The Rödl-Nibble method

April 23, 2026

Exercise 1. Coupon Collector

Suppose you are collection coupons, coloured with colours $\{1, \dots C\}$. Prove that for every $\varepsilon > 0$,

- after collecting $(1+\varepsilon)C \ln C$ coupons, one collected coupons of every colour with probability $1 - o(1)$ (as C tends to infinity)
- after collecting $(1-\varepsilon)C \ln C$ coupons, misses one colour with propability $1 - o(1)$.

Solution 1.

Let X be the number of missing coupons. For $c \in [C]$, let I_c be the indicator variable that the coupon colour c is missing. We have by linearity of expectation, $\mathbb{E}[X] = \sum_c \mathbb{E}[I_x] = C(1 - \frac{1}{C})^{(1\pm\varepsilon)C\ln C}$

- If we collected a total of $(1+\varepsilon)C \ln C$ coupons,

$$\begin{aligned} \mathbb{E}[X] &= C(1 - \frac{1}{C})^{(1+\varepsilon)C\ln C} \\ &\leq Ce^{-(1+\varepsilon)\ln C} && \text{using } 1 + x \leq e^x \\ &\leq C^{-\varepsilon} = o(1) \end{aligned}$$

  So by first moment method, $X = 0$ with probability $1 - o(1)$

- If we collected a total of $(1-\varepsilon)C \ln C$ coupons,

$$\begin{aligned} \mathbb{E}[X] &= C(1 - \frac{1}{C})^{(1-\varepsilon)C\ln C} \\ &\geq C\left(e^{-1/C} - \frac{1}{2C^2}\right)^{(1-\varepsilon)C\ln C} && \text{using } 1 + x + x^2/2 \geq e^x \qquad \geq C(1-o(1))e^{-(1-\varepsilon)\ln C} \\ &\geq (1 - o(1))C^{\varepsilon} \end{aligned}$$

  The Simple concentration bound and Talagrand's inequality do not work here. We use Chebyshev's inequality instead. The variables I_c are negatively correlated: $\mathbb{P}[I_c = 1 | I_d = 0] \geq \mathbb{P}[I_c = 1]$. The variance of X is

$$\begin{aligned} \operatorname{Var}[X] &= C \operatorname{Var}[I_c] + \binom{C}{2} \operatorname{Cov}[I_c, I_d] \\ &\leq C\mathbb{E}[I_c] = \mathbb{E}[X] && \text{Because the covariance is negative} \end{aligned}$$

  Hence, By Chebyshev's inequality, $\mathbb{P}[X \leq \mathbb{E}[X]/2] = \frac{4}{\mathbb{E}[X]} = O(n^{-\varepsilon})$, so $\mathbb{P}[X > 0] = 1 - o(1)$.

A second possibility is to explicit X as a sum of geometric random variables.

1 Independent sets in triangle-free graphs of bounded average degree

The average degree of a graph G is $\operatorname{avgdeg}(G) = \frac{2|E(G)|}{|V(G)|}$.
Ajtai, Komlós and Szemerédi proved in 1981 the following theorem:

Theorem 1. *Every triangle-free graph G on n vertices, with average degree at most d has an independent set of size $\Omega(\frac{n \log d}{d})$.*

Exercise 2.

Give an example of graphs showing that the triangle-free assumption is necessary.

Solution 2.

A collection of $n/(d+1)$ cliques of size d has average degree d and independence number $n/(d+1) = o(n \log d/d)$.

Exercise 3.

Give a rough sketch of how to use the Nibble method to prove Theorem 1.

Solution 3.

We repeatedly apply the following procedure:
Let S be a random subset of vertices obtained by selecting each vertex independently at random with probability p (to be adjusted later). Let I be the set of isolated vertices in $G[S]$. The set I is independent and w.h.p, $|I|$ will be close to εn. Let G' be the graph induced by $V(G) \setminus N(S)$. By concentration, w.h.p, G' will have around $(1-\varepsilon)dn$ vertices and average degree around d.

Exercise 4.

Deduce an upper bound on the off-diagonal Ramsey number $R(3,t)$ from Theorem 1.

Solution 4.

Let $n = O(t^2/\log t)$. Let G be a triangle-free n-vertex graph. If G has no independent set of size t, than avgdeg$(G) \leq \Delta(G) < t$. Hence, by Theorem 1, G has a independent set of size $O(n \log t/t) = t$, a contradiction.

Exercise 5. The Nibble step of Theorem 1

Let G be a n-vertex graph with average degree d and maximum degree $\Delta \leq 10d$. Prove that there exists a subset S of vertices such that the following conditions are verified.

- $|S| \geq \frac{n}{100d}$
- The set S spans at most $\frac{|S|}{50}$ edges,
- At least $n/2$ vertices are at distance at least 2 from S.
- Let H be the graph induced by the vertices that are neither in S nor neighbours of S. Let n', e' and d' be the number of vertices, of edges and the average degree of H. We have

$$\frac{n'^2}{2e'} = \frac{n'}{d'} > \nu\frac{n}{d} = \frac{n^2}{2e} \quad \text{where } \nu = 1 - 1/d - O(\sqrt{d/n})$$

Hint: For the last point, use the following inequality. For every n-vertex triangle-free graph of maximum degree Δ, for every $x \in [0, \frac{1}{10\Delta}]$,

$$\frac{1}{|E(G)|} \cdot \sum_{uv \in E(G)} e^{-x(\deg(u)+\deg(v))} \leq \left(\frac{1}{n} \cdot \sum_{u \in V(G)} e^{-x\deg(u)}\right)^2 \tag{1}$$

Solution 5.

Let S be a set of vertices chosen independently at random with probability $p = \frac{1}{10d}$. Let X count the number of edges spanned by S. Let H be the graph induced by the vertices that are neither in S nor neighbours of S. Let n' and e' be count the number of vertices and edges in H. We will prove that

1. $\mathbb{E}[|S|] = \frac{n}{10d}$ and for every δ, $\mathbb{P}(||S| - \mathbb{E}[|S|]| > \delta\mathbb{E}[S]) = o(1)$
2. $\mathbb{E}[X] = \frac{n}{200d}$ and for every δ, $\mathbb{P}(|X - \mathbb{E}[X]| > \delta\mathbb{E}[X]) = o(1)$

3. for every δ, $\mathbb{P}(n' < (1-\delta)\mathbb{E}[n']) = o(1)$
4. for every δ, $\mathbb{P}(e' > (1+\delta)\mathbb{E}[e']) = o(1)$

1. The size of S is a binomial random variable of parameters n and p, so $\mathbb{E}[|S|] = \frac{n}{10d}$ and by Chernoff's bound, for every δ,
$$\mathbb{P}\left(||S| - \mathbb{E}[S]| \geq \delta \mathbb{E}[S]\right) \leq 2e^{-\frac{\delta^2 n}{10d}} = o(1)$$
In particular, $|S| \geq \frac{n}{100d}$ w.h.p.
2. For every edge uv of G, let X_{uv} be the indicator variable that u and v both belong to S. The expectation of X is by linearity of expectation $\mathbb{E}[X] = p^2|E(G)| = \frac{n}{200d}$. The variance of X is
$$\begin{aligned}\mathrm{Var}[X] &= \sum_{uv \in E(G)} \mathrm{Var}[X_{uv}] + \sum_{uv,uw \in E(G)} \mathrm{Cov}[X_{uv}, X_{uw}] \\ &\leq |E(G)|p^2 + \left(\sum_{u \in V(G)} \binom{\deg(u)}{2}\right) \cdot (p^3 - p^4) \\ &\leq \mathbb{E}[X] + \frac{n\Delta^2}{2} p^3 \\ &\leq \frac{n}{200d} + \frac{n}{20d} \leq \frac{n}{10d} = \frac{\mathbb{E}[X]}{20}\end{aligned}$$
So by Chebyshev's inequality, for every δ,
$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \delta \mathbb{E}[X]) \leq \frac{1}{20\delta^2 \mathbb{E}[X]} = \frac{10d}{n\delta^2} = o(1)$$
In particular, the set S spans at most $\frac{|S|}{50}$ edges w.h.p.
3. The expectation of n' is by linearity of expectation $\mathbb{E}[n'] = \sum_{u \in V(G)} (1-p)^{\deg(u)+1}$. We have
$$\begin{aligned}\mathbb{E}[n'] &\geq n - \mathbb{E}[|S|] - \mathbb{E}[|N(S)|] \\ &\geq n - \frac{n}{10d} - (1 - (1-p)^2)|E(G)| \\ &\geq n - \frac{n}{10d} - 2p|E(G)| \\ &\geq 0.8n\end{aligned}$$

   Compute the variance of n'

4. The expectation of e' is by linearity of expectation $\mathbb{E}[e'] = \sum_{uv \in E(G)} (1-p)^{\deg(u)+1+\deg(v)+1}$. By (1),
$$\mathbb{E}[e'] \leq |E(G)|(1-p)^2 \left(\frac{1}{n} \sum_{u \in V(G)} (1-p)^{\deg(u)}\right)^2 \leq \frac{|E(G)|}{n^2} \mathbb{E}[n']^2$$

   Compute the variance of e'

Let $\delta = 1000\sqrt{t/n}$ and let S such that the four conditions are verified. We have

$$\begin{aligned}\frac{n'^2}{2e'} &\geq \frac{(1-\delta)^2 \mathbb{E}[n']^2}{(1+\delta) 2\mathbb{E}[e']} \\ &\geq (1 - 3\delta - o(\delta)) \frac{\mathbb{E}[n']^2}{2\mathbb{E}[e']} \\ &\geq (1 - 3\delta - o(\delta)) \frac{n^2}{|E(G)|}\end{aligned}$$

2 Rödl's theorem on approximate block designs

★ Exercise 6. Pippenger

The goal of this exercise is to prove the Nibble step of Pippenger's theorem. Prove the following:

For every integer $r \geq 2$, reals $K \geq 1$, $\varepsilon > 0$ and $\delta' > 0$, there exists $\delta = \delta(r, K, \varepsilon, \delta')$ and $D_0 = D_0(r, K, \varepsilon, \delta')$ such that the following holds. For every n and D such that $n \geq D \geq D_0$, for every n-vertex r-uniform hypergraph $\mathcal{H}$ such that

(i) For every vertex $x \in V(\mathcal{H})$ but at most δn of them, $\deg(x) = (1 \pm \delta)D$,

(ii) For every vertex $x \in V(\mathcal{H})$, $0 < \deg(x) < KD$

(iii) For every distinct vertices x and $y \in V(\mathcal{H})$, $\deg(x, y) \leq \delta D$

there exists a set E' of edges such that

1. $|E'| = (1 \pm \delta')\frac{\varepsilon n}{r}$
2. The set $V' = V(\mathcal{H}) \setminus \bigcup_{e \in E'} e$ has size $(1 \pm \delta')ne^{-\varepsilon}$
3. In the subhypergraph $\mathcal{H}'$ induced by V', for every vertex x but at most $\delta'|V'|$ of them, $\deg_{\mathcal{H}'}(x) = (1 \pm \delta')De^{-\varepsilon(r-1)}$.

★ Solution 6.

See Alon Spencer page 58.