Chapter 6: The Rödl Nibble

The Probabilistic Method

Summer 2020

Freie Universität Berlin

Chapter Overview

- Introduce the Erdős-Hanani Conjecture
- Prove it with the Rödl Nibble

§1 The Erdős-Hanani Conjecture

Chapter 6: The Rödl Nibble
The Probabilistic Method

§2 The Nibble

Chapter 6: The Rödl Nibble
The Probabilistic Method

§3 The Lemma

Chapter 6: The Rödl Nibble
The Probabilistic Method

§1 The Erdős-Hanani Conjecture

Chapter 6: The Rödl Nibble

The Probabilistic Method

Edge-disjoint Triangles

Recall

- Bounding the probability of $G(n,p)$ being K_3-free
- Restricted our attention to mutually independent events
- $\Leftrightarrow$ edge-disjoint triangles

Lemma 5.4.1

There exists a family of $\frac{1}{3}\binom{n-1}{2}$ pairwise edge-disjoint triangles in K_n.

Larger cliques

- Can run the same argument for the probability of being K_k-free
- Want to find a large collection of edge-disjoint k-cliques

Hypergraphs and Packings

"Graphs are for babies" - Tom Trotter, 2017

Random t-uniform hypergraph $H^{(t)}(n,p)$

- Vertex set $V = [n]$
- Edges: each t-set in $\binom{[n]}{t}$ an edge independently with probability p

Clique containment

- Can ask for threshold for $\left\{K_k^{(t)} \subseteq H^{(t)}(n,p)\right\}$
- Upper bound on probability: use edge-disjoint hypercliques

Definition 6.1.1 (Packings)

A (k,t)-packing in $[n]$ is a family of k-sets $\mathcal{F} \subseteq \binom{[n]}{k}$ such that every t-set is contained in at most one member of the family.

An Extremal Problem

Maximum packings

- For effective bounds, want as large a packing as possible
- $m(n,k,t) = \max \{|\mathcal{F}| : \mathcal{F} \text{ is a } (k,t)\text{–packing on } [n]\}$

Proposition 6.1.2

For all $n \geq k \geq t$, we have $m(n,k,t) \leq \frac{\binom{n}{t}}{\binom{k}{t}}$.

Proof

- Given packing $\mathcal{F}$, double-count pairs (F,T) with $F \in \mathcal{F}$ and $T \in \binom{F}{t}$
- Each $F \in \mathcal{F}$ has $\binom{k}{t}$ subsets of size $t \Rightarrow |\mathcal{F}|\binom{k}{t}$ pairs
- Each t-set covered at most once $\Rightarrow$ at most $\binom{n}{t}$ pairs ■

The Case of Equality

Proposition 6.1.2

For all $n \geq k \geq t$, we have $m(n,k,t) \leq \frac{\binom{n}{t}}{\binom{k}{t}}$.

Remarks

- With our earlier construction, shows $\frac{1}{3}\binom{n-1}{2} \leq m(n,3,2) \leq \frac{1}{3}\binom{n}{2}$
- Can we do better?
- Tightness in Proposition 6.1.2: every t-set covered exactly once

Definition 6.1.3 (Designs)

A t-$(n,k,1)$ design is a family of k-sets $\mathcal{F} \subseteq \binom{[n]}{k}$ such that every t-set $T \in \binom{[n]}{t}$ is contained in exactly one set $F \in \mathcal{F}$.

The Utility of Designs

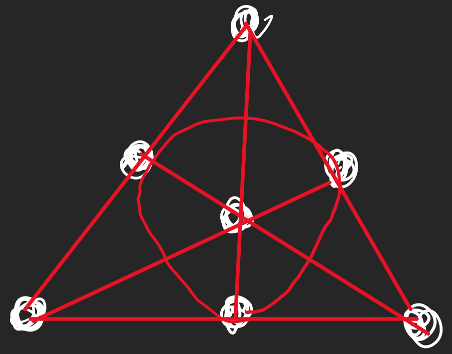
Definition 6.1.3 (Designs)

A t-$(n, k, 1)$ design is a family of k-sets $\mathcal{F} \subseteq \binom{[n]}{k}$ such that every t-set $T \in \binom{[n]}{t}$ is contained in exactly one set $F \in \mathcal{F}$.

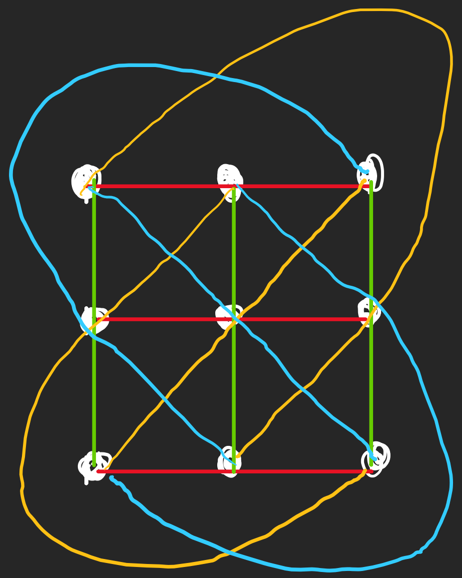
Useful objects

- Study originated in field of experiment design

Examples ($k = 3, t = 2$)



$n = 7$



$n = 9$

Divisibility Restrictions

Proposition 6.1.4

If a t-$(n, k, 1)$ design exists, then, for every $0 \leq i \leq t - 1$, $\binom{n-i}{t-i}$ is divisible by $\binom{k-i}{t-i}$.

Proof

- Fix a design $\mathcal{F} \subseteq \binom{[n]}{k}$, and consider $[i] \subseteq [n]$
- There are $\binom{n-i}{t-i}$ t-sets T with $[i] \subseteq T$
  - Each such T is contained in exactly one set $F \in \mathcal{F}$
- Each such F contains $\binom{k-i}{t-i}$ t-sets T with $[i] \subseteq T$
- $\Rightarrow |\mathcal{F}|\binom{k-i}{t-i} = \binom{n-i}{t-i}$ $\blacksquare$

- e.g.: a 2-$(n, 3, 1)$ design can only exist when $n \equiv 1,3 \pmod 6$

Asymptotic Designs

Difficulties

- Probabilistic method is blind to arithmetic conditions
  - Suggests designs will be hard to construct

Approximation

- How large a packing can we find?
- Can we ensure that almost all t-sets are contained in a k-set from the family?

Conjecture 6.1.5 (Erdős-Hanani, 1963)

For fixed $k \geq t \geq 1$, as $n \to \infty$, we have

$$m(n,k,t) = \big(1 - o(1)\big)\frac{\binom{n}{t}}{\binom{k}{t}}.$$

A Dual Problem

Types of set families

- (k,t)-packings: k-sets that cover every t-set *at most* once
- t-$(n,k,1)$ designs: k-sets that cover every t-set *exactly* once

Definition 6.1.6 (Coverings)

A (k,t)-covering of $[n]$ is a family of k-sets $\mathcal{F} \subseteq \binom{[n]}{k}$ such that every t-set $T \in \binom{[n]}{t}$ is contained in at least one set $F \in \mathcal{F}$. The size of the smallest (k,t)-covering of $[n]$ is denoted by $M(n,k,t)$.

Proposition 6.1.7

For all $n \geq k \geq t$, we have $M(n,k,t) \geq \frac{\binom{n}{t}}{\binom{k}{t}}$.

Asymptotic Packings and Coverings

Proposition 6.1.8

For fixed $k \geq t$, we have

$$\lim_{n\to\infty} \frac{m(n,k,t)\binom{k}{t}}{\binom{n}{t}} = 1 \Leftrightarrow \lim_{n\to\infty} \frac{M(n,k,t)\binom{k}{t}}{\binom{n}{t}} = 1.$$

Proof (⇒)

- Let $\mathcal{F}$ be a (k,t)-packing of size $\left(1-\mathrm{o}(1)\right)\frac{\binom{n}{t}}{\binom{k}{t}}$
- Then $\mathcal{F}$ covers $|\mathcal{F}|\binom{k}{t} = \left(1-o(1)\right)\binom{n}{t}$ of the t-sets
- Form a cover $\mathcal{F}'$ by adding a k-set covering each uncovered t-set
- $|\mathcal{F}'| = \left(1-o(1)\right)\frac{\binom{n}{t}}{\binom{k}{t}} + o(1)\binom{n}{t} = \left(1+o(1)\right)\frac{\binom{n}{t}}{\binom{k}{t}}$ ∎

Asymptotic Packings and Coverings

Proposition 6.1.8

For fixed $k \geq t$, we have

$$\lim_{n\to\infty} \frac{m(n,k,t)\binom{k}{t}}{\binom{n}{t}} = 1 \Leftrightarrow \lim_{n\to\infty} \frac{M(n,k,t)\binom{k}{t}}{\binom{n}{t}} = 1.$$

Proof ($\Leftarrow$)

- Let $\mathcal{F}$ be a (k,t)-covering of size $\left(1+\mathrm{o}(1)\right)\frac{\binom{n}{t}}{\binom{k}{t}}$
- For each t-set T, let $d_T = |\{F \in \mathcal{F} : T \subseteq F\}|$ be its degree in $\mathcal{F}$
- Form a (k,t)-packing $\mathcal{F}'$ by deleting for each t-set T any excess covering sets
- # deleted sets $\leq \sum_T (d_T - 1) = \left(\sum_T d_T\right) - \binom{n}{t} = \binom{k}{t}|\mathcal{F}| - \binom{n}{t} = o\left(\binom{n}{t}\right)$
- $\Rightarrow |\mathcal{F}'| = \left(1 - o(1)\right)\frac{\binom{n}{t}}{\binom{k}{t}}$ ■

The Random Hypergraph

- Does $H^{(k)}(n,p)$ form a good cover?

Covering sets

- A fixed t-set $T \in \binom{[n]}{t}$ is contained in $\binom{n-t}{k-t}$ sets of size k
- $\Rightarrow \mathbb{P}\left(T \text{ uncovered by } H^{(k)}(n,p)\right) = (1-p)^{\binom{n-t}{k-t}} \geq \exp\left(-2p\binom{n-t}{k-t}\right)$
- $\Rightarrow \mathbb{E}[\#\text{ uncovered } t\text{−sets}] \geq \binom{n}{t} \exp\left(-2p\binom{n-t}{k-t}\right)$
- $\Rightarrow$ to cover all t-sets, need $p = \Omega\left(\frac{\log\binom{n}{t}}{\binom{n-t}{k-t}}\right)$

Size of cover

- $\left|H^{(k)}(n,p)\right| \sim \mathrm{Bin}\left(\binom{n}{k}, p\right)$
- $\Rightarrow$ with high probability, size of cover $= \Omega\left(\frac{\binom{n}{k}\log\binom{n}{t}}{\binom{n-t}{k-t}}\right) = \Omega\left(\frac{\binom{n}{t}\log\binom{n}{t}}{\binom{k}{t}}\right)$

Summary So Far

Corollary 6.1.9

For $k \geq t$, we have $\frac{\binom{n}{t}}{\binom{k}{t}} \leq M(n,k,t) = O\left(\log\binom{n}{t}\right)\frac{\binom{n}{t}}{\binom{k}{t}}$.

Lower bound

- Double counting: each k-set covers only $\binom{k}{t}$ of the $\binom{n}{t}$ t-sets

Upper bound

- Random hypergraph $H^{(k)}(n,p)$ of appropriate density

Conjecture 6.1.5’ (Erdős-Hanani, 1963)

For fixed $k \geq t$, as $n \to \infty$, we have $M(n,k,t) = \left(1 + o(1)\right)\frac{\binom{n}{t}}{\binom{k}{t}}$.

Any questions?

§2 The Nibble

Chapter 6: The Rödl Nibble

The Probabilistic Method

Rödl to the Rescue

Conjecture 6.1.5’ (Erdős-Hanani, 1963)

For fixed $k \geq t$, as $n \to \infty$, we have $M(n,k,t) = \left(1 + o(1)\right)\frac{\binom{n}{t}}{\binom{k}{t}}$.

Theorem 6.2.1 (Rödl, 1985)

The Erdős-Hanani Conjecture is true.

Generalisation

- Rödl’s objective was to prove the Erdős-Hanani Conjecture
- His method, the Rödl Nibble, applies in more general settings
- We shall see a generalisation due to Pippinger (1989)

Hypergraph Covers

Definition 6.2.2 (Cover)

Let $H = (V, E)$ be an r-uniform n-vertex hypergraph without isolated vertices. A *cover* of H is a collection of edges $\mathcal{F} \subseteq E(H)$ that covers all the vertices; that is, $\bigcup_{e \in \mathcal{F}} e = V(H)$.

Remarks

- A cover of H is an $(n, r, 1)$-covering, whose sets are edges of H
- Each cover must contain at least $\frac{n}{r}$ edges
- Trivial to find covers of this size when $H = K_n^{(r)}$
  - Take a maximum matching
  - If needed, add one edge with remaining vertices
- Can we guarantee small covers in sparser hypergraphs?

Pippinger's Theorem

Theorem 6.2.3 (Pippinger, 1989)

For every $r \geq 2$ and large enough $D \in \mathbb{N}$, any r-uniform n-vertex hypergraph H without isolated vertices that satisfies the following conditions:

1. Almost all vertices have degree approximately D,
2. All vertices have degree $O(D)$,
3. Every pair of vertices have $o(D)$ common edges,

has a cover of size $\left(1 + o(1)\right)\frac{n}{r}$.

A Non-example

- Bounded degrees and co-degrees are necessary

Construction

- Consider a star – all edges containing some fixed vertex v_0
- Almost all vertices have degree $\binom{n-2}{r-2}$
  - But $\deg v_0 = \binom{n-1}{r-1} \gg \binom{n-2}{r-2}$
- Most pairs of vertices have co-degree $\binom{n-3}{r-3}$
  - However, v_0 and any other vertex have co-degree $\binom{n-2}{r-2}$

Large covers

- Each edge covers $r-1$ vertices from $V(H) \setminus \{v_0\}$
  - $\Rightarrow$ each cover has size at least $\frac{n-1}{r-1} \approx \left(1 + \frac{1}{r-1}\right)\frac{n}{r}$

Pippenger's Precise Theorem

Theorem 6.2.3 (Pippinger, 1989)

For every integer $r \geq 2$ and reals $\kappa \geq 1$ and $\alpha > 0$, there are $\gamma = \gamma(r, \kappa, \alpha) > 0$ and $D_0 = D_0(r, \kappa, \alpha)$ such that for every $n \geq D \geq D_0$, any r-uniform n-vertex hypergraph H without isolated vertices that satisfies the following conditions:

1. All but at most γn vertices have degree $(1 \pm \gamma)D$,
2. All vertices have degree at most κD,
3. Every pair of vertices have co-degree at most γD,

has a cover of size at most $(1 + \alpha)\frac{n}{r}$.

Small Coverings

Conjecture 6.1.5’ (Erdős-Hanani, 1963)

For fixed $k \geq t$, as $n \to \infty$, we have $M(n,k,t) = \big(1 + o(1)\big)\frac{\binom{n}{t}}{\binom{k}{t}}$.

Proof

- Build an auxiliary r-graph H, for $r \coloneqq \binom{k}{t}$
  - $V(H) = \binom{[n]}{t}$ and $E(H) = \left\{\binom{F}{t} : F \in \binom{[n]}{k}\right\}$
  - Cover of $H \leftrightarrow (k,t)$-covering of $[n]$
- Hypergraph is D-regular for $D \coloneqq \binom{n-t}{k-t} \Rightarrow \kappa = 1$
- Co-degrees are at most $\binom{n-(t+1)}{k-(t+1)} = \left(\frac{k-t}{n-t}\right) D \leq \gamma D$ when n is large
- Satisfy Pippinger’s conditions for any α
  - $\Rightarrow$ cover (hence covering) of size at most $(1+\alpha)\frac{\binom{n}{t}}{\binom{k}{t}}$

■

Proving Pippinger

"There is only one way to eat an elephant, a bite at a time."

– Desmond Tutu

The failure of randomness

- Cover some vertices several times before covering others
- Fix: prevent the random process from doing so
  - Remove covered vertices from consideration

An iterative approach

- Choose a small number of edges at random
  - Hope that they are mostly disjoint
- Remove the covered vertices from the hypergraph
  - Hope that the remaining edges are still well-distributed
- Repeat until everything is covered

One Step at a Time

Lemma 6.2.4

For every integer $r \geq 2$ and reals $\lambda \geq 1$, $\varepsilon > 0$ and $\delta' > 0$, there are $\delta = \delta(r, \lambda, \varepsilon, \delta')$ and $D_0 = D_0(r, \lambda, \varepsilon, \delta')$ such that, for every $n \geq D \geq D_0$, every r-uniform n-vertex hypergraph $H = (V, E)$ satisfying

1. For all vertices $v \in V$ except at most δn, $\deg(v) = (1 \pm \delta)D$,
2. For all vertices $v \in V$, $\deg(v) < \lambda D$, and
3. For any pair of vertices $u, v \in V$, $\deg(u, v) < \delta D$,

has a set E' of edges with the properties

a. $|E'| = (1 \pm \delta')\left(\frac{\varepsilon n}{r}\right)$,

b. for $V' = V \setminus (\cup_{e \in E'} e)$ we have $|V'| = (1 \pm \delta')ne^{-\varepsilon}$, and

c. For all but at most $\delta'|V'|$ vertices $v \in V'$, the degree of v in $H[V']$ is $(1 \pm \delta')De^{-\varepsilon(r-1)}$.

Using the Lemma

Plan of attack

- Start with original hypergraph $H_0 = H$ on vertex set $V_0 = \mathrm{V}$
- Given a hypergraph H_i, apply Lemma 6.2.4 to obtain a set of edges E_i
  - Let $V_{i+1} = V_i \setminus \left(\cup_{e \in E_i} e\right)$ be the uncovered vertices
  - $H_{i+1} = H[V_{i+1}]$ the induced hypergraph
- Once V_t is sufficiently small, cover each remaining vertex greedily
  - $\Rightarrow$ total size of cover is $|V_t| + \sum_{i<t} |E_i|$

Parameters

- With every application of the lemma, control over the distribution worsens
- Initial distribution of edges very good $\Rightarrow$ lemma can be used throughout
  - Work backwards to determine what is needed

Evolution of Parameters

Before applying the lemma

- n vertices, all but δn have degree $(1 \pm \delta)D$
- Maximum degree $< \lambda D$
- Maximum codegree $< \delta D$

After applying the lemma

- $(1 \pm \delta')ne^{-\varepsilon}$ vertices, all but δ' proportion have degree $(1 \pm \delta')De^{-\varepsilon(r-1)}$
- Maximum degree $< \lambda D$, maximum codegree $< \delta D$

Change of parameters

- $D_{i+1} \coloneqq D_i e^{-\varepsilon(r-1)}$
- $\Rightarrow \lambda_{i+1} \coloneqq \lambda_i e^{\varepsilon(r-1)}, \delta_{i+1} \geq \delta_i e^{\varepsilon(r-1)}$
- Need $\delta_i \leq \delta(r, \lambda_i, \varepsilon, \delta_{i+1})$ to apply lemma

Size of Vertex and Edge Sets

Vertex sets

- By Lemma 6.2.4, $|V_i| \leq (1+\delta_i)|V_{i-1}|e^{-\varepsilon}$
- $\Rightarrow |V_i| \leq \left(\prod_{j=1}^{i}(1+\delta_j)\right)ne^{-i\varepsilon} \leq \left(1+\sum_{j=1}^{i}\delta_j\right)ne^{-i\varepsilon}$
  - By growing the δ_i fast enough, can ensure $\sum_{j=1}^{i}\delta_j \leq 2\delta_t$

Edge sets

- Lemma 6.2.4: $|E_i| \leq (1+\delta_{i+1})\frac{\varepsilon|V_i|}{r}$

$$\leq (1+\delta_{i+1})(1+2\delta_t)\frac{\varepsilon n e^{-i\varepsilon}}{r}$$

$$\leq (1+4\delta_t)\frac{\varepsilon n e^{-i\varepsilon}}{r}$$

Size of the Cover

Recall

- $|V_i| \le (1+2\delta_t)ne^{-i\varepsilon}$ and $|E_i| \le (1+4\delta_t)\frac{\varepsilon n e^{-i\varepsilon}}{r}$

Total size of cover

- $|V_t| + \sum_{i=0}^{t-1}|E_i| \le (1+2\delta_t)ne^{-t\varepsilon} + (1+4\delta_t)\frac{\varepsilon n}{r}\sum_{i=0}^{t-1} e^{-i\varepsilon}$
  $$\le (1+4\delta_t)\left(re^{-t\varepsilon} + \frac{\varepsilon}{1-e^{-\varepsilon}}\right)\frac{n}{r}$$
- Choosing t large, can ensure $re^{-t\varepsilon} \le \varepsilon$
- $1 - e^{-\varepsilon} \ge 1 - \left(1 - \varepsilon + \frac{1}{2}\varepsilon^2\right) = \varepsilon\left(1 - \frac{1}{2}\varepsilon\right)$
  - $\Rightarrow \frac{\varepsilon}{1-e^{-\varepsilon}} \le \frac{1}{1-\frac{1}{2}\varepsilon} \le 1+\varepsilon$
- $\Rightarrow$ cover has size at most $(1+4\delta_t)(1+2\varepsilon)\frac{n}{r}$
  - By choosing ε, δ_t sufficiently small, we can ensure this is at most $(1+\alpha)\frac{n}{r}$

Piecing It Together

Theorem 6.2.3 (Pippinger, 1989)

For every integer $r \geq 2$ and reals $\kappa \geq 1$ and $\alpha > 0$, there are $\gamma = \gamma(r, \kappa, \alpha) > 0$ and $D_0 = D_0(r, \kappa, \alpha)$ such that for every $n \geq D \geq D_0$, any r-uniform n-vertex hypergraph H with well-distributed edges has a cover of size at most $(1 + \alpha)\frac{n}{r}$.

Proof

- Choose ε, δ so that $(1 + 4\delta)(1 + 2\varepsilon) < 1 + \alpha$, and t so that $re^{-t\varepsilon} \leq \varepsilon$
- Set $\lambda_i := \kappa e^{i\varepsilon(r-1)}$ and $D_i := De^{-i\varepsilon(r-1)}$ for each $0 \leq i \leq t$
- Set $\delta_t := \delta$, and, for $i = t - 1, t - 2, \ldots, 0$, choose δ_i such that
  - $\delta_i \leq \delta(r, \lambda_i, \varepsilon, \delta_{i+1})$ from Lemma 6.2.4, $\delta_i \leq e^{-\varepsilon(r-1)}\delta_{i+1}$ and $\delta_i \leq \frac{1}{2}\delta_{i+1}$
- Set $\gamma := \delta_0$ and D_0 such that $D_i := D_0 e^{-i\varepsilon(r-1)} \geq D(r, \lambda_i, \varepsilon, \delta_{i+1})$ for all i
- We can then iterate the lemma t times, giving the small cover ■

Any questions?

§3 The Lemma

Chapter 6: The Rödl Nibble

The Probabilistic Method

Recalling the Statement

Lemma 6.2.4

For every integer $r \geq 2$ and reals $\lambda \geq 1$, $\varepsilon > 0$ and $\delta' > 0$, there are $\delta = \delta(r, \lambda, \varepsilon, \delta')$ and $D_0 = D_0(r, \lambda, \varepsilon, \delta')$ such that, for every $n \geq D \geq D_0$, every r-uniform n-vertex hypergraph $H = (V, E)$ satisfying

1. For all vertices $v \in V$ except at most δn, $\deg(v) = (1 \pm \delta)D$,
2. For all vertices $v \in V$, $\deg(v) < \lambda D$, and
3. For any pair of vertices $u, v \in V$, $\deg(u, v) < \delta D$,

has a set E' of edges with the properties

a. $|E'| = (1 \pm \delta')\binom{\varepsilon n}{r}$,

b. for $V' = V \setminus (\cup_{e \in E'} e)$ we have $|V'| = (1 \pm \delta')ne^{-\varepsilon}$, and

c. For all but at most $\delta'|V'|$ vertices $v \in V'$, the degree of v in $H[V']$ is $(1 \pm \delta')De^{-\varepsilon(r-1)}$.

Proof Strategy

Selection of edges

- Select each edge to be in E' independently at random

Analysis

- Estimate $|E'|$, $\mathbb{P}(v \in V')$ and $\mathbb{P}(e \in H[V'])$
- Concentration inequalities $\Rightarrow$ hypergraph statistics close to expectations
- Quantifying over vertices
  - Polynomial concentration suffices
  - Can use Chebyshev's Inequality

Proof of Lemma, Part a

Lemma 6.2.4.a

a. $|E'| = (1 \pm \delta')\binom{\varepsilon n}{r}$

Proof

- Select each edge to be in E' independently with probability $p = \frac{\varepsilon}{D}$
- $\Rightarrow |E'| \sim \mathrm{Bin}(e(H), p)$
- Handshake Lemma $\Rightarrow e(H) = \frac{1}{r}\sum_v \deg(v)$
- Sum of degrees
  - At least $(1-\delta)n \cdot (1-\delta)D + \delta n \cdot 0 = (1-\delta)^2 nD \geq (1-2\delta)nD$
  - At most $(1-\delta)n \cdot (1+\delta)D + \delta n \cdot \lambda D = (1-\delta^2+\delta\lambda)nD$
  - $\Rightarrow e(H) = (1 \pm \delta_1)\frac{nD}{r}$ for some $\delta_1 = \delta_1(\delta, \lambda) \to 0$ as $\delta \to 0$

Proof of Lemma, Part a

Lemma 6.2.4.a

$$|E'| = (1 \pm \delta')\left(\frac{\varepsilon n}{r}\right).$$

Proof (cont'd)

- Recall
  - Each edge selected with probability $p = \frac{\varepsilon}{D}$
  - $e(H) = (1 \pm \delta_1)\frac{nD}{r}$
- $\Rightarrow \mathbb{E}[|E'|] = e(H)p = (1 \pm \delta_1)\frac{\varepsilon n}{r}$
- $\mathrm{Var}(|E'|) = e(H)p(1-p) \leq \mathbb{E}[|E'|] = o(\mathbb{E}[|E'|]^2)$
- $\therefore$ Chebyshev $\Rightarrow$ with high probability, $|E'| = (1 \pm 2\delta_1)\frac{\varepsilon n}{r}$ $\blacksquare$

Proof of Lemma, Part b

Lemma 6.2.4.b

For $V' = V \setminus (\cup_{e \in E'} e)$ we have $|V'| = (1 \pm \delta')ne^{-\varepsilon}$.

Proof

- $|V'| = \sum_{v \in V} 1_{\{v \in V'\}}$
- $\mathbb{E}\left[1_{\{v \in V'\}}\right] = \mathbb{P}(v \in V') = (1-p)^{\deg(v)}$
- When $\deg(v) = (1 \pm \delta)D$:
  - $\mathbb{E}\left[1_{\{v \in V'\}}\right] = \left(1 - \frac{\varepsilon}{D}\right)^{(1\pm\delta)D} = (1 \pm \delta_2)e^{-\varepsilon}$ for some $\delta_2 = \delta_2(\varepsilon, \delta) \to 0$ and D large
- At most δn exceptional vertices, for which $0 \leq \mathbb{E}\left[1_{\{v \in V'\}}\right] \leq 1$
- $\Rightarrow \mathbb{E}[|V'|] = (1 \pm \delta_3)ne^{-\varepsilon}$ for some $\delta_3 = \delta_3(\varepsilon, \delta) \to 0$

Proof of Lemma, Part b

Lemma 6.2.4.b

For $V' = V \setminus (\cup_{e \in E'} e)$ we have $|V'| = (1 \pm \delta')ne^{-\varepsilon}$.

Proof (cont'd)

- $\mathrm{Var}(|V'|) = \sum_{v\in V} \mathrm{Var}\left(1_{\{v\in V'\}}\right) + \sum_{u\neq v} \mathrm{Cov}\left(1_{\{u\in V'\}}, 1_{\{v\in V'\}}\right)$
- $\sum_{v\in V} \mathrm{Var}\left(1_{\{v\in V'\}}\right) \leq \sum_{v\in V} \mathbb{E}\left[1_{\{v\in V'\}}\right] = \mathbb{E}[|V'|]$
- $$\begin{aligned} \mathrm{Cov}\left(1_{\{u\in V'\}}, 1_{\{v\in V'\}}\right) &= \mathbb{E}\left[1_{\{u\in V'\}} 1_{\{v\in V'\}}\right] - \mathbb{E}\left[1_{\{u\in V'\}}\right] \mathbb{E}\left[1_{\{v\in V'\}}\right] \\ &= (1-p)^{\deg(u)+\deg(v)-\deg(u,v)} - (1-p)^{\deg(u)+\deg(v)} \\ &\leq (1-p)^{-\deg(u,v)} - 1 \leq \left(1 - \frac{\varepsilon}{D}\right)^{-\delta D} - 1 \end{aligned}$$
- $\Rightarrow \mathrm{Cov}\left(1_{\{u\in V'\}}, 1_{\{v\in V'\}}\right) \leq \delta_4$ for some $\delta_4 = \delta_4(\varepsilon, \delta) \to 0$

Proof of Lemma, Part b

Lemma 6.2.4.b

For $V' = V \setminus (\cup_{e \in E'} e)$ we have $|V'| = (1 \pm \delta')ne^{-\varepsilon}$.

Proof (cont'd)

- Recall
  - $\mathbb{E}[|V'|] = (1 \pm \delta_3)ne^{-\varepsilon}$
  - $\mathrm{Cov}\left(1_{\{u \in V'\}}, 1_{\{v \in V'\}}\right) \leq \delta_4$
- $\Rightarrow \mathrm{Var}(|V'|) \leq \mathbb{E}[|V'|] + \delta_4 n^2 \leq 2\delta_4 n^2$
- Chebyshev: $\mathbb{P}(|V'| \neq (1 \pm \delta_5)ne^{-\varepsilon}) \leq \frac{2\delta_4 n^2}{(\delta_5 - \delta_3)^2 n^2 e^{-2\varepsilon}}$
  - This can be made arbitrarily small for appropriate choice of $\delta_5 \to 0$ $\blacksquare$

Proof of Lemma, Part c

Lemma 6.2.4.c

For all but at most $\delta'|V'|$ vertices $v \in V'$, the degree of v in $H[V']$ is $(1 \pm \delta')De^{-\varepsilon(r-1)}$.

Proof (outline)

- Fix a vertex $v \in V$, and condition on $v \in V'$
- Need to study how many edges $e \ni v$ survive in $H[V']$
  - Edge e survives if and only if $u \in V'$ for all $u \in e$
- We have good control over vertices of degree $(1 \pm \delta)D$
  - $\Rightarrow$ can control edges whose vertices are all of typical degree
  - Call such edges *good*, and *bad* otherwise
- $\Rightarrow$ can control $\deg(v)$ if most edges $e \ni v$ are good
- Shall show that degree conditions $\Rightarrow$ most vertices are mostly in good edges

Good Edges

Claim 6.3.1

There is some $\delta_6 \to 0$ such that:

i. all but at most $\delta_6 n$ vertices have $\deg(v) = (1 \pm \delta_6)D$, and are in at most $\delta_6 D$ bad edges.

ii. if an edge e is good, then given some $v \in e$, we have
$$|\{f \in E: v \notin f, f \cap e \neq \emptyset\}| = (1 \pm \delta_6)(r-1)D.$$

Proof of i.

- At most δn vertices have $\deg(v) \neq (1 \pm \delta)D$
- $\Rightarrow$ there are at most $\delta n \cdot \lambda D$ bad edges.
- $\Rightarrow$ at most $\frac{\delta \lambda n D}{\delta_6 D}$ vertices can be in more than $\delta_6 D$ bad edges
- For a suitable choice of $\delta_6 \to 0$, this is less than $(\delta_6 - \delta)n$ ∎

Good Edges

Claim 6.3.1

There is some $\delta_6 \to 0$ such that:

i. all but at most $\delta_6 n$ vertices have $\deg(v) = (1 \pm \delta_6)D$, and are in at most $\delta_6 D$ bad edges.

ii. if an edge e is good, then given some $v \in e$, we have
$$|\{f \in E : v \notin f, f \cap e \neq \emptyset\}| = (1 \pm \delta_6)(r-1)D.$$

Proof of ii.

- e good $\Rightarrow$ for the $r-1$ vertices $u \in e, u \neq v$, we have $\deg(u) = (1 \pm \delta)D$
- $\Rightarrow |\{f \in E : v \notin f, f \cap e \neq \emptyset\}| \leq (1+\delta)(r-1)D$
- Overcounted: edges f that meet two vertices of e
  - Co-degree bound $\Rightarrow$ at most $\binom{r}{2}\delta D$ such edges
- $\Rightarrow |\{f \in E : v \notin f, f \cap e \neq \emptyset\}| \geq (1-\delta)(r-1)D - \binom{r}{2}\delta D$ ∎

Survival of Good Edges

Claim 6.3.2

There is some $\delta_7 \to 0$ such that, if we condition on $v \in V'$, and e is a good edge containing v, then $\mathbb{P}(e \subseteq V') = (1 \pm \delta_7)e^{-\varepsilon(r-1)}$.

Proof

- $v \in V' \Rightarrow$ no edge containing v was selected in E'
- $e \subseteq V' \Rightarrow$ every $u \in e$ is also in V'
  - $\Rightarrow$ no edge $f \in E$ with $f \cap e \neq \emptyset$ is selected in E'
- By assumption, this is true for every $f \ni v$
  - $\Rightarrow$ need only consider $\{f \in E: v \notin f, f \cap e \neq \emptyset\}$
- Claim 6.3.1.ii $\Rightarrow$ there are $(1 \pm \delta_6)(r-1)D$ such edges
- Probability none are selected in E' is $(1-p)^{(1\pm\delta_6)(r-1)D}$
- $p = \frac{\varepsilon}{D} \Rightarrow$ this is $(1 \pm \delta_7)e^{-\varepsilon(r-1)}$ ■

Expected Degrees

Claim 6.3.3

There is some $\delta_8 \to 0$ such that, if v is a vertex as in Claim 6.3.1.i and we condition on $v \in V'$, then the expected degree $\deg'(v)$ of v in $H[V']$ is $(1 \pm \delta_8)De^{-\varepsilon(r-1)}$.

Proof

- For each edge $e \in E$, let 1_e be the indicator for the event $e \subseteq V'$
- $\Rightarrow$ degree of v in $H[V']$ is $\sum_{e \ni v} 1_e$
- At most $\delta_6 D$ bad edges containing v
  - $\deg'(v) = \sum_{e \ni v, e \text{ good}} 1_e \pm \delta_6 D$
- Number of good edges containing v is $(1 \pm \delta \pm \delta_6)D$
- Claim 6.3.2 $\Rightarrow \mathbb{E}[1_e] = (1 \pm \delta_7)e^{-\varepsilon(r-1)}$ for every good $e \ni v$
- $\Rightarrow \mathbb{E}[\deg'(v)] = (1 \pm \delta \pm \delta_6)(1 \pm \delta_7)De^{-\varepsilon(r-1)} \pm \delta_6 D$ ■

Variance in Degrees

Claim 6.3.4

There is some $\delta_9 \to 0$ such that, if v is a vertex as in Claim 6.3.1.i and we condition on $v \in V'$, then $\mathrm{Var}(\deg'(v)) \leq \delta_9 D^2$.

Proof

- As usual, $\mathrm{Var}(\deg'(v)) \leq \mathbb{E}[\deg'(v)] + \sum_{v \in e,f; e \neq f} \mathrm{Cov}(1_e, 1_f)$
- Contribution to sum from bad edges is at most $\delta_6(1+\delta)D^2$
- Fix good $e \ni v$, and estimate $\sum_{f \text{ good}: v \in f \neq e} \mathrm{Cov}(1_e, 1_f)$
- Let $T(e) = \{h \in E : v \notin h, v \cap e \neq \emptyset\}$, and let $t(e,f) = |T(e) \cap T(f)|$
- $\mathrm{Cov}(1_e, 1_f) = \mathbb{E}[1_e 1_f] - \mathbb{E}[1_e]\mathbb{E}[1_f]$

$$= (1-p)^{|T(e) \cup T(f)|} - (1-p)^{|T(e)| + |T(f)|}$$

$$\leq (1-p)^{-t(e,f)} - 1$$

Variance in Degrees

Claim 6.3.4

There is some $\delta_9 \to 0$ such that, if v is a vertex as in Claim 6.3.1.i and we condition on $v \in V'$, then $\mathrm{Var}(\deg'(v)) \leq \delta_9 D^2$.

Proof (cont'd)

- Recall
  - $T(e) = \{h \in E: v \notin h, v \cap e \neq \emptyset\}$ and $t(e,f) = |T(e) \cap T(f)|$
  - $\mathrm{Cov}(1_e, 1_f) \leq (1-p)^{-t(e,f)} - 1$
- For each $u \in e$, $\deg(u,v) \leq \delta D$
  - $\Rightarrow$ at most $(r-1)\delta D$ edges f with $|e \cap f| \geq 2$
- Otherwise $e \cap f = \{v\}$
  - $\Rightarrow$ for all $h \in T(e) \cap T(f)$ there are $u \in e \setminus \{v\}, u' \in f \setminus \{v\}$ with $u, u' \in h$
  - $\Rightarrow t(e,f) \leq (r-1)^2 \delta D$
  - $\Rightarrow \mathrm{Cov}(1_e, 1_f) \leq \left(1 - \frac{\varepsilon}{D}\right)^{-(r-1)^2 \delta D} - 1 \leq r^2 \varepsilon \delta$

Variance in Degrees

Claim 6.3.4

There is some $\delta_9 \to 0$ such that, if v is a vertex as in Claim 6.3.1.i and we condition on $v \in V'$, then $\mathrm{Var}(\deg'(v)) \leq \delta_9 D^2$.

Proof (cont'd)

- $\mathrm{Var}(\deg'(v)) \leq 2\delta_6 D^2 + \sum_{v\in e, e \text{ good}} \sum_{v \in f, f \text{ good}} \mathrm{Cov}(1_e, 1_f)$

$$\leq 2\delta_6 D^2 + \sum_{v\in e, e \text{ good}} \left((r-1)\delta D + \sum_{f \text{ good}, f\cap e = \{v\}} \mathrm{Cov}(1_e, 1_f)\right)$$

$$\leq 2\delta_6 D^2 + \sum_{v\in e, e \text{ good}} \left((r-1)\delta D + \sum_{f \text{ good}, f\cap e = \{v\}} r^2 \varepsilon \delta\right)$$

$$\leq 2\delta_6 D^2 + (1+\delta)D \cdot \left((r-1)\delta D + (1+\delta)D \cdot r^2 \varepsilon\delta\right)$$

- For appropriate $\delta_9 \to 0$, this is at most $\delta_9 D^2$ ■

Completing the Proof

Lemma 6.2.4.c

For all but at most $\delta'|V'|$ vertices $v \in V'$, the degree of v in $H[V']$ is $(1 \pm \delta')De^{-\varepsilon(r-1)}$.

Proof

- All but at most $\delta_6 n$ vertices are as in Claim 6.3.1.i; can ignore the rest
- For such a vertex v, conditioning on $v \in V'$:
  - Claim 6.3.3: $\mathbb{E}[\deg'(v)] = (1 \pm \delta_8)De^{-\varepsilon(r-1)}$
  - Claim 6.3.4: $\mathrm{Var}(\deg'(v)) \leq \delta_9 D^2$
- Chebyshev: for some $\delta_{10} \to 0$, $\mathbb{P}\left(\deg(v) \neq (1 \pm \delta_{10})De^{-\varepsilon(r-1)}\right) \leq \delta_6$
- Markov: the probability of having more than $2\delta_6|V'|$ such vertices in V' whose degree is not $(1 \pm \delta_{10})De^{-\varepsilon(r-1)}$ is less than $\frac{1}{2}$ ■

Epilogue

Central question

- For which n does a t-$(n, k, 1)$ design exist?
- Divisibility conditions ⇒ infinite sequence of possible values
  - These conditions are necessary, but *not* sufficient

Erdős-Hanani Conjecture / Rödl's Theorem

- ⇒ for all large n, *asymptotic* designs exist

Exact results

- Wilson (1972-1975): $t = 2, k \geq 3, n$ large and satisfying divisibility conditions
- Keevash (2014+): generalised Wilson to all t
  - Follows the steps of Rödl Nibble, but uses an algebraic construction to complete design
- Glock, Kühn, Lo and Osthus (2016+): new proof of existence of designs
  - Proof is purely combinatorial/probabilistic

Any questions?