}^{i-1} \overline{A_i}\right] \geq \prod_i (1 - x_i)$$

So we prove the claim by induction on $|S|$:
$|S| = 0$ we get from the assumption.
Define $S_1 = \{j \in S \mid (i, j) \in \mathcal{D}\}$ and $S_2 = S \setminus S_1$, and let $B = \bigcap_{j\in S_1} \overline{A_j}$ and $C = \bigcap_{j\in S_2} \overline{A_j}$. Then:

$$\Pr[A_i \mid B \cap C] = \frac{\Pr[A_i \cap B \mid C]}{\Pr[B \mid C]} \leq \frac{\Pr[A_i \mid C]}{\Pr[B \mid C]} = \frac{\Pr[A_i]}{\Pr[B \mid C]} \leq \frac{x_i \prod_{i\to j}(1 - x_j)}{\Pr[B \mid C]}$$

So we need to show $\Pr[B \mid C] \geq \prod_{j\in S_1}(1 - x_j)$. Denote $S = \{j_k\}_{k\in[t]}$, then:

$$\Pr[B \mid C] = \prod_{k\in[t]} \Pr\left[\overline{A_{j_k}} \mid \prod_{k'=1}^{k-1} \overline{A_{j_{k'}}} \cap C\right] \overset{\text{induction}}{\geq} \prod_{k=1}^{t} (1 - x_{j_k}) = \prod_{j\in S_1} (1 - x_j)$$

□

### 5.2.1 The Algorithmic Version of the Lemma

Assume $A_1 \dots A_n$ are events in some product space $\Sigma^N$ That satisfy the conditions of 5.7. Is it possible to efficiently find $(\sigma_1, \dots, \sigma_N) \in \Sigma^N$ such that no $A_i$ holds? This is a generalization of 5.1.

**The Moser-Tardös Algorithm for $SAT$:** Consider a random $\sigma$. If some $A_i$ holds - resample all of the $\sigma_i$ on which he is dependent.

**Claim 5.2.1.** *The expected number of times $A_i$ is resampled is at most $\frac{x_i}{1-x_i}$ - so the runtime of the algorithm is linear.*

*Proof.* No forml Proof. Consider the "log"[III] of the events we've taken care of: $A, B, A, C, B, D, C, A' \ldots,$ and we ask "why is $A'$ resampled?" - we build a tree rooted at $A'$ - one of its predecessors in the log "broke it", say $C$ (so $C$ is a child of $A'$ in the tree), and maybe $D$ as well. Continue in this manner. We ask the probability for such a tree to occurre - and bound this. In fact, $\sum_{T \text{ Tree}} \Pr\left[T\right] \leq \frac{x_I}{1-x_i}$ □

[III]Computer-wise

# Chapter 6

# Concentration of Measure and Martingales

## 6.1 Chernoff Bound

**Theorem 6.1** (Chernoff Bound, Hoeffding). *Let $X_1 \dots X_n \underset{i.i.d}{\sim} \{\pm 1\}$, then*

$$Pr\left[\sum_{i=1}^{n} X_i > t\right] \leq e^{-\frac{t^2}{2n}}$$

*In particular - $t = \alpha\sqrt{n}$ then the bound is $e^{-\frac{\alpha^2}{2}}$, like the Gaussian Tail. If $t = \alpha \cdot n$ then the bound is $e^{-\frac{\alpha^2 n}{2}} \overset{n\to\infty}{\longrightarrow} 0$ exponentially.*

*Remark.* This can be generalized: Let $X_1 \dots X_n$ be Bernouli independent RV, with $\mu = \mathbb{E}\left[\sum X_i\right]$, then

$$\Pr\left[\sum X_i > (1+\varepsilon)\mu\right] \leq e^{-C_\varepsilon \mu^2}$$

*Proof.*

$$\Pr\left[\sum_{i=1}^{n} X_i > t\right] \overset{\lambda\geq 0}{=} \Pr\left[e^{\lambda \sum_{i=1}^{n} X_i} > e^{\lambda t}\right] \overset{\text{Markov}}{<} \frac{\mathbb{E}\left[e^{\lambda \sum_{i=1}^{n} X_i}\right]}{e^{\lambda t}} =$$
$$\frac{\mathbb{E}\left[\prod_{i=1}^{n} e^{\lambda X_i}\right]}{e^{\lambda t}} \overset{\text{i.i.d}}{=} \frac{\mathbb{E}\left[e^{\lambda X_1}\right]^n}{e^{\lambda t}}$$

Note that $\mathbb{E}\left[e^{\lambda X_1}\right]^n$ is the *Moment Generating Function.* Since $\mathbb{E}\left[e^{\lambda X_1}\right] = e^{\frac{\lambda^2}{2}}$, plugging it in results in

$$e^{\frac{\lambda^2}{2}n - \lambda t} \overset{\lambda = t/n}{=} e^{-\frac{t^2}{2n}}$$

□

### 6.1.1 Discrepancy

**Theorem 6.2.** *Let $A_1 \dots A_m \subset [n]$, then there exists $f : [n] \to \{\pm 1\}$ such that for all $i$,*

$$\left|\sum_{x \in A_i} f(x)\right| \leq \sqrt{3n \log m}$$

*Remark* (Spencer). When $m = n$, can reduce to $6\sqrt{n}$. [I]

[I] Six Standard Deviations Suffice

*Proof.* Let $f$ be chosen randomly. Then:

$$\Pr\left[\left|\sum_{x\in A_i} f(x)\right| > \sqrt{3n\log m}\right] \le 2e^{\frac{3n\log m}{2|A_i|}} \le 2m^{-\frac{3}{2}}$$

So

$$\Pr\left[\exists i \text{ such that } \left|\sum_{x\in A_i} f(x)\right| > \sqrt{3n\log m}\right] \le m\cdot 2m^{-\frac{3}{2}} \overset{\text{Large enough } m}{<} 1$$

□

**Theorem 6.3.** *Let $G \sim \mathcal{G}(n,\frac{1}{2})$. Then with probability $1-o(1)$, for any $U\subset[n]$ of size $u$:*

$$(\star) \qquad \left||E(U)| - \frac{1}{2}\binom{u}{2}\right| \le u^{\frac{3}{2}}\sqrt{\log\left(\frac{en}{u}\right)}$$

*Proof.*

$$\Pr\left[\text{The theorem Fails}\right] \le \sum_{u=1}^{n}\binom{n}{u}\cdot\Pr\left[\neg(\star)\right] \le \sum_{u=1}^{n}\left(\frac{en}{u}\right)^{u}\cdot e^{-\frac{\left(2u^{\frac{3}{2}}\sqrt{\log\frac{en}{u}}\right)^{2}}{2\binom{u}{2}}} \le$$

$$\sum_{u}$$

[redacted] □

### 6.1.2 Hidwiger's Conjecture

Recall the 4 color theorem:

**Theorem 6.4** (4-color theorem). *If $\chi(G)\ge 5$ then $G$ is nonplanar.*

**Theorem 6.5** (Wagner). *$G$ is nonplanar iff $G$ contains $K_5$ or $K_{3,3}$ as a* minor[II]*, and we denote $K_5, K_{3,3} \le G$*

Combining these results show that if $\chi(G)\ge 5$ then $G$ contains $K_5$ or $K_{3,3}$ as a minor. In fact - it can be shown that

**Theorem 6.6.** $\chi(G)\ge 5 \Rightarrow K_5 \le G$

**Conjecture.** $\chi(G)\ge t \Rightarrow K_t \le G$

This is an open problem for $t\ge 7$.

**Conjecture** (Hajos). *If $\chi(G)\ge t$ then $G$ contains a* partitioning[III] *of $K_t$, and we denote $K_t \hookrightarrow G$.*

Funny thing - this conjecture is way, way off. The following theorem shows this.

**Theorem 6.7.** *There exists a graph $G$ over $n$ vertices such that $\chi(G)\ge\frac{n}{2\log n}$, and $G$ does not contain $K_{10\sqrt{n}}$ as a topological minor.*

*Proof.* Let $G\sim\mathcal{G}(n,1/2)$ and $K_{10\sqrt{n}} \hookrightarrow G$. Then there exists $A\subset[n]$ with $|A| = 10\sqrt{n}$ that are the vertices of $K_{10\sqrt{n}}$ in the partitioning of its edges. Note that $|E(K_{10\sqrt{n}})| \approx 50n$. But at most $n$ edges of this embedding use vertices outside of $A$, so we need to show that w.h.p there is no such $A$ that contains $\ge 49n$ edges. This follows from Chernoff 6.3 □

[II] A subgraph obtained by deleting vertices, deleting edges or contracting edges
[III] Topological Minor

## 6.2 Martingales

**Definition 6.1** (Martingale)**.** A sequence of random variables $z_0, z_1 \ldots$ is called a *Martingale* if $\mathbb{E}[|z_i|] < \infty$ and

$$\mathbb{E}[z_{i+1} \mid z_0 \ldots z_i] = z_i$$

Equivalently,

$$\mathbb{E}[z_{i+1} - z_i \mid z_0 \ldots z_i] = 0$$

*Remark.* The second equality is much stronger than saying $\mathbb{E}[z_{i+1} - z_i] = 0$, as the equality in the definition is between *Random Variables*.

**Example 6.1.** Let $x_1 \ldots x_n \overset{\text{ind.}}{\sim} \{\pm 1\}$ and $z_i = \sum_{j=1}^{i} x_j$. Then:

$$\mathbb{E}[z_{i+1} \mid z_0 \ldots z_i] = \mathbb{E}[z_i + x_{i+1} \mid z_0 \ldots z_i] = \mathbb{E}[z_i \mid z_0 \ldots z_i] + \overbrace{\mathbb{E}[x_{i+1} \mid z_0 \ldots z_i]}^{0} = z_i$$

**Example 6.2.** For any gambling strategy in a fair Casino, [redacted]

**Example 6.3** (Doob, exposure martingale)**.** Let $x_1 \ldots x_n \longrightarrow \Omega$ be independent, and $f : \Omega^n \to \mathbb{R}$. Define

$$z_i(x) := \mathbb{E}_{x_j | j > i}[f(x_1, \ldots x_i, x_{i+1}, \ldots x_n]$$

That is, in the $i$'th state we were given the values of $x_1 \ldots x_i$, and $z_i$ gives his best guess to the value of $f(x)$ by taking expectation. Note that $z_0 = \mathbb{E}[f(x)]$ and $z_n = f(x)$.
Consider $x_{i,j} \in \{0, 1\}$ for any $1 \le i < j \le n$ that encode a graph. Assume $f$ is some parameter of graph (say $\chi$). Then the exposure martingale is called the *Edge Exposure Martingale*.
Consider $X_i = \{j < i \mid j \sim_G i\}$ for some underlying graph $G$. Then they can define (similarly) a martingale, that's called the *Vertex Exposure Martingale*.

**Theorem 6.8** (Azuma's inequality)**.** *Let $z_i$ be a martingale such that $|z_{i+1} - z_i| < c_{i+1}$*[IV]*, and that $z_0$ is deterministic (that is - a constant RV). Then*

$$Pr[z_n - z_0 > t] \le e^{-\frac{t^2}{2\sum c_i^2}}$$

*That is - the place we "end the process" much further than where we started behaves like centralized RVs.*

*Remark.* In the context of exposure martingales: Let $x_1, \ldots, x_n \in \Omega$ are independent, and $f$ is $C = (c_1, \ldots c_n)$*-Lipschitz* in the sense that if $x, x'$ differ only in $x_i$, then $|f(x) - f(x')| \le c_i$. Then

$$\Pr[f(x) - \mathbb{E}[f(x)] > t] \le e^{-\frac{t^2}{2\sum c_i^2}}$$

**Lemma 6.8.1.** *Assume $Y$ is a random variable with expectation $0$ and $|Y| \le c$. Then*

$$\mathbb{E}\left[e^Y\right] \le \frac{e^c + e^{-c}}{2} \le e^{-\frac{c^2}{2}}$$

*with equality iff $Y \overset{U}{\sim} \{\pm c\}$*

*Proof.* Consider the graph $e^Y$, say that the line connecting $e^c$ with $e^{-c}$ has the equation $a + by$, then by convexity - $e^Y \le a + bY$ so $\mathbb{E}\left[e^Y\right] \le \mathbb{E}[a + bY] \overset{\mathbb{E}[Y]=0}{=} a$, and $a = \frac{e^c + e^{-c}}{2}$. $\square$

[IV] This assumption is called *Bounded differences*

*Proof (Of 6.8).* WLOG $z_0 = 0$. Now:

$$\mathbb{E}\left[e^{\lambda z_n}\right] = \mathbb{E}\left[e^{\lambda(z_n - z_{n-1})} \cdot e^{\lambda z_{n-1}}\right] = \tag{6.1}$$

$$= \underset{z_1 \dots z_{n-1}}{\mathbb{E}}\left[\mathbb{E}\left[e^{\lambda(z_n - z_{n-1})} \cdot e^{\lambda z_{n-1}} \mid z_1 \dots z_{n-1}\right]\right] = \tag{6.2}$$

$$= \underset{z_1 \dots z_{n-1}}{\mathbb{E}}\left[\mathbb{E}\left[e^{\lambda(z_n - z_{n-1})} \mid z_1 \dots z_{n-1}\right] \cdot e^{\lambda z_{n-1}}\right] \tag{6.3}$$

Now, for every sample $z_1, \dots z_{n-1}$,$\mathbb{E}\left[\lambda(z_n - z_{n-1}) \mid z_1 \dots z_{n-1}\right] = 0$ by definition. Moreover, $|\lambda(z_n - zn - 1)| < |\lambda| c_n$. So by 6.8.1 taking $Y = \lambda(z_n - z_{n-1})$ *conditioned by* $z_1 \dots z_{n-1}$, we have:

$$\underset{z_1 \dots z_{n-1}}{\mathbb{E}}\left[\mathbb{E}\left[e^{\lambda(z_n - z_{n-1})} \mid z_1 \dots z_{n-1}\right] \cdot e^{\lambda z_{n-1}}\right] \leq \underset{z_1 \dots z_{n-1}}{\mathbb{E}}\left[e^{\frac{\lambda^2 c_n^2}{2}} \cdot e^{\lambda z_{n-1}}\right]$$

Continuing inductively, we get

$$\leq \dots \leq e^{\lambda^2 \left(\sum_i c_i^2\right)/2}$$

And now

$$\Pr\left[z_n - z_0 > t\right] = \Pr\left[e^{\lambda(z_n - z_0)} > e^{\lambda t}\right] \overset{\text{Markov}}{\leq} \frac{e^{\lambda^2 \left(\sum_i c_i^2\right)/2}}{e^{-\lambda t}}$$

Optimizing $\lambda$ we get the desired bound. ☐

**Example 6.4.** We throw $n$ balls into $n$ cells independently. Denote the number of empty cells by $L = \sum L_i$. Then $\mathbb{E}\left[L\right] = \sum_{i=1}^{n} \mathbb{E}\left[L_i\right] = n \cdot \left(1 - \frac{1}{n}\right)^n \in \left[\frac{n-1}{e}, \frac{n}{e}\right]$.

**Claim 6.2.1.** *$Pr\left[\left|L - \frac{n}{e}\right| > 1 + t\sqrt{n}\right] \leq e^{-\frac{t^2}{2}}$*

*Proof.* Let $X_i \in [n]$ be the cell the $i$th ball went into. Note that $L(X_1, \dots X_n)$ is 1-Lispchitz, and the claim follows by Azuma. ☐

**Example 6.5.** Consider $\chi(G)$ with $G \sim \mathcal{G}(n, p)$.

**Claim 6.2.2.** *For all $p$, let $X = \chi(G \sim \mathcal{G}(n, p))$. Then*

$$Pr\left[\left|X - \mathbb{E}\left[X\right]\right| > t\sqrt{n}\right] < e^{-\frac{t^2}{2}}$$

*Proof.* We use the Vertex Exposure Martingale. $\chi$ is 1-Lipschitz, and the claim follows by Azuma. ☐

In fact, this result can be strengthened. We do not prove this.

**Claim 6.2.3.** *If $p = n^{-\alpha}$ with $\alpha > \frac{1}{2}$, then there exists $\mu = \mu(n, p)$ such that*

$$Pr\left[\mu \leq \chi(\mathcal{G}(n, p)) \leq \mu + 1\right] \overset{n \to \infty}{\longrightarrow} 1$$

*This type of claim is called* Two Point Concentration

**Claim 6.2.4** (Relaxation)**.** *If $p = n^{-\alpha}$ with $\alpha > \frac{5}{6}$, then there exists $\mu = \mu(n, p)$ such that*

$$Pr\left[\mu \leq \chi\left(\mathcal{G}(n, p)\right) \leq \mu + 3\right] \overset{n \to \infty}{\longrightarrow} 1$$

*Proof.* Let $G \sim \mathcal{G}(n,p)$, and fix $\varepsilon$. Let $\mu$ be the maximal number such that $\Pr[\chi(G) < \mu] < \varepsilon$. We show that $\Pr[\chi(G) > \mu. + 3] < \varepsilon$ and get the result. Let $Y$ be the minimal size of a subset $S \subset V(G)$ such that $G \setminus S$ is $\mu$-colorable. Note that $Y$ is 1-Lipschitz with respect to vertex-exposure. Moreover

$$e^{-\frac{\lambda^2}{2}}\varepsilon \le \Pr[\chi(G) \ge \mu] = \Pr[Y = 0] \overset{\text{Azuma}}{\le} e^{-\frac{-\mathbb{E}[Y]^2}{2n}}$$

and therefore $\mathbb{E}[Y] \le \lambda\sqrt{n}$. Now:

$$\Pr\left[Y > 2\lambda\sqrt{n}\right] = \Pr\left[Y > \mathbb{E}[Y] + \lambda\sqrt{n}\right] \le e^{-\frac{\lambda^2}{2}} = \varepsilon$$

We not prove this, but under the assumption on $p$ - any set of vertices of size $\le 2\lambda\sqrt{n}$ is 3-colorable (because it's sparse), and the theorem follows. $\square$

**Example 6.6.** Consider the Hamming Cube $H^n = \{0,1\}^n$ with Hamming distance $d_H(x,y)$.

**Definition 6.2.** Let $A \subset H^n$ and $t > 0$. Define $A_t := \{x \mid d_H(x,A) \le t\}$

**Question 6.2.1** (Isoperimetric inequality)**.** *What is the minimal size of $A_t$ given $|A|$?*

In Euclidian space this is asking the minimal perimeter body for a given volume? The answer is a sphere.

**Theorem 6.9** (Harper)**.** *The optimal $A$ in $H^n$ is a ball.*

**Theorem 6.10.** *If $|A| = \varepsilon \cdot 2^n$, then there exists $2\sqrt{2\log\frac{1}{\varepsilon}} = t > 0$ such that $|A_{t\sqrt{n}}| > (1-\varepsilon)2^n$.*

*Proof.* Let $x \in \{0,1\}^n$, and let $f(x) = d_H(x,A)$. Clearly $f$ is 1-Lipschitz, and $\Pr[f(x) = 0] = \frac{|A|}{2^n} = \varepsilon$. But also,

$$\Pr[f(x) = 0] \le \Pr[f(x) - \mathbb{E}[f(x)] \le -\mathbb{E}[f(x)]] \overset{\text{Azuma for } -t}{\le} e^{-\frac{\mathbb{E}[f(x)]^2}{2n}}$$

So $\mathbb{E}[f(x)] \le \sqrt{2\log\frac{1}{\varepsilon}\cdot n}$, so now:

$$\frac{|\overline{A_{t\sqrt{n}}}|}{2^n} = \Pr\left[f(x) > t\sqrt{n}\right] = \Pr\left[f(x) - \mathbb{E}[f(x)] > \sqrt{2\log\frac{1}{\varepsilon}\cdot n}\right] \overset{\text{Azuma + calculations}}{\le} \varepsilon$$

$\square$

**Theorem 6.11.** *For any $A \subset H^n$, for any $t > 0$, $\frac{|A|}{2^n} \cdot \frac{|\overline{A_{t\sqrt{n}}}|}{2^n} \le e^{-\frac{t^2}{4}}$*

## 6.3 Talagrand's Inequality

### 6.3.1 The LIS problem

Talagrand's motivation was the *Longest Increasing Subsequence* problem: Given $\pi \sim \mathrm{Uni}[S_n]$, what is the longest length of an increasing subsequence in $\pi$? Denote it $L$. It's relatively easy to show that w.h.p $L = \Theta(\sqrt{n})$. For the upper bound:

$$\mathbb{E}[\#\text{increasing subsequences of length } k] = \binom{n}{k}\cdot\frac{1}{k!} \le \left(\frac{en}{k}\right)^k \cdot \left(\frac{e}{k}\right)^k = \left(\frac{e^2 n}{k^2}\right)^k \xrightarrow{k=\Omega(\sqrt{n})} 0$$

Now, take $x_i \sim \mathrm{Uni}[[0,1]]$ and order them increasingly. Then $L$ is 1-Lipschitz, and then the concentration of measure is $\approx \sqrt{n}$.

**Definition 6.3** (Talagrand Distance)**.** Let $A \subset \Omega^n$. The *Talagrand Distance* between $A$ and $x \in \Omega^n$ is

$$d_T(A, x) = \max_{\alpha \in \mathbb{R}_+^n, \|\alpha\|_2 = 1} \min_{y \in A} \sum_{i \text{ s.t } x_i \neq y_i} \alpha_i$$

*Remark.* $d_T(A, x) \geq \frac{1}{\sqrt{n}} d_H(A, x)$ by taking $\alpha_i = 1/\sqrt{n}$

Therefore, take $B_t = \{x \in \Omega^n \mid d_T(A, x) \geq t\}$, then $B_t^H = \{x \in \Omega^n \mid d_H(x, A) \geq \sqrt{n}t\}$ is contained in $B_t$.

**Theorem 6.12** (Talagrand's Inequality)**.** *For any $A \subset \Omega^n, t > 0$, $Pr[A] \cdot Pr[B] \leq e^{-\frac{t^2}{4}}$*

**Definition 6.4.** We say $f : \Omega^n \longrightarrow \mathbb{R}$ is *h-certifiable*, where $h : \mathbb{R} \to \mathbb{R}$, if for all $x \in \Omega^n$ and $s \in \mathbb{R}$, if $f(x) \geq s$ there exists $I \subset [n]$ of cardinality $h(s)$ such that

$$\forall y \in \Omega^n \quad y \mid_I = x \mid_I \Rightarrow f(y) \geq s$$

That is - there is a subset of indices of size $h(s)$ that implies $f(y) \geq s$.

Let $f$ be $1-$Lipschitz and $h$ certifiable, define $A := \left\{x \mid f(x) \leq r - t\sqrt{h(r)}\right\}, B := \{x \mid f(x) \geq r\}$.

**Claim 6.3.1.** $B \subset B_t$

*Proof.* We need to show that for any $x$ for which $f(x) \geq r$, there exists $\alpha \in \mathbb{R}_+^n$ such that for any $y \in A$, $t \leq \sum_{i:x_i \neq y_i} \alpha_i$. Let $x \in B$, and let $I$ be the certificate to $f(x) \geq r$, and let $\alpha = \frac{1}{\sqrt{|I|}} \mathbb{1}_I$. For any $y \in A$, $y$ and $x$ differ in at least $t\sqrt{h(r)}$ coordinates of $I$ since $f$ is Lipschitz, therefore

$$\sum_{i:x_i \neq y_i} \alpha_i \geq t\sqrt{h(r)} \geq t$$

$\square$

**Corollary 6.13.** $Pr[A] \cdot Pr[B] \leq e^{-\frac{t^2}{4}}$

**Corollary 6.14.** *If $L$ is the LIS of a random permutation, then*

$$Pr\left[L \leq 2\sqrt{n} - t\sqrt[4]{n}\right] \cdot Pr\left[L \geq 2\sqrt{n}\right] \leq e^{-\frac{t^2}{4}}$$

### 6.3.2 Random Matrices Application

The question is how concentrated the maximal eigenvalue of a random symmetric matrix $X \in M_{n \times n}([-1, 1])$? $\lambda_1(X) = \max_{\|v\|_2=1} v^\top X v$, and therefore 2-Lipschitz.

**Claim 6.3.2.** *For any $m \in \mathbb{R}$, $t > 0$, $Pr[\lambda_1 \leq m] \cdot Pr[\lambda_1 \geq m + t] \leq e^{-\frac{t^2}{64}}$*

*Proof.* Let $B$ be the event $\lambda_1 \geq m + t$, and $A$ the event $\lambda_1 \leq m$. We show that $B \subset B_{t/4}$ by showing for any $X$ such that $B$ holds, there exists $\alpha \in \mathbb{R}_+^{\binom{n}{2}+n}$ with $\|\alpha\| = 1$ such that for any $Y$ with $\lambda_1(Y) \leq m$ it holds that $\sum_{i,j} \alpha_{i,j} \mathbb{1}_{X_j^i \neq Y_j^i} \geq t$: Let $v \in S^{n-1}$ such that $v^\top X v = \lambda_1(X) \geq m + t$. Note that now, $v^\top Y v \leq m$, hence

$$t \leq v^\top (X - Y) v = \sum_{i,j} v_i v_j (x_j^i - y_j^i) \leq \sum_{i,j} |v_i| \cdot |v_j| \cdot 2\mathbb{1}_{X_j^i \neq Y_j^i} \leq 4 \cdot \sum_{i,j} \alpha_j^i \mathbb{1}_{X_j^i \neq Y_j^i}$$

By taking

$$\alpha_j^i = \begin{cases} 2\left|v_i\right|^2 & i = j \\ 4\left|v_i\right|\left|v_j\right| & i \neq j \end{cases}$$

so we get what we want by Talagrand. □

### 6.3.3 Geometric Interpretation of Talagrand Distance

Let $A \subset \Omega^n, x \in \Omega^n$, define

$$U(A,x) = \{s \in \{0,1\}^n \mid \exists y \in A \quad s_i = 1 \iff x_i \neq y_i\}$$

Then $d_T(A,x) = \max_\alpha \min_{y \in U(A,x)} \sum s_i \alpha_i$. This is minimizing a linear functional on some subset of the cube. Define $V(A,x) = \operatorname{conv}(U(A,x))$, then

$$d_T(A,x) = \max_\alpha \min_{v \in V(A,x)} \sum v_i \alpha_i$$

**Claim 6.3.3.** $d_T(A,x) = \min_{v \in V(A,x)} \|v\|$

*Proof.* Let $v$ be such minimizer. Let $\alpha = \frac{v}{\|v\|}$. For any $u \in V(A,x)$ we have:

$$\sum \alpha_i u_i \geq \sum \alpha_i v_i = \sum v_i^2 / \|v\| = \|v\|$$

So we get $\geq$. On the other hand, take any $\alpha$, in particular the one that maximizes the RHS. We need to show that there is some $s \in U(A,x)$ such that $\sum \alpha_i s_i \leq \|v\|$. We write $v = \sum_{s \in U(A,x)} \lambda_s s$ (a convex combination). Then:

$$\|v\| \geq \sum \alpha_i v_i = \sum_{s \in U(A,x)} \lambda_s \langle \alpha, s \rangle$$

And since the RHS is convex combination, there must be some $s$ for which the inner product is at most $\|v\|$. □

## 6.4 Harris Inequality (FKG ineq.)

**Definition 6.5.** An event $A \subset \mathbb{R}^n$ is called *increasing* if for all $x, y$ such that for any $i$, $x_i \leq y_i$:

$$x \in A \Rightarrow y \in A$$

Let $x_1, \ldots x_n \in \mathbb{R}$ independent random variables.

**Theorem 6.15** (Harris ineq.)**.** *if $A, B$ are increasing events, then:*

$$Pr[A \cap B] \geq Pr[A]\, Pr[B]$$

*If $B$ is of positive probability, then:*

$$Pr[A \mid B] \geq Pr[A]$$

*Proof.* We show that for any $f, g : \mathbb{R}^n \to \mathbb{R}$ that are monotone (in each coordinate), then $\mathbb{E}[fg] \geq \mathbb{E}[f]\mathbb{E}[g]$, we proceed by induction on $n$. For $n = 1$:

$$\mathbb{E}[f(x)g(x)] - \mathbb{E}[f(x)]\mathbb{E}[g(x)] = \frac{1}{2}\mathop{\mathbb{E}}_{x,y\text{i.i.d}}[(f(x) - f(y))(g(x) - g(y))] \overset{\text{same sign}}{\geq} 0$$

Denote $x = (\overbrace{x_1, \ldots x_{n-1}}^{\tilde{x}}, x_n)$. Now:

$$\mathbb{E}[f(x)g(x)] = \mathbb{E}_{x_n}[\mathbb{E}_{\tilde{x}} f(\tilde{x}, x_n)g(\tilde{x}, x_n)] \overset{\text{induction}}{\geq} \mathbb{E}_{x_n}[\mathbb{E}_{\tilde{x}}[f(\tilde{x}, x_n)]\mathbb{E}_{\tilde{x}}[g(\tilde{x}, x_n)]] = \star$$

Note that $\mathbb{E}_{\tilde{x}}[f(\tilde{x}, x_n)]$ is an increasing function of $x_n$, so by induction hypothesis:

$$\star \geq \mathbb{E}_{x_n}[\mathbb{E}_{\tilde{x}}[f(\tilde{x}, x_n)]]\,\mathbb{E}_{x_n}[\mathbb{E}_{\tilde{x}}[g(\tilde{x}, x_n)]] = \mathbb{E}[f]\,\mathbb{E}[g]$$

□

*Remark.* If $f$ is increasing and $g$ is decreasing, it's easy to show that the opposite inequality holds, and if $f, g$ both decreasing, then the same ineq. holds.

**Example 6.7.** What is the probability that $G \sim \mathcal{G}(n,p)$ is triangle free? Denote $T_{i,j,k}$ the event "the triangle $i, j, k$ appears", then $\overline{T_{i,j,k}}$ are decreasing, hence:

$$\Pr\left[\cap_{i,j,k}\overline{T_{i,j,k}}\right] \overset{FKG}{\geq} \prod_{i,j,k} \Pr\left[\overline{T_{i,j,k}}\right] = (1 - p^3)^{\binom{n}{3}} \overset{p=o(1)}{=} e^{-(1+o(1))\frac{p^3 n^3}{6}} = e^{-(1+o(1))\mathbb{E}[\#\triangle]}$$

**Example 6.8.** Let $G \sim \mathcal{G}(n, \frac{1}{2})$. Then:

$$\Pr\left[\cap_{v\in V} \deg(v) \geq \frac{n-1}{2}\right] \overset{FKG}{\geq} \prod_{v\in V} \Pr\left[\deg(v) \geq \frac{n-1}{2}\right] = \left(\frac{1}{2}\right)^n$$

Surprisingly - the true answer is $(c + o(1))^n$ for $c \approx 0.61$.

**Example 6.9.** Let $\mathcal{F} \subset 2^{[n]}$ such that all $A, B \in F$ intersect. Then the maximal size of $\mathcal{F}$ is $2^{n-1}$ (a nice question in extremal combinatorics). How about $\mathcal{G}$ such that no two sets cover $[n]$? same question - take complement. Cliteman asked how large can $\mathcal{H}$ be such that both properties hold? $|\mathcal{H}| \leq 2^{n-2}$ - take all sets that include 1 but does not include 2. Can be proven using FKG.

# Chapter 7

# The Poisson Paradigm

Recall that a Poisson random variable with parameter $\lambda$, denoted $X \sim \text{Pois}[\lambda]$, can be though of as a limit of $X_n \sim \text{Bin}[n,p]$. Note that setting $\lambda = np$

$$\Pr[X_n = k] = \binom{n}{k} p^k (1-p)^{n-k} = (1+o(1)) \frac{n^k}{k!} \cdot \left(\frac{c}{k}\right)^k \cdot e^{-\lambda} \overset{n\to\infty}{\to} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

That is - a sum of (many) random variable, almost$^{\text{TM}}$ that each occur with small probability is close to a Poisson Random Variable

## 7.1 Janson Inequalities

Let $A_1, \ldots A_k$ an increasing sequence of events, denote $\mu = \mathbb{E}\left[\sum_i \mathbb{1}_{A_i}\right]$, and $\Delta = \sum_{i \sim j} \Pr[A_i \cap A_j]$, where $i \sim j$ if $A_i, A_j$ are dependent. Recall the following:

1. If $X = \sum_i \mathbb{1}_{A_i}$, then $\mu = \mathbb{E}[X]$, and $Var(X) \leq \mu + \Delta$, hence $\Pr[X=0] \leq \frac{\mu+\Delta}{\mu^2}$
2. (Harris) $\Pr[X=0] \geq \prod_i \Pr\left[\overline{A_i}\right] = \prod 1 - \Pr[A_i] \approx e^{-(1+o(1))p}$

**Theorem 7.1** (Janson first inequality)**.** *Under the same setting,*

$$Pr[X=0] = Pr\left[\cap_i \overline{A_i}\right] \leq e^{-\mu + \frac{\Delta}{2}}$$

**Theorem 7.2** (Janson second inequality)**.** *If $\mu \leq \Delta$,*

$$Pr[X=0] = Pr\left[\cap_i \overline{A_i}\right] \leq e^{-\frac{\mu^2}{2\Delta}}$$

*Proof.* First, we show that 7.1 implies 7.2. For any $T \subset [k]$ we have:

$$\Pr\left[\cap_i \overline{A_i}\right] \leq \Pr\left[\cap_{i\in T} \overline{A_i}\right] \leq e^{-\mu_T + \frac{\Delta_T}{2}}$$

Now, choose $i \in T$ with probability $q$. Then

$$\mathbb{E}\left[-\mu_T + \frac{\Delta_T}{2}\right] = -q \cdot \mu + \frac{q^2 \Delta}{2} = \frac{-\mu}{2\Delta}$$

Taking $q = \frac{\mu}{\Delta} \leq 1$ by assumption gives the last equality. So there must be some $T$ with $-\mu_T + \frac{\Delta_T}{2} \leq \frac{-\mu}{2\Delta}$, ending this claim.

**Lemma 7.2.1.** $Pr\left[A_i \mid \overline{A_j} : j < i\right] \geq Pr[A_i] - \sum_{j<i, j\sim i} Pr[A_j \cap A_i]$

*Proof.* Denote $B = \bigcap_{j<i,j\sim i} \overline{A_j}, C = \bigcap_{j<i,j\not\sim i} \overline{A_j}$. Then:

$$\Pr\left[A_i \mid BC\right] = \frac{\Pr\left[A_i BC\right]}{\Pr\left[BC\right]} \geq \frac{\Pr\left[A_i BC\right]}{\Pr\left[C\right]} = \Pr\left[A_i B \mid C\right] = \Pr\left[A_i \mid C\right] - \Pr\left[A_i\overline{B} \mid C\right] =$$

$$\Pr\left[A_i\right] - \Pr\left[A_i\overline{B} \mid C\right] \overset{\star}{\geq} \Pr\left[A_i\right] - \Pr\left[A_i\overline{B}\right] = \Pr\left[A_i\right] - \Pr\left[A_i \cap \bigcup_{j\sim i,j<i} A_j\right] \overset{\text{Union Bound}}{\geq}$$

$$\Pr\left[A_i\right] - \sum_{i<j,i\sim j} \Pr\left[A_i \cap A_j\right]$$

with $\star$ is since $A_i$ are increasing, $C$ is decreasing and $\overline{B}$ is increasing. Then by FKG, they are of negative correlation, so the inequality holds ☐

Now, denoting $p_i = \Pr\left[A_i \mid \overline{A_j} : j < i\right]$ we have

$$\Pr\left[\cap\overline{A_i}\right] = \prod(1 - p_i) \leq e^{\sum p_i} = e^{-\mu + \frac{\Delta}{2}}$$

and we are done. ☐

*Remark.* It always holds that $\Pr\left[X = 0\right] \leq e^{-\frac{\mu}{2}} + e^{-\frac{\mu^2}{2\Delta}}$. This is an exponential bound, much better than the polynomial bound obtained by second moment methods.

### 7.1.1 Triangles in $\mathcal{G}(n, p)$ - again!

Let $G \sim \mathcal{G}(n, p)$, what is the probability that $G$ is triangle free? Set $A_{i,j,k}$ be the event $i, j, k$ are a triangle in $G$, hence $\mu = \frac{n^3p^3}{6}$ and $\Delta = \frac{n^4p^5}{2}$ (we've done this before). When is $\mu > \Delta$? When $p < \frac{1}{\sqrt{3n}}$. Hence

$$\Pr\left[\text{Triangle Free}\right] \leq \begin{cases} e^{-\mu(1+o(1))} & p \ll \frac{1}{\sqrt{3n}} \\ e^{-cn^2p} & p \gg \frac{1}{\sqrt{3n}} \end{cases}$$

We've already said that when $p \ll \frac{1}{n}$ then $\Pr\left[\text{Triangle Free}\right] \to 1$, and when $p \gg \frac{1}{n}$ then $\Pr\left[\text{Triangle Free}\right] \to 0$. What happens when $p = \Theta(1/n) = \frac{c}{n}$? In that case, $\mu \to \frac{c^3}{6}$, and $\Delta = \mathcal{O}(1/n)$. So

$$e^{-\frac{c^3}{6}} \longleftarrow \left(1 - \left(\frac{c}{n}\right)^3\right)^{\frac{n}{3}} \overset{\text{Harris}}{\leq} \Pr\left[\text{Triangle Free}\right] \overset{\text{Jansen}}{\leq} e^{-\frac{c^3}{6} + \mathcal{O}(1/n)}$$

### 7.1.2 Chromatic number of $\mathcal{G}(n, \frac{1}{2})$

We've already established that $\chi(G) \geq (1 + o(1))\frac{n}{2\log n}$

**Lemma 7.2.2.** *WHP, any $S \subset [n]$ of size no less of $\frac{n}{\log^2 n}$ contains a coclique of size $(1 + o(1))2\log n$.*

*Proof.*

$$\Pr\left[\text{no such } S \text{ exists}\right] = \binom{n}{\frac{n}{\log^2 n}} \cdot \Pr\left[S \text{ does not include such coclique}\right]$$

and continue using Janson... ☐

**Example 7.1** (Using Jansen)**.**

**Claim 7.1.1.** *Let $C > 2$, if $p = \left(\frac{c \cdot \log n}{n^5}\right)^{1/6}$, then w.h.p between any two vertices of $G(n, p)$ there is a path of length 6.*

*Proof.* Fix two vertices $a, b$, and let $X$ be the number of paths of length 6 between $a, b$. Then:

$$\mu = \mathbb{E}[X] = n^5(1 + o(1))p^6 = c \cdot \log n$$
$$\Delta = \mathcal{O}\left(n^5 \sum_{v=1}^{5} n^{5-v} \cdot p^{12-v}\right) = \mu^2 \mathcal{O}\left(\sum v = 1^5 (np)^{-v}\right) = \mu^2 \mathcal{O}\left(n^{-\frac{1}{6}}\right)$$

So the probability that there is no such path, that is $\Pr[X = 0]$, satisfies:

$$\Pr[X = 0] \overset{\text{Chebishev}}{\leq} \frac{1}{c \log n} + \mathcal{O}\left(n^{-\frac{1}{6}}\right)$$

So $\Pr[\text{there are } a, b \text{ with no such path}] \leq \binom{n}{2} \frac{1}{c \log n}$ - not helpful at all! But using Jansen, we have:

$$\Pr[X = 0] \leq e^{-c \log n + \mathcal{O}\left(\frac{\log^2 n}{n^{1/6}}\right)}$$

Which gives an informative bound. □

# Chapter 8

# Bits and Pieces

## 8.1 Entropy Method

**Definition 8.1** (Entropy)**.** Let $X$ be a (finitely supported) random variable. It's *Entropy* is $\mathcal{H}(X) := -\sum_{x\in X} p(x)\log_w(p(x))$

Some facts about entropy:

**Theorem 8.1.** *1. $\mathcal{H}(X) \le \log_2(|X|)$, with equality iff $X \sim Uni[Supp(X)]$*

*2. $\mathcal{H}(p) = \mathcal{H}(Ber[p]) = -p\log p - (1-p)\log(1-p)$, and an important fact is that*

$$\binom{n}{p\cdot n} \approx 2^{\mathcal{H}(p)\cdot n}$$

*3. $X \sim Bin\left[n, \frac{1}{2}\right]$, then $\mathcal{H}(X) = \frac{1}{2}\log n + o(1)$* [I]

*4. **Chain Rule:** $\mathcal{H}(X,Y) = \mathcal{H}(X) + \mathcal{H}(Y \mid X)$ with $\mathcal{H}(Y \mid X) :== \mathbb{E}_{x\sim X}\left[\mathcal{H}(Y \mid X = x)\right]$*

*5. $\mathcal{H}(Y \mid X) \le \mathcal{H}(Y)$*

*6. $\mathcal{H}(X_1,\ldots,X_n) \le \sum_i \mathcal{H}(X_i \mid X_{<i}) \le \sum_i \mathcal{H}(X_i)$*

*Proof.* All properties are a corollary of Jensen and convexity etc. □

### 8.1.1 Fake Coins

Let $A \subset [n]$ be a collection of *fake* coins, we can sample $S \subset [n]$ and get $|A \cap S|$.

**Claim 8.1.1.** *If $S_1 \ldots S_k$ satisfy that for any $A \ne A'$ there exists $i$ such that $|A\cap S_i| \ne |A' \cap S_i|$, then $k \ge (1+o(1))\frac{2n}{\log n}$. That is, we need at least $k$ samples to classify $A$ precisely.*

*Proof.* Let $A$ be some random subset of $[n]$. Let $X_i = |S_i \cap A|$. We claim that there is some bijection $(X_1 \ldots X_k) \leftrightarrow 2^{[n]}$ (that is, they encode all subsets). $(X_1 \ldots X_k)$ is supported by $2^n$ elements and is distributed uniformly on them. Hence:

$$2 = \log(2^n) = \mathcal{H}(X_1 \ldots X_k) \le \sum_i \mathcal{H}(X_i) \overset{X_i \sim \text{Bin}\left[|S_i|, \frac{1}{2}\right]}{\le} \sum_i (1+o(1))\log|S_i| \le k\left(\frac{1}{2} + o(1)\right)\log n$$

□

[I]With the intuition that since Bin $\left[n, \frac{1}{2}\right]$ is incredibly concentrated at $\mu$, we only need to understand the std, so $\sqrt{n}$.

### 8.1.2 Shearer Inequality

**Theorem 8.2** (Shearer)**.** $2\mathcal{H}(X,Y,Z) \leq \mathcal{H}(X,Y) + \mathcal{H}(X,Z) + \mathcal{H}(Y,Z)$

*Proof.*

$$\mathcal{H}(X,Y,Z) = \mathcal{H}(X) + \mathcal{H}(Y \mid X) = \mathcal{H}(Z \mid X,Y)$$

so

$$2\mathcal{H}(X,Y,Z) = \mathcal{H}(X) + \mathcal{H}(Y \mid X) = \mathcal{H}(Z \mid X) + \mathcal{H}(X) + \mathcal{H}(Y) + \mathcal{H}(Z \mid Y)$$

As required □

**Corollary 8.3.** *Let $S \subset \mathbb{R}^3$, then $Vol(S)^2 \leq Ar(S_{x,y})Ar(S_{x,z})Ar(S_{y,z})$ with $S_{x,z}$ is the shadow of $S$ on $x,z$ plane.*

*Proof.* Take some finite $P \subset \mathbb{R}^3$. We show $|P|^2 \leq |P_{x,y}| \cdot |P_{x,z}| \cdot |P_{y,z}|$: Let $(X,Y,Z)$ be a random point in $P$. Then $\log(|P|) = \mathcal{H}(X,Y,Z)$ so

$$\log|P|^2 = 2\mathcal{H}(X,Y,Z) \leq \mathcal{H}(X,Y) + \mathcal{H}(X,Z) + \mathcal{H}(Y,Z) \leq \log|P_{x,y}| + \log|P_{x,z}| + \log|P_{y,z}|$$

Taking exponent on both sides finishes the proof. For the continuous case - take a tight enough lattice $\varepsilon\mathbb{Z}^3 \cap S$ approximating $S$ □

**Theorem 8.4** (Shearer, generalized)**.** *Let $X_1 \dots X_n$ be random variables and $A_1, \dots A_m \subset [n]$ such that any $i \in [n]$ belongs to at least $k$ of the $A_j$'s. Denote by $X_{A_j}$ the (multivariate) random variable comprised of all $X_i$ such that $i \in A_j$ .Then:*

$$\sum_{j=1}^{m} \mathcal{H}\left(X_{A_j}\right) \geq k\mathcal{H}(X)$$

*Remark.* The previous Shearer inequality is a private case of this - taking $A_1 = \{1,2\}, A_2 = \{1,3\}, A_3 = \{2,3\}$

*Proof.*

$$\sum_{j=1}^{m} \mathcal{H}\left(A_{X_j}\right) = \sum_{j=1}^{m} \sum_{i \in A_j} \mathcal{H}\left(X_i \mid X_t \text{ such that } t < i, t \in A_j\right) \geq$$
$$\sum_{j=1}^{m} \sum_{i \in A_j} \mathcal{H}\left(X_i \mid X_t \text{ such that } t < i\right) \geq \sum_{i=1}^{n} \mathcal{H}\left(X_i \mid X_{<i}\right) \cdot k = k \cdot \mathcal{H}(X)$$

□

**Theorem 8.5** (Kruskal Katona, easy case)**.** *Let $G$ be a graph over $e$ edges and $t$ triangles. Then $t \leq \frac{(2e)^{\frac{3}{2}}}{6}$*

*Remark.* The hard(er) case states that if $e = \binom{x}{2}$ for some $x$, then $t \leq \binom{x}{3}$ which is the intuition from taking $e$ be the edges of a complete graph. Thinking of $e = \frac{x^2}{2}$ and $\binom{x}{3}$ as $\frac{x^3}{6}$ gives the intuition to the theorem

*Proof.* Let $X_1, X_2, X_3$ be a vertices of a triangle. Then

$$\log_2(6t) = \mathcal{H}(X_1, X_2, X_3) \overset{8.4}{\leq} \frac{1}{2}\left(\mathcal{H}(X_1,X_2) + \mathcal{H}(X_2,X_3) + \mathcal{H}(X_1,X_3)\right) = \star$$

Note that $(X_1, X_2)$ is some distribution on the edges - so $\mathcal{H}(X_1, X_2) \leq \log(2e)$. Then

$$\log_2(6t) \leq \star \leq \frac{3}{2}\log(2e)$$

Completing the proof. □

### 8.1.3 Jeff Khan's Theorem

**Theorem 8.6** (Jeff Khan). *Let $G$ be a bipartite $d$–regular graph over $n$ vertices with parts $A, B$. Denote by $\iota(G)$ the number of independent sets in $G$. Then*

$$\iota(G) \le (2^{d+1} - 1)^{\frac{n}{2d}}$$

*Remark.* Take $K_{d,d}$, then $\iota(K_{d,d}) = 2^{d+1} - 1$ (since we counted the empty set twice), so the theorem is in fact "the best case is to take copies of $K_{d,d}$".

*Proof.* Let $X = (X_1 \dots X_n)$ be indicator vector of a randomly chosen independent set. Then

$$\log(\iota(G)) = \mathcal{H}(X) = \mathcal{H}(A_A) + \mathcal{H}(X_B \mid X_A) \overset{8.4}{\le} \frac{1}{d}\sum_{b\in B}\mathcal{H}(X_{N_b}) + \sum_{b\in B}\mathcal{H}(X_b \mid X_{N_b})$$

So we need to show that for any $b$, $\mathcal{H}(X_{N_b}) + d\mathcal{H}(X_b \mid X_{N_b}) \le \log(2^{d+1} - 1)$. Let $Y$ be sampling of $X_{N_b}$ and of $d$ independent copies of $(X_b \mid X_{N_b})$, $X_b^{(1)}, X_b^{(2)}, \dots X_b^{(d)}$. Note that $Y$ is supported on $\iota(K_{d,d})$. Now, $\mathcal{H}(Y) = \mathcal{H}(X_{N_b}) + \sum_{j=1}^{d}\mathcal{H}(X_b^{(i)} \mid X_{N_b})$ since $X_b^{(i)}$ are independent. So we are done. $\square$

### 8.1.4 Latin Squares

**Definition 8.2** (LatinSquare)**.** A Latin square is an $n \times n$ array in which every sign $i \in [n]$ appears exactly once in every row and every column.

**Question 8.1.1.** *How many Latin squares are there?*

Denote the number of Latin squares over $n$ symbols by $L_n$. A guess would be

$$L_n \approx \overbrace{n^{n^2}}^{\text{write every symbols}} \cdot \overbrace{\left(\frac{n!}{n^n}\right)^{2n}}^{\text{cosntrains}} \approx \left(\frac{n}{e^2}\right)^{n^2}$$

**Theorem 8.7.** $L_n \le \left((1 + o(1))\dfrac{n}{e^2}\right)^{n^2}$

*Proof.* Let $X_{i,j}$ be a random Latin square. Then

$$\log(L_n) = \mathcal{H}(X) \overset{\star}{=} \mathop{\mathbb{E}}_{\text{random order} <}\left[\sum_{i,j}\mathcal{H}(X_{i,j} \mid X_{<i,j})\right] =$$

$$\sum_{i,j}\mathbb{E}_<\left[\mathcal{H}(X_{i,j} \mid X_{<i,j})\right] = n^2\mathbb{E}_<\left[\mathcal{H}(X_{1,1} \mid X_{<1,1})\right] = n^2\mathbb{E}_< \overbrace{\mathop{\mathbb{E}}_{L}\left[\mathcal{H}(X_{1,1} \mid X_{<11} = L_{<11}\right]}^{\text{Expose according to random Latin square}} \le$$

$$\mathbb{E}_<\mathbb{E}_L\left[\log \overbrace{(N_{1,1}(L,<))}^{\#\text{of signs not taken in } L_{1,1}}\right] = \mathbb{E}_L\mathbb{E}_<\log(N_{1,1}(L,<)) = \heartsuit$$

$\star$ - we sample $<$ by sampling $t_{i,j} \sim \text{Uni}\,[[0,1]]$ i.i.d. Therefore

$$\mathbb{E}_<\log(N_{1,1}) = \int_0^1 \mathbb{E}_{[0,1]^{n^2-1}}\log(N_{1,1}(t_{1,1}))dt_{1,1} \overset{\text{Jensen}}{\le} \int_0^1 \log\mathbb{E}\left[N_{1,1}(t_{1,1})\right]dt_{1,1} = \triangle$$

Now

$$\mathbb{E}\left[N_{1,1}(t_{1,1},<,L)\right] = 1 + (n-1)(1-t_{1,1})^2$$

Hence

$$\triangle = \int_0^1 \log(n(1-t_{1,1})^2(1+o(1)))dt_{1,1} = \log(n) + -2 + o(1)$$

We now have

$$\heartsuit \leq n^2(\log(n) - 2/\ln(2) + o(1))$$

So calculations work. □

### 8.1.5 Bergman Inequality

**Theorem 8.8** (Bergman Inequality)**.** *Let $G$ be a bipartite graph over $L \sqcup R$ with $|L| = |R| = n$, and let $d_1, \ldots d_n$ be the sequence of degrees in $L$. Then the number of perfect matchings in $G$ satisfies*

$$PM(G) \leq \prod_{i=1}^{n} (d_i!)^{\frac{1}{d_i}}$$

*Remark.* $PM(G)$ can be though of as the *permanent* (determinant without signs) of the $n \times n$ bi-adjacency matrix of $G$.

**Corollary 8.9.** *The number of latin squares over $n$ symbols satisfies $L_n \leq \left((1+o(1)\frac{n}{e^2}\right)^{n^2}$*

*Proof.* Think of an empty matrix $n \times n$ and choose some permutation to fix the symbol 1. Then, the empty cells of the matrix are a bipartite graph (with parts "rows" and "columns" as the parts). Assigning the symbol 2 is to find a perfect matching in this graph, and by Bergman there are $\prod_{i\in[n]}(n-1)^{1/(n-1)}$ matchings. This gives

$$L_n \leq \prod_{d=n}^{1} (d!)^{\frac{n}{d}}$$

and the bound is achieved by Sterling approximation. □

*Proof (Bergman, 8.8).* Let $X = (X_1, \ldots X_n)$ where $X_i$ is the neighbor of $i \in L$ in a random perfect matching. Then:

$$\log(PM(G)) = \mathcal{H}(X) = \mathbb{E}_< \sum_{i=1}^{n} \mathcal{H}(X_i \mid X_{<i}) - \sum_{i=1}^{n} \mathbb{E}_< \mathcal{H}(X_i \mid X_{<i}) =$$

$$\sum_{i=1}^{n} \mathbb{E}_< \left[\mathop{\mathbb{E}}_{M \text{ matching}} \mathcal{H}(X_i \mid X_{<i} = M_{<i})\right] \leq \sum_{i=1}^{n} \mathbb{E}_{<,M} \log(\overbrace{N_i(<,M)}^{\heartsuit}) =$$

$$\sum_{i=1}^{n} \frac{1}{d_i} \sum_{j=1}^{d_i} \log(j) = \sum_{i=1}^{n} \frac{1}{d_i} \log(d_i!) = \sum_{i=1}^{n} \log\left((d_i!)^{\frac{1}{d_i}}\right)$$

Where $\heartsuit$ is the number of available neighbors of $i$ when we reach it in the order, which is the support of $X_i$ conditioned by $X_{<i} = M_{<i}$. Note that $N_i(<,M) \sim \mathrm{Uni}\left[[d_i]\right]$: Consider $N(i)$; then any permutation affects $N_i(<,M)$ by asking "how many vertices in $N(i)$ were exposed before $i$?". Since $<$ is uniform, then the place where $i$ was exposed is uniform. Taking exponent on both sides gives the result □

### 8.1.6 GCFS

**Theorem 8.10** (Easy theorem)**.** *If $\mathcal{F} \subset 2^{[n]}$ is intersecting, then $|Ff| \leq 2^{n-1}$*

*Question* 1. What is the maximal size of a family of graphs $\mathcal{G} \subset 2^{\binom{[n]}{2}}$ such that any $G_1, G_2 \in \mathcal{G}$ intersect in a triagel?

**Theorem 8.11** (Alice, Filmus, Fridgit, 2012)**.** $|\mathcal{G}| \leq 2^{\binom{n}{2}-3}$

**Theorem 8.12** (Gram-Chang-Frenkel-Shearer)**.** $|\mathcal{G}| \leq 2^{\binom{n}{2}-2}$

*Proof.* Assume $n \equiv 0 \mod 2$. For any $S \subset [n]$ with $|S| = \frac{n}{2}$, define $E_S := \left\{e \in E \mid e \in S \vee e \in \overline{S}\right\}$. Let $K$ be the number of $S$'s fir which a given edge belongs to $E_S$ Note that

$$\binom{n}{2} K = \binom{n}{\frac{n}{2}} \cdot |E_S|$$

Then:

$$\mathcal{H}(G \in \mathcal{G}) \overset{Shearer\ 8.4}{\leq} \frac{1}{K} \sum_{S \in \binom{[n]}{\frac{n}{2}}} \mathcal{H}( \overbrace{G\, |_{E_S} |\, G \in \mathcal{G}}^{\text{Intersecting Family!}} ) \leq \frac{1}{K} \binom{n}{\frac{n}{2}} (|E_S| - 1)$$

∎

## 8.2 Phase Transition in Random Graphs

The question at hand is how do connected components look in random graphs.

**Theorem 8.13** (Erdoös, Rényi, '59)**.** *Let $G \sim \mathcal{G}(n, p)$ with $p = \frac{c}{n}$ for some $c > 0$. Denote by $C_i$ the size of the $i$'th connected components ($C_1$ being the largest). Then with high probability:*

1. *If $c < 1$ then $C_1 = O(\log(n))$.*
2. *If $c = 1$, not that important but $C_1 = \Theta(n^{2/3})$. In fact, for any fixed $j$, $C_j = \Theta(n^{2/3})$*
3. *If $c > 1$ then $C_1 = \Theta(n)$, we call it the **giant** components, and $C_2 = O(\log(n))$.*

*Remark.* The intuition here is that $c = np \approx \mathbb{E}[\deg(v)]$.

**Definition 8.3.** A Galton-Watson poisson tree with parameter $c$ is a tree generated by taking a root $r$, and generating Pois $[c]$ successors, and continuing with each of the generated vertices. Denote the tree by $GW(c)$.
More formally, let $Z_1, Z_2 \ldots \overset{i.i.d.}{\sim}$ Pois $[c]$. Denote $Y_0 = 1$, and $Y_t := Y_{t-1} - 1 + Z_t$, that is $Y_t = 1 + \sum_{i=1}^{t}(Z_i - 1)$. We think of $z_t$ as the children of the $t$'th vertex, and $Y_t$ as the bumber of vertices we've seen in time $t$. Let $T = \inf\{t \mid Y_t = 0\} \in \mathbb{N} \cup \{\infty\}$.

**Theorem 8.14** (Erdoös, Rényi reformulated)**.** *We have:*

1. *If $c \leq 1$ then $Pr[T < \infty] \longrightarrow 1$*
2. *If $c > 1$ then $Pr[T = \infty] = y > 0$ with $1 - y = e^{-cy}$.*

*First Proof, using BFS.* Note that $\deg(v) \sim \text{Bin}[n-1,p] \approx \text{Pois}[c]$. Then running BFS on $G$ results in (sort of) $GW(c)$ - in the BFS process, once we've exposed enough vertices, the successors of a vertex does not distribute as $\text{Pois}[c]$ since the binomial distribution is not with parameter $n-1$.

1. Assume $c < 1$. Note that $Y_t > 0 \iff \sum_{i=1}^{t}(z_i - 1) \geq 0$, but this is a sum of i.i.d random variables with expectation $c < 1$, so by the law of large numbers, $\Pr\left[\sum_{i=1}^{t}(z_1 - 1) \geq t\right] \longrightarrow 0$

2. Now assume $c \geq 1$. Denote $y = \Pr[T = \infty]$, then:

$$1 - y = \Pr[T < \infty] = \sum_{i=1}^{\infty} \Pr[z_i = i](1-y)^i = e^{-cy}$$

With the last equality by opening the definition of poisson distribution. Then if $c = 1$ we are done (y=0), and if $c > 1$ then we have $y > 0$.

$\square$

We can strengthen (3) in 8.13:

**Claim 8.2.1.** *If $c > 1$, then $C_1 = (y + o(1))n$ for $y$ that satisfies $1 - y = e^-cy$.*

*Remark.* If $c = 1 + \varepsilon$, then $y \approx 2\varepsilon$.

**Theorem 8.15.** *Let $c > 0$, denote the size of $GW(c)$ by $T_c$. Then for any $k \in \mathbb{N}$,*

$$Pr[T_c = k] = \lim_{n\to\infty} Pr\left[\left|C_v(\mathcal{G}\left(n, \frac{c}{n}\right)\right| = k\right] = \frac{e^{-ck}(ck)^{k-1}}{k!} \approx \frac{1}{\sqrt{2\pi}}k^{-\frac{3}{2}}\left(ce^{1-c}\right)^k$$

*for some $v \in V$.*

*Proof.* Let $z_1, \dots, z_k$ be a sequence of numbers such that if $Y_t := 1 + \sum_{i=1}^{t}(z_i - 1)$, then $Y_t \geq 0$ for any $t < k$, and $Y_k = 0$. Then

$$\Pr[T_c = k] = \sum_z \Pr[\forall t \in [k] \quad Z_i = z_i] = \sum_z \prod_{i=1}^{k} \Pr[\text{Pois}[c] = z_i]$$

Now,

$$\Pr\left[\left|C_v(\mathcal{G}\left(n, \frac{c}{n}\right)\right| = k\right] = \sum_z \Pr[(\text{any } i) \text{ the } i\text{'th vtx in the BFS starting from } v \text{ had } z_i \text{ new neighbors}] =$$

$$\sum_z \prod_{i=1}^{k} \Pr\left[\text{Bin}\left[n - 1 - \sum_{j<i} z_j, \frac{c}{n}\right] = z_i\right] = \sum_z \prod_{i=1}^{k}(1 + o(1))\Pr[\text{Pois}[c] = z_i]$$

Which shows the first equality. Now, what is $\Pr[|C_v(G)| = k]$?

$$\Pr[|C_v(G)| = k, \text{and not a tree}] \leq \binom{n}{k-1} 2^{\binom{k}{2}} \left(\frac{c}{n}\right)^k \longrightarrow 0$$

$$\Pr[|C_v(G)| = k, \text{and a tree}] = \binom{n}{k-1} k^{k-1} \left(\frac{c}{n}\right)^{k-1} \cdot \left(1 - \frac{c}{k}\right)^{k(n-k)} \longrightarrow \frac{(ck)^{k-1}e^{-ck}}{k!}$$

$\square$

# Appendix A

# Extras

## A.1 Crossing number

Turan worked on this in WWII.

**Definition A.1** (Crossing Number)**.** For a graph $G$, the *Crossing Number* $c(G)$ is the minimal amount of intersections between edges when drawn in $\mathbb{R}^2$

**Claim A.1.1.** *For any $G$, $c(G) \geq e - (3v - 6) \geq e - 3v$*

*Proof.* From Euler's Formula $\square$

And now consider taking a random induced subgraph $G'$ we have that $\mathbb{E}\left[c(G')\right] = p^4 c \geq p^2 e - 3pv$, thus

$$c \geq \frac{e}{p^2} - \frac{3v}{p^3}$$

optimizing $p$ we get $p = \frac{9v}{2e}$, so when $e > \frac{9v}{2}$, then $c \geq \alpha \cdot \frac{e^3}{v^2}$ Plugging into LHS (when $p < 1$). Therefore when taking a dense subgraph $e \sim v^2$, then $c \gtrsim v^4$