The Rödl Nibble

Frédéric Havet

COATI, INRIA, I3S, CNRS, Univ. Nice Sophia Antipolis
Sophia Antipolis, France

17ème JCALM –Sophia Antipolis – 4-5 mai 2017
Beyond design

Design = perfect case \hspace{1cm} \textbf{BUT} \hspace{1cm} do not always exists.

No decomposition of $K_5$ into copies of $K_3$.

\[ |E(K_5)| = 10 \text{ and } |E(K_3)| = 3. \]

No decomposition of $K_6^{(3)}$ into copies of $K_4^{(3)}$.

\[ \text{(degree in } K_6^{(3)} = 10 \text{ and degree in } K_4^{(3)} = 3) \]
Approximate design

**packing** of $F$ in $G$ : set of edge-disjoint copies of $F$ in $G$.

**covering** of $G$ by $F$ : set of copies of $F$ such that every edge of $G$ is in one of the copies.
Approximate design

**packing** of $F$ in $G$: set of edge-disjoint copies of $F$ in $G$.

**covering** of $G$ by $F$: set of copies of $F$ such that every edge of $G$ is in one of the copies.
Approximate design

**packing** of $F$ in $G$ : set of edge-disjoint copies of $F$ in $G$.

**covering** of $G$ by $F$ : set of copies of $F$ such that every edge of $G$ is in one of the copies.

$m(n, k, t) : \text{maximum size of a packing of } K_k^{(t)} \text{ in } K_n^{(t)}$

$M(n, k, t) : \text{minimum size of a covering of } K_n^{(t)} \text{ by } K_k^{(t)}$. 
Trivial inequalities

\[
m(n, k, t) \leq \frac{\binom{n}{t}}{\binom{k}{t}} \leq M(n, k, t)
\]

with equalities iff there is a \( t-(n, k, 1) \)-design.
Examples

\[ m(5, 3, 2) = 2 < \frac{\binom{5}{2}}{\binom{3}{2}} = \frac{10}{3} \leq M(5, 3, 2) = 4 \]

\[ m(6, 4, 3) = 3 < \frac{\binom{6}{3}}{\binom{4}{3}} = 5 < M(6, 4, 3) = 6 \]
Erdős–Hanani Conjecture

\[ m(n, k, t) \leq \frac{\binom{n}{t}}{\binom{k}{t}} \leq M(n, k, t) \]

Conjecture (Erdős–Hanani, 1963)

\[ \lim_{n \to +\infty} \frac{m(n, k, t)}{\binom{n}{t} / \binom{k}{t}} = 1 \quad \text{and} \quad \lim_{n \to +\infty} \frac{M(n, k, t)}{\binom{n}{t} / \binom{k}{t}} = 1. \]

Erdős–Hanani: True for \( t = 2 \) (graphs).
Rödl Theorem

\[ m(n, k, t) \leq \binom{n}{t} / \binom{k}{t} \leq M(n, k, t) \]

Theorem (Rödl, 1985)

\[ \lim_{n \to +\infty} \frac{m(n, k, t)}{(n \choose t) / (k \choose t)} = 1 \quad \text{and} \quad \lim_{n \to +\infty} \frac{M(n, k, t)}{(n \choose t) / (k \choose t)} = 1. \]
The auxiliary hypergraph $\mathcal{H}$

**Hypergraph** $\mathcal{H}$: vertices $= t$-subsets of $[n]$
hyperedges $= \text{sets of } \binom{k}{t} \text{ } t\text{-subsets of a } k\text{-subset of } [n]$.

packing of $K_k^{(t)}$ in $K_n^{(t)}$ $\leftrightarrow$ **matching** in $\mathcal{H}$
**matching**: set of pairwise vertex-disjoint hyperedges.

covering of $K_n^{(t)}$ by $K_k^{(t)}$ $\leftrightarrow$ **cover** of $\mathcal{H}$
**cover**: set of hyperedges s.t. every vertex is in one of them.

**Theorem (Rödl, 1985)**

\[ \mathcal{H} \text{ has a matching of size } (1 - o(1)) \frac{N}{r} \quad \text{and a cover of size } (1 - o(1)) \frac{N}{r} \]

\[ N = \binom{n}{t} \quad r = \binom{k}{t}. \]
A more general result

\( \mathcal{H} \) has \( N \) vertices with \( N = \binom{n}{t} \).

\( \mathcal{H} \) is \( r \)-uniform with \( r = \binom{k}{t} \).

\( \mathcal{H} \) is \( D \)-regular with \( D = \binom{n-t}{k-t} \).

**Observation**: The existence of the desired matching and cover also holds for large class of \( r \)-uniform hypergraphs \( D \)-regular graphs.

Frank and Rödl, Pippenger and Spencer 1989, Kahn 1996.
Equivalence of the two statements

In a $D$-regular $r$-uniform hypergraph of order $N$

matching of size $(1 - o(1)) \frac{N}{r} \Leftrightarrow$ cover of size $(1 + o(1)) \frac{N}{r}$

$(\Rightarrow)$ matching of size $\frac{N(1-\epsilon)}{r}$, ⇒ at most $\epsilon N$ non-covered vertices
  ⇒ cover of size at most $\frac{N(1-\epsilon)}{r} + \epsilon N$.

$(\Leftarrow)$ cover of size $\frac{N(1+\epsilon)}{r}$. Each vertex $x$ covered $c(x)$ times.
For each $x$ choose $c(x) - 1$ hyperedges and remove them.
At most $\sum_{v \in V} (c(x) - 1) = \frac{(1+\epsilon)N}{r} \times r - N = \epsilon N$ are removed.
⇒ matching of size at least $\frac{N(1+\epsilon)}{r} - \epsilon N$. 
General idea: nibbling

Fix $\epsilon$.

Take a random set of $\epsilon N/r$ edges. W.h.p. only $O(\epsilon^2 N)$ vertices covered more than once. So at least $\epsilon N - O(\epsilon^2 N)$ covered vertices.

Remove the covered vertices.

Choose again a random set of edges covering roughly an $\epsilon$-fraction of the vertices with almost no overlap.

And so on until at most $\epsilon N$ vertices remain.

Cover each of the remaining vertices with a dedicated hyperedge.
The Theorem

$r > 2$ fixed.
For $\kappa \geq 1$ and $a > 0$, there exists $\gamma > 0$ and $d_0$ s.t. the following holds for all $N \geq D \geq d_0$.

Every $r$-uniform hypergraph $H = (V, E)$ on $N$ vertices such that

(0) $d(x) > 0$ for all $x \in V$.
(1) All vertices $x$ except at most $\gamma N$ satisfy $d(x) = (1 \pm \gamma)D$.
(2) For all $x \in V$, $d(x) < \kappa D$.
(3) For each pair $(x, y)$ of distinct vertices, $d(x, y) < \gamma D$.

has a cover of size $(1 + a)\frac{N}{r}$.

d(x, y) : codegree = number of edges containing both $x$ and $y$.

In $\mathcal{H}$, we have $d(x, y) = \binom{n-t-1}{k-t-1} = o(D)$. 
The Nibble Lemma

\( r > 2 \) fixed.
For \( K \geq 1, \epsilon > 0 \) and \( \delta' > 0 \), there exists \( \delta > 0 \) and \( D^* \) s.t. the following holds for all \( N \geq D \geq D^* \).

Every \( r \)-uniform hypergraph \( H = (V, \mathcal{E}) \) on \( N \) vertices such that
(1) All vertices \( x \) except at most \( \delta N \) satisfy \( d(x) = (1 \pm \delta)D \).
(2) For all \( x \in V \), \( d(x) < KD \).
(3) For each pair \( (x, y) \) of distincts vertices, \( d(x, y) < \delta D \).

contains a set \( \mathcal{E}' \) of hyperedges s.t.
(iv) \( |\mathcal{E}'| = \frac{N}{r} \epsilon(1 \pm \delta') \).
(v) \( V' = V \setminus \bigcup_{S \in \mathcal{E}'} S \) has size \( Ne^{-\epsilon}(1 \pm \delta') \).
(vi) All vertices \( x \) of \( V' \) except at most \( \delta' |V'| \) the degree \( d'(x) \) of \( x \) in \( H[V'] \) satisfies \( d'(x) = De^{-\epsilon(r-1)}(1 \pm \delta') \).
Proving the Theorem with the Nibble Lemma

\( a > 0 \). Take \( \delta, \epsilon \) very small, and \( p \) s.t. \( e^{-\epsilon p} < \epsilon \).

\( K_i = \kappa e^{i(r-1)} \), \( D_i = De^{-\epsilon i(r-1)} \) with \( D \geq d_0 \).

Choose \( \delta = \delta_p > \delta_{p-1} > \cdots > \delta_0 \) s. t. we can apply the Nibble Lemma each time and \( \delta_{i-1} \leq \delta_i e^{-\epsilon(r-1)} \).

Step \( i \): \( H_{i-1} \rightarrow H_i \).

(1) All \( x \) but at most \( \delta_{i-1} N_{i-1} \) satisfy \( d(x) = (1 \pm \delta_{i-1}) D_{i-1} \).

(2) For all \( x \in V_{i-1} \), \( d(x) < K_{i-1} D_{i-1} \).

(3) For each pair \( (x, y) \) of distincts vertices, \( d(x, y) < \delta_{i-1} D_{i-1} \).

A set \( E'_i \) of hyperedges s.t.

(iv) \( |E'_i| = \frac{|V_{i-1}|}{r} \epsilon(1 \pm \delta_i) \).

(v) \( V_i = V_{i-1} \setminus \bigcup_{S \in E_i} S \) has size \( |V_{i-1}| \epsilon^{-\epsilon}(1 \pm \delta_i) \).

(vi) All vertices \( x \) of \( V_i \) except at most \( \delta_i |V_i| \) the degree of \( x \) in \( H_i \) satisfies \( d(x) = D_{i-1} e^{-\epsilon(r-1)}(1 \pm \delta_i) = (1 \pm \delta_i) D_i \).
Proving the Theorem with the Nibble Lemma

\[ a > 0. \]  Take \( \delta, \epsilon \) very small, and \( p \) s.t. \( e^{-\epsilon p} < \epsilon \).

\[ K_i = \kappa e^{i(r-1)}, \quad D_i = D e^{-i(r-1)} \text{ with } D \geq d_0. \]

Choose \( \delta = \delta_p > \delta_{p-1} > \cdots > \delta_0 \) s.t. we can apply the Nibble Lemma each time and \( \delta_{i-1} \leq \delta_i e^{-\epsilon (r-1)}. \)

\[ P = \prod_{i=0}^{p} (1 + \delta_i) \leq \prod_{i=0}^{p} (1 + \delta e^{-i\epsilon (r-1)}) \leq \frac{1 + 4\delta}{1 + 2\delta}. \]

\[ |V_i| \leq N e^{-i\epsilon} P \leq N e^{-i\epsilon} (1 + 2\delta). \]

So \( |\mathcal{E}_i'| = \left( \frac{\epsilon}{r} |V_{i-1}| \right) (1 \pm \delta_i) \leq \frac{\epsilon}{r} N e^{-(i-1)\epsilon} (1 + 2\delta) P \]

\[ \leq \frac{\epsilon}{r} N e^{-(i-1)\epsilon} (1 + 4\delta). \]
Proving the Theorem with the Nibble Lemma

$a > 0$. Take $\delta, \epsilon$ very small, and $p$ s.t. $e^{-p} < \epsilon$.

$$|V_i| \leq Ne^{-i\epsilon}(1 + 2\delta) \quad \text{and} \quad |E'_i| \leq \frac{\epsilon}{r} Ne^{-(i-1)\epsilon}(1 + 4\delta).$$

Cover of size $\sum_{i=1}^{p} |E'_i| + |V_p|$.

$$(1 + 4\delta)\frac{\epsilon}{r} N \sum_{i=0}^{p-1} e^{-i\epsilon} + |V_p| \leq (1 + 4\delta)\frac{\epsilon N}{r} \frac{1}{1 - e^{-\epsilon}} + (1 + 2\delta) Ne^{-p}$$

$$\leq \frac{N}{r} (1 + 4\delta) \left( \frac{1}{1 - e^{-\epsilon}} + r\epsilon \right)$$

$$< (1 + a)\frac{N}{r}.$$
Proving the Nibble Lemma

\( r > 2 \) fixed.
For \( K \geq 1, \epsilon > 0 \) and \( \delta' > 0 \), there exists \( \delta > 0 \) and \( D^* \) s.t. the following holds for all \( N \geq D \geq D^* \).

Every \( r \)-uniform hypergraph \( H = (V, E) \) on \( N \) vertices such that

1. All vertices \( x \) except at most \( \delta N \) satisfy \( d(x) = (1 \pm \delta)D \).
2. For all \( x \in V \), \( d(x) < KD \).
3. For each pair \( (x, y) \) of distinct vertices, \( d(x, y) < \delta D \).

contains a set \( E' \) of hyperedges s.t.

4. \( |E'| = \frac{N}{r} \epsilon (1 \pm \delta') \).
5. \( V' = V \setminus \bigcup_{e \in E'} e \) has size \( Ne^{-\epsilon}(1 \pm \delta') \).
6. All vertices \( x \) of \( V' \) except at most \( \delta'|V'| \) the degree \( d'(x) \) of \( x \) in \( H[V'] \) satisfies \( d'(x) = De^{-\epsilon(r-1)}(1 \pm \delta') \).
Proving the Nibble Lemma

$\mathcal{E}'$ a random subset of $\mathcal{E}$: every $e \in \mathcal{E}$ is picked randomly, independently with probability $p = \epsilon/D$.

- With very high probability (say $\geq 0.9$) (iv) holds.
- With very high probability (say $\geq 0.9$) (v) holds.
- With very high probability (say $\geq 0.9$) (vi) holds.

$\implies$ with positive probability ($\geq 0.7$), (iv), (v) and (vi) hold.
With very high probability, (iv) holds

(iv) \(|E'| = \frac{N}{r} \epsilon (1 \pm \delta')\).

(1): All vertices \(x\) except at most \(\delta N\) satisfy \(d(x) = (1 \pm \delta)D\).

\[
\Rightarrow \quad \frac{1}{r} (1 - \delta)^2 DN \leq |E| \leq \frac{1}{r} (1 + \delta) DN + \delta KDN
\]

\[
|E| = (1 \pm \delta_1) \frac{DN}{r}.
\]

\[
E(|E'|) = |E|/\rho = (1 \pm \delta_1) \frac{\epsilon N}{r}
\]

Need to prove that \(|E'|\) is concentrated around its expected value.
2nd Moment Method

Variance : \( \text{Var}(X) = E \left( (X - E(X))^2 \right) \).

Chebyshev Inequality : For any \( \lambda \),

\[
\Pr \left( |X - E(X)| \geq \lambda \sqrt{\text{Var}(X)} \right) \leq \frac{1}{\lambda^2} .
\]

\[
\Pr \left( X = E(X) \pm \lambda \sqrt{\text{Var}(X)} \right) \geq 1 - \frac{1}{\lambda^2} .
\]
Variance and covariance

Assume $X = X_1 + \cdots + X_m$.

$$\text{Var}(X) = \sum_{i=1}^{m} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$$

Covariance : $\text{Cov}(Y, Z) = \mathbb{E}(YZ) - \mathbb{E}(Y) \mathbb{E}(Z)$.

If $Y$ and $Z$ are independent, then $\text{Cov}(Y, Z) = 0$.

If $X_i = 1$ with probability $p_i$ and 0 otherwise, then

$$\text{Var}(X_i) = p_i(1 - p_i) \leq p_i = \mathbb{E}(X_i)$$

$$\text{Var}(X) \leq \mathbb{E}(X) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$$
With very high probability, (iv) holds

(iv) $|\mathcal{E}'| = \frac{N}{r} \epsilon (1 \pm \delta').$

(1): All vertices $x$ except at most $\delta N$ satisfy $d(x) = (1 \pm \delta) D.$

$$\implies \quad \frac{1}{r} (1 - \delta)^2 DN \leq |\mathcal{E}| \leq \frac{1}{r} (1 + \delta) DN + \delta KDN$$

$$|\mathcal{E}| = (1 \pm \delta_1) \frac{DN}{r}.$$

$$\mathbb{E}(|\mathcal{E}'|) = |\mathcal{E}|/p = (1 \pm \delta_1) \frac{\epsilon N}{r}$$

$$\text{Var}(|\mathcal{E}'|) = |\mathcal{E}| p (1 - p) \leq (1 \pm \delta_1) \frac{\epsilon N}{r}.$$

By Chebyshev Inequality,

$$\Pr \left( |\mathcal{E}'| = (1 \pm \delta_2) \frac{\epsilon N}{r} \right) > 0.9.$$
With very high probability, (v) holds

\[(v) \quad V' = V \setminus \bigcup_{e \in \mathcal{E}'} e \text{ has size } Ne^{-\epsilon}(1 \pm \delta').\]

\[I_x = 1 \text{ if } x \not\in \bigcup_{e \in \mathcal{E}'} e \text{ and } I_x = 0 \text{ otherwise.}\]

\[|V'| = \sum_{x \in V} I_x.\]

\[x \text{ good if } d(x) = (1 \pm \delta D), \text{ bad otherwise.}\]

\[x \text{ good: } E(I_x) = \Pr(I_x = 1) = (1 - p)^{d(x)} = (1 - \frac{\epsilon}{D})^{(1\pm\delta D)} = e^{-\epsilon}(1 \pm \delta_3).\]

\[x \text{ bad: } 0 \leq E(I_x) \leq 1, \text{ but at most } \delta N \text{ bad vertices.}\]

**Linearity of the Expected Value:** \[E(|V'|) \leq Ne^{-\epsilon}(1 \pm \delta_4).\]
With very high probability, (v) holds

\[
\Var(|V'|) = \sum_{x \in V} \Var(I_x) + \sum_{x,y \in V, x \neq y} \Cov(I_x, I_y)
\leq \E(|V'|) + \sum_{x, y \in V, x \neq y} \Cov(I_x, I_y)
\]

\[
\Cov(I_x, I_y) = \E(I_x I_y) - \E(I_x) \E(I_y)
= (1 - p)^{d(x)+d(y)-d(x,y)} - (1 - p)^{d(x)+d(y)}
\leq (1 - p)^{-d(x,y)} - 1 \leq \left(1 - \frac{\epsilon}{D}\right)^{-\delta(D)} - 1 \leq \delta_5.
\]

\[
\Var(|V'|) \leq \E(|V'|) + \delta_5 N^2 \leq \delta_6 \left(\E(|V'|)\right)^2.
\]

**Chebyshev** : With proba \( \geq 0,9 \),

\[
|V'| = (1 \pm \delta_7) \E(|V'|) = (1 \pm \delta_8) Ne^{-\epsilon}.
\]