The Container Method

Case in point: Mantel theorem for random graphs.

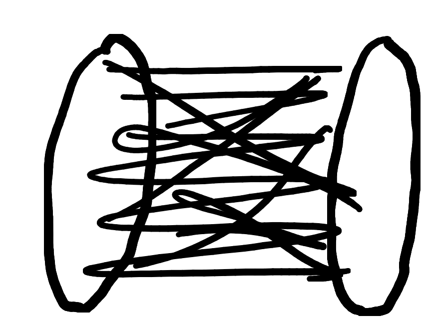
$$\mathrm{ex}(G, H) := \max\left(e(G') : G' \subseteq G \text{ and } H \not\subseteq G\right).$$

when $G = K_n$, we write $\mathrm{ex}(n, H)$.

Mantel's theorem 1907: $\mathrm{ex}(n, K_3) = \left\lceil \frac{n^2}{4} \right\rceil$.

Turán's theorem 1940's: $\mathrm{ex}(n, K_{r+1}) = \left(1 - \frac{1}{r} + o(1)\right)\frac{n^2}{2}$

The largest triangle free graph is $K_{\frac{n}{2},\frac{n}{2}}$



$$\boxed{\text{What is } \mathrm{ex}(G(n,p), \triangle)\ ?}$$

Observations: 1. $\mathrm{ex}(G(n,p), \triangle) \geq \mathrm{Bin}\left(\frac{n^2}{4}, p\right)$

by intersecting $G(n,p)$ and $K_{\frac{n}{2},\frac{n}{2}}$.

So $\mathrm{ex}(G(n,p), \triangle) \geq \left(\frac{1}{4} + o(1)\right) p n^2$ w.h.p.

2. Triangle removal process:

Let T be the random variable counting the number of triangles in $G(n,p)$.

If $\mathbb{E}[T] = o\left(\mathbb{E}\left[e(G(n,p))\right]\right)$, then we can remove one edge in every triangle. We get a subgraph $G' \subseteq G(n,p)$

with no triangles and almost the same number of edges:

if $\underbrace{\mathbb{E}[T]}_{p^3 \frac{n^3}{6}} \ll \underbrace{\mathbb{E}[e(G(n,p))]}_{p\frac{n^2}{2}}$ then $ex(G(n,p), K_3) \geq p\binom{n}{2}(1+o(1))$

$$\geq \left(\frac{1}{2}+o(1)\right)pn^2$$

i.e. $p \ll n^{-1/2}$

And this is sharp because $G(n,p)$ has typically $\left(\frac{1}{2}+o(1)\right)pn^2$ edges!

$\rightarrow$ What happens when we don't have $p \ll n^{-1/2}$?

Frankel - Rödl 1986: If $p \gg n^{-1/2}$, then with probability $1-o(1)$, $ex(G(n,p), K_3) = \left(\frac{1}{4}+o(1)\right)pn^2$.

1) The usual approach: First moment method

Let X_m be the random variable counting the number of triangle-free subgraphs of $G(n,p)$ with $m = \left(\frac{1}{4}+o(1)\right)pn^2$ edges.

If $\mathbb{E}[X_m] = o(1)$ then by Markov's inequality,

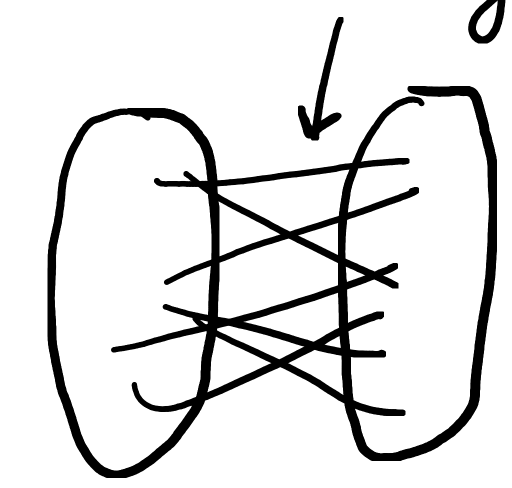
$$\mathbb{P}\left(ex(G(n,p), K_3) \leq \left(\frac{1}{4}+o(1)\right)pn^2\right)$$

$$= \mathbb{P}(G(n,p) \text{ has no triangle-free subgraph with } m \text{ edges})$$

$$= \mathbb{P}(X_m = 0) = 1 - \mathbb{P}(X_m \geq 1) \geq 1 - \frac{\mathbb{E}[X_m]}{1} = 1 - o(1)$$

$\rightarrow$ Let's compute $\mathbb{E}[X_m]$.

Let Y_m be the number of subgraphs of $K_{\frac{n}{2},\frac{n}{2}}$ with m edges.



We have $X_m \geq Y_m$ so

$$\mathbb{E}[X_m] \geq \mathbb{E}[Y_m] = \binom{\frac{n^2}{4}}{m} p^m$$

Linearity of expectation

$$= \left((1+o(1))\frac{e n^2}{4m}\right)^m p^m$$

Stirling

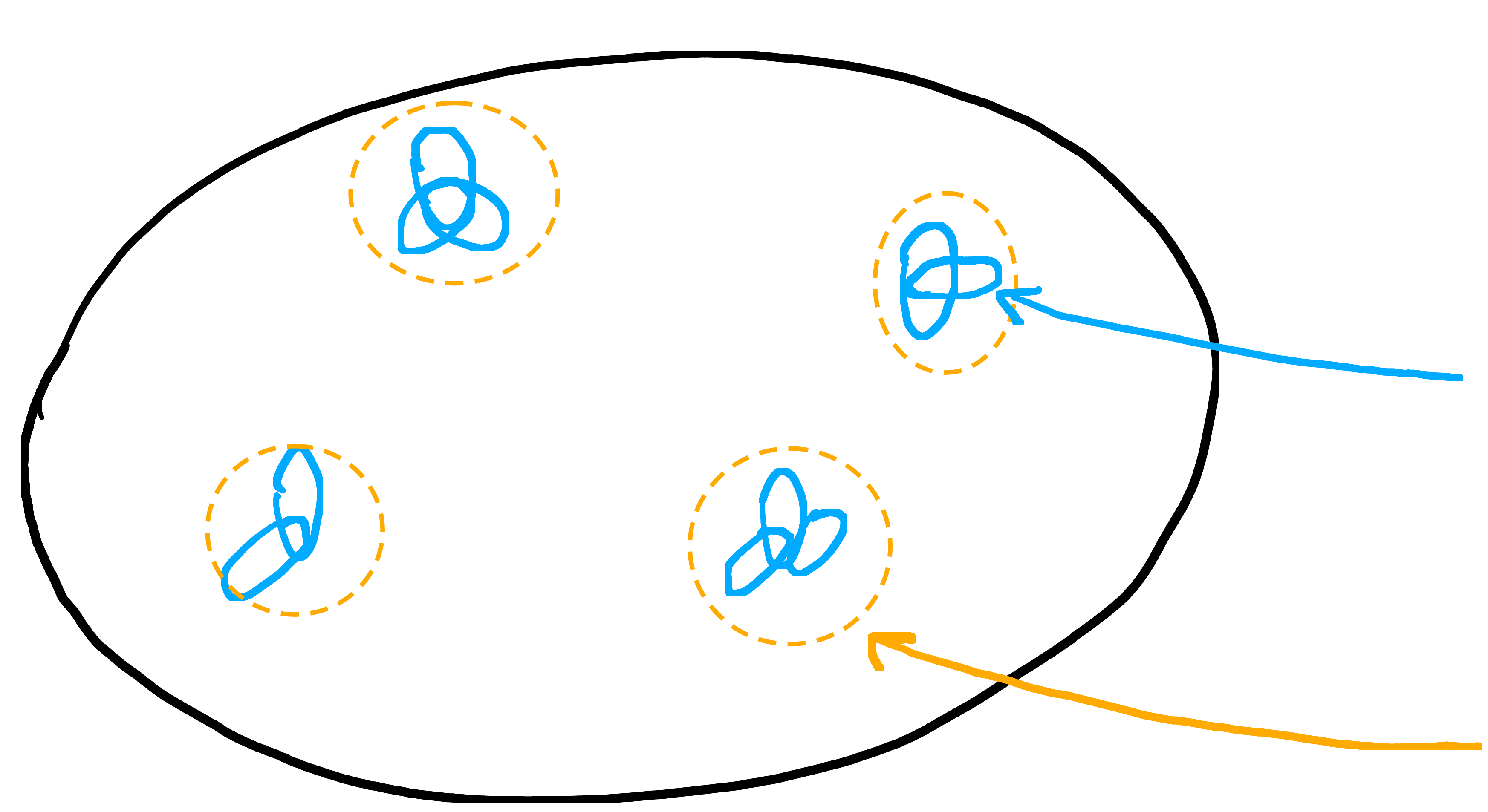
$$= \left((1+o(1))e\right)^m \gg 1.$$

So the first moment method does not work here: there are many large triangle-free graphs.

What should we do when the first moment method fails i.e. $\mathbb{E}[X_m]$ is large?

Option 1: Maybe what we are trying to prove is false: maybe X_m is close to $\mathbb{E}[X_m]$ $\rightarrow$ second moment method.

Option 2: Maybe X_m is large but the probability of containing such a subgraph is still small.



Space of subgraphs of K_n

triangle free subgraphs

concentrated in a small number of clusters called containers

2) A weak container lemma for triangle-free subgraphs

Theorem: (weak container lemma for Δ-free subgraphs):
For every $n \geq 0$, there exists $\mathcal{G}_n$ a collection of graphs on n vertices such that:

(i) $|\mathcal{G}_n| \leq n^{O(n^{3/2})}$

(ii) Each container $G \in \mathcal{G}_n$ contains $o(n^3)$ triangles

(iii) Each triangle-free graph is contained in some $G \in \mathcal{G}_n$

Before proving Frankl-Rödl theorem, we need one more thing:

Supersaturation lemma for triangles: For every $\varepsilon > 0$, there exist $\gamma > 0$ such that every n-vertex graph with at least $\left(\frac{1}{4} + \varepsilon\right)n^2$ edges has at least γn^3 triangles.

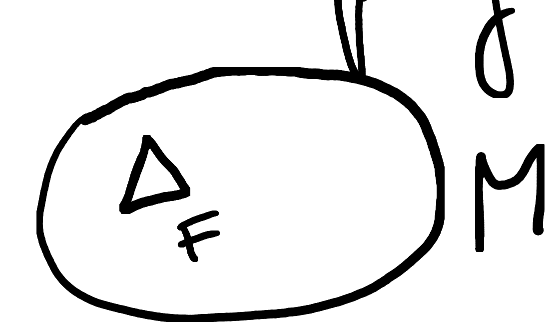
Proof 1: Regularity lemma & Graph removal lemma.

Proof 2: Goodman's bound (double counting + convexity)

Proof 3: Let C be a large constant and M be a random subset of C vertices s.t.

$$\mathbb{P}(G[M] \text{ contains a triangle}) = \delta > 0$$

Let F be a random triple of vertices. Sample F by first sampling M before sampling $F \subseteq M$.



$$\begin{aligned}
\#\text{triangles in } G &= \mathbb{P}(F \text{ is a triangle}) \cdot \binom{n}{3} \\
&= \binom{n}{3} \mathbb{P}(F \text{ is a triangle} \mid K_3 \subseteq G[M]) \cdot \mathbb{P}(K_3 \subseteq G[M]) \\
&= \binom{n}{3} \binom{C}{3}^{-1} \delta = \Omega(n^3).
\end{aligned}$$

$\blacksquare$

Proof that $p \gg \frac{\log n}{\sqrt{n}} \implies ex(G(n,p), K_3) \leq \left(\frac{1}{4}+o(1)\right) p n^2$

Let $\varepsilon > 0$ and $m = \left(\frac{1}{4}+\varepsilon\right) p n^2$.

Goal: $\mathbb{P}(\exists H \subseteq G(n,p)$ with no triangles and $e(H) = m) = o(1)$.

By supersaturation, for every $G \in \mathcal{G}_n$, $e(G) \leq \left(\frac{1}{4}+o(1)\right) n^2$.

Let X_G be the number of edges in $G \cap G(n,p)$.

$X_G \sim Bin(e(G), p)$ so $\mathbb{E}[X_G] = p\, e(G)$

$$\mathbb{P}(X_G \geq m) \leq e^{\Omega(-\varepsilon p n^2)} \text{ by Chernoff bound.}$$

$\mathbb{P}(G(n,p)$ contains a triangle-free graph H with m edges$)$

$$\underset{\substack{\text{union bound} \\ \text{on } \mathcal{G}_n}}{\leq} \sum_{G \in \mathcal{G}_n} \mathbb{P}(X_G \geq m)$$

$$\leq n^{O(n^{3/2})} e^{-\Omega(\varepsilon p n^2)} = o(1)$$

if $p \gg \frac{\log n}{\sqrt{n}}$. ∎

3) Encoding the problem with a hypergraph:

Let $\mathcal{H}$ be the 3-uniform hypergraph on the vertex set $E(K_n)$ with one hyperedge $\{e_1, e_2, e_3\}$ if $(e_i)_{i \in [3]}$ forms a triangle of K_n.

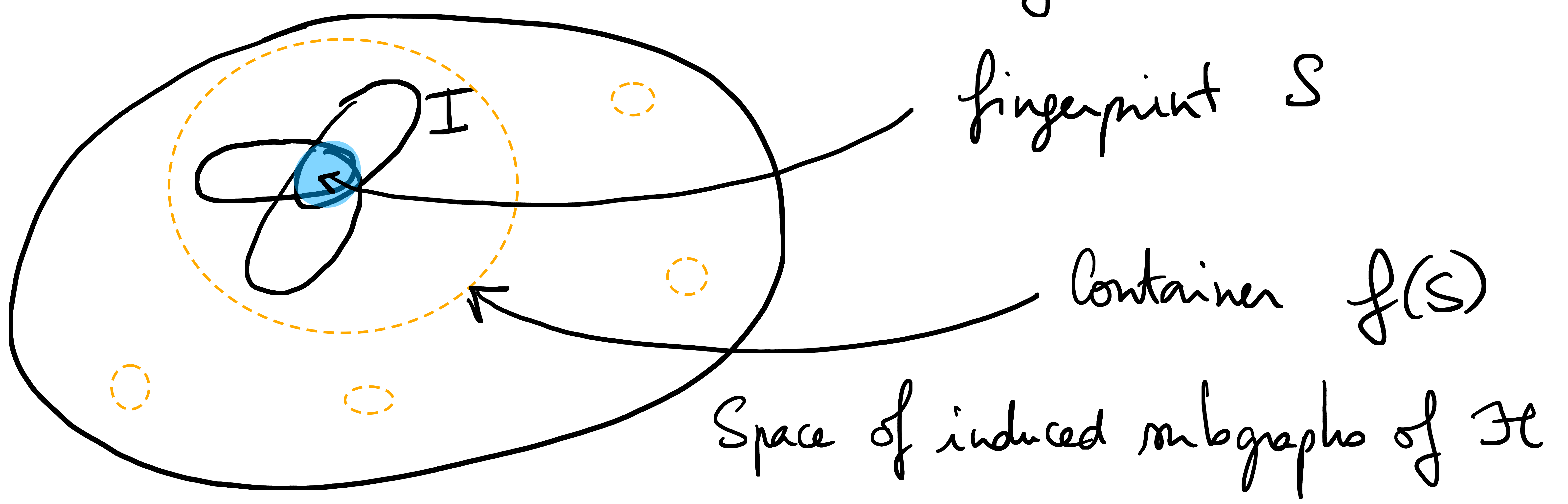
A triangle free subgraph of K_n corresponds to an independent set of $\mathcal{H}$.

4) A strong hypergraph container lemma

Theorem (Strong 3-uniform hypergraph container lemma).
For every $c > 0$, there exists $\delta > 0$ such that:
For every hypergraph $\mathcal{H}$ of average degree $\geq d$, with $\Delta_1(\mathcal{H}) \leq cd$ and $\Delta_2(\mathcal{H}) \leq c\sqrt{d}$,
there exists a collection $\mathcal{C}$ of subsets of $V(\mathcal{H})$ and a function $f : \{0,1\}^{V(\mathcal{H})} \to \mathcal{C}$ s.t.

(i) For every independent set I of $\mathcal{H}$, there exists a fingerprint S with
$$S \subseteq I \subseteq f(S) \quad \text{and} \quad |S| \leq \frac{v(\mathcal{H})}{\sqrt{d}}.$$

(ii) $|C| \leq (1-\delta)\, v(\mathcal{H})$ for every $C \in \mathcal{C}$.



Theorem (Strong container lemma for Δ-free graphs).
For every $\gamma > 0$
There exists a collection $\mathcal{G}_n$ of containers C and a function $f : \{0,1\}^{E(K_n)} \to \mathcal{G}_n$ with

(i) $|\mathcal{G}_n| \leq n^{O(n^{3/2})}$

(ii) Every $C \in \mathcal{G}_n$ has at most γn^3 triangles.

(iii) Every triangle-free graph G has a fingerprint S such that $S \subseteq G \subseteq f(S)$ and $e(S) = O(n^{3/2})$

<u>Proof</u>: we have $e(\mathcal{H}) = \binom{n}{3} \sim \frac{n^3}{6}$ so $\mathrm{avgdeg}(\mathcal{H}) \approx \frac{n^3/6}{n^2/2} = \frac{n}{3}$.

More generally, every induced subgraph $\mathcal{H}'$ of $\mathcal{H}$ with at least γn^3 hyperedges has average degree at least $2\gamma n = d$

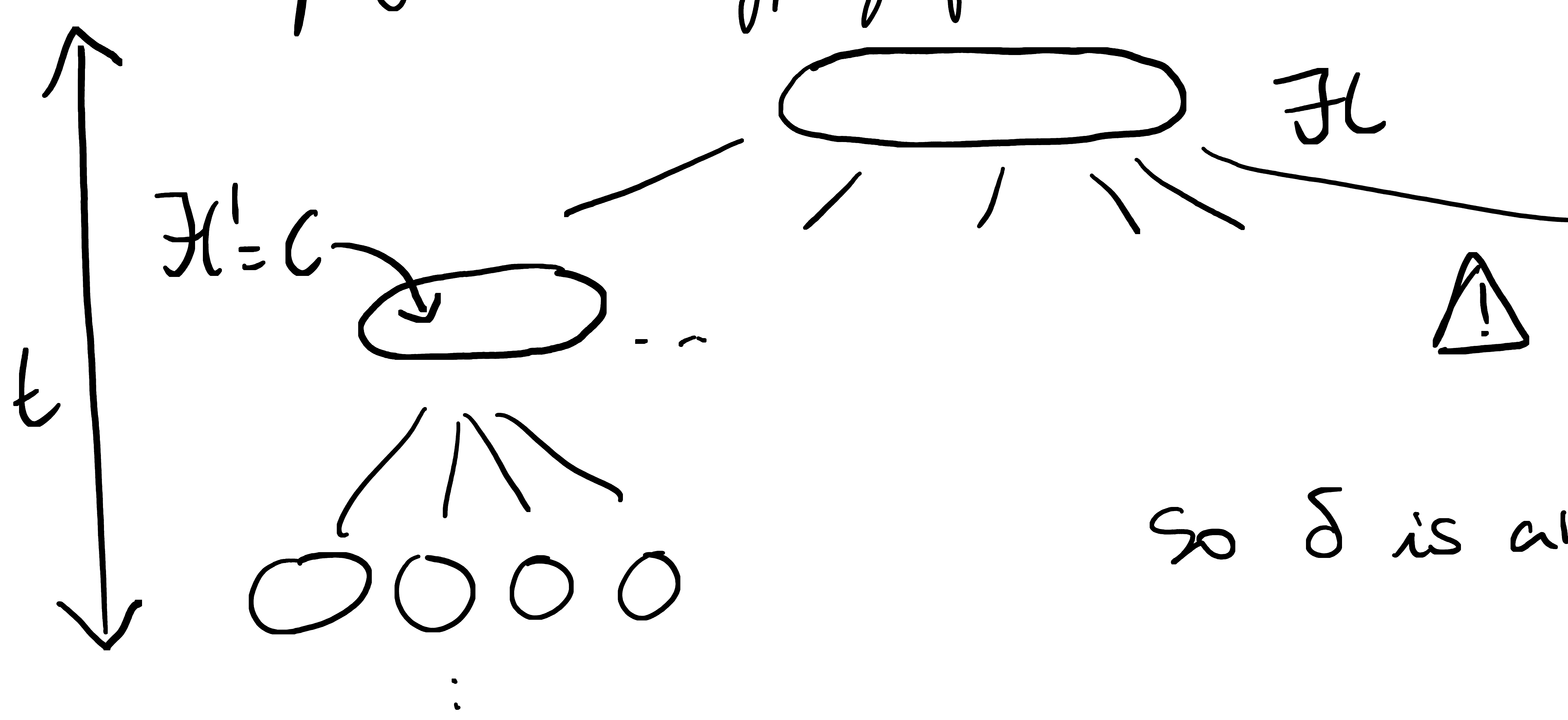
- $\Delta_1(\mathcal{H}') \leq \Delta_1(\mathcal{H}) = n-2 \leq \frac{1}{2\gamma} d$
- $\Delta_2(\mathcal{H}') \leq \Delta_2(\mathcal{H}) = 1$.

We apply the 3-uniform hypergraph container lemma with $c = \frac{1}{2\gamma}$.

We get at most $\sum_{k=1}^{\binom{n}{2}/\sqrt{d}} \binom{\binom{n}{2}}{k} = n^{O(n^{3/2})}$ containers.

By (ii), each container has at most $(1-\delta)\binom{n}{2}$ edges.

<u>Key idea</u>: if some container C has more than γn^3 triangles, we apply the hypergraph container lemma to $\mathcal{H}' = \mathcal{H}[E(C)]$.



⚠ $c = \frac{1}{2\gamma}$ all the time!

So δ is an absolute constant.

After t recursive applications, each container has

- at most γn^3 triangle
- or at most $(1-\delta)^t \binom{n}{2}$ edges, and hence at most $\Delta_1 \cdot (1-\delta)^t \binom{n}{2} \leq \frac{(1-\delta)^t}{2} n^3 \leq \gamma n^3$ triangles

for t a large enough constant.

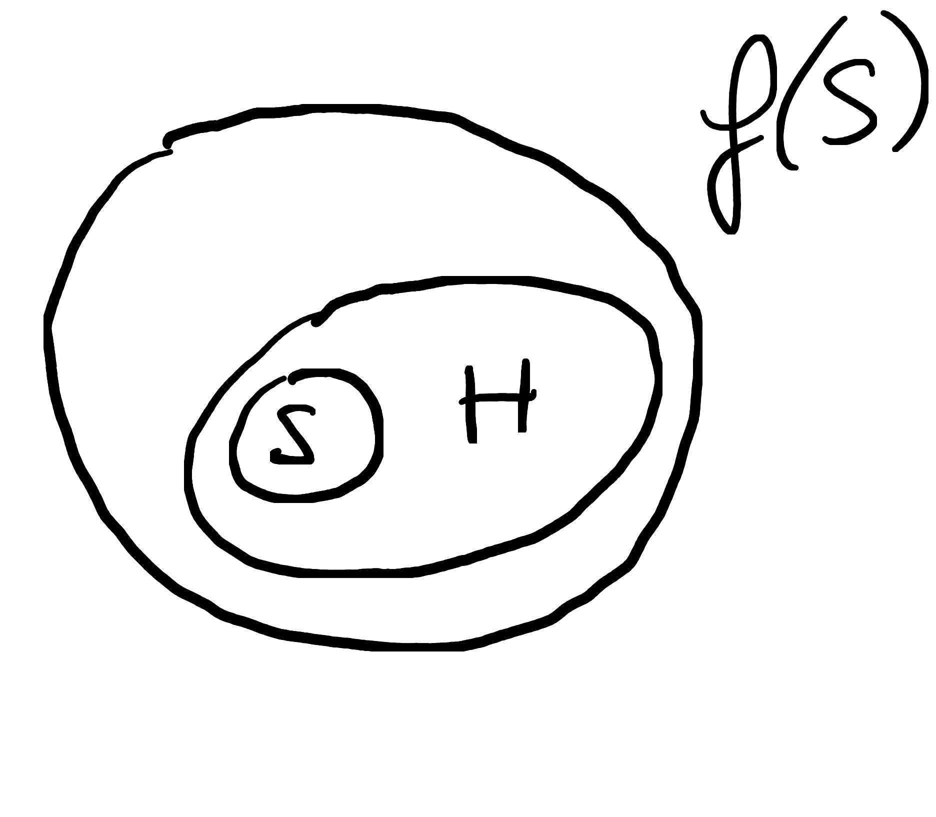
Moreover, $|\mathcal{G}_n| \leq \left(n^{O(n^{3/2})}\right)^t \leq n^{O(n^{3/2})}$ ▨

5) Proof of Frankl–Rödl (without the log factor)

Let $\varepsilon > 0$ and $m = \left(\frac{1}{4} + \varepsilon\right) p n^2$.

For every fingerprint S the probability that $G(n,p)$ contains a triangle-free subgraph H that has m edges and fingerprint S is at most

$$\underbrace{\mathbb{P}\left(S \subseteq G(n,p)\right)}_{= p^{|S|}} \cdot \mathbb{P}\left(e\left(f(S) \cap G(n,p) - S\right) \geq m - e(S)\right)$$



As $f(S)$ contains at most $o(n^3)$ triangles, $f(S)$ has at most $\left(\frac{1}{4} + \frac{\varepsilon}{2}\right) n^2$ edges by supersaturation.

Let $X = e\left(f(S) \cap G(n,p) - S\right) \sim \mathrm{Bin}\left(e(f(S) \setminus S), p\right)$

so $\mathbb{E}[X] \leq \left(\frac{1}{4} + \frac{\varepsilon}{2}\right) p n^2 - e(S)$ $\quad e(S) \ll m$ because $p \gg n^{-1/2}$

$$\mathbb{P}\left(X \geq m - e(S)\right) \leq \mathbb{P}\left(X \geq \left(\frac{1}{4} + \frac{\varepsilon}{2}\right) p n^2\right)$$
$$\leq e^{-\Omega(\varepsilon p n^2)} \quad \text{by Chernoff bound.}$$

So $\mathbb{P}(\exists H \subseteq G(n,p)$ with no triangle and $e(H) = m)$

$$\underset{\text{union bound on } S}{\leq} \sum_{k=1}^{O(n^{3/2})} \binom{\binom{n}{2}}{k} p^k e^{-\Omega(\varepsilon p n^2)}$$

$$\underset{\text{Stirling}}{\leq} \sum_{k=1}^{O(n^{3/2})} \left(\frac{O(p n^2)}{k}\right)^k e^{-\Omega(\varepsilon p n^2)} = o(1)$$

for $p \gg n^{-1/2}$. $\blacksquare$

6) Property testing:

G is ε-close from having property Π if the addition or removal of εn^2 edges gives a graph satisfying Π.

Canonical ε-tester for Π: Monte Carlo algorithm that samples s vertices and guess whether G satisfies Π

- if G satisfies Π, $\quad \mathbb{P}(\text{YES}) \geq 2/3$
- if G is ε-far from Π, $\quad \mathbb{P}(\text{NO}) \geq 2/3$

Sample complexity

Blais - Seth 2025 The sample complexity of ρ-independent set is $\tilde{O}\left(\frac{\rho^3}{\varepsilon^2}\right)$.

Sketch of proof: Sample $|S| = s$ vertices. Answer yes if $G[S]$ has an independent set of size ρs.

- If G has a ρn-independent set, by concentration, it is likely that $G[S]$ has a ρs-independent set.

- Container lemma: If G is ε-far from having a ρn-independent set, then every independent set I has a fingerprint F with

$$F \subseteq I \subseteq f(F)$$

$f(F)$: size $\leq \left(\rho - |F| \dfrac{\varepsilon}{8\rho \ln(2\rho\varepsilon^{-1})}\right) n$

F: size $\leq \dfrac{8\rho^2 \ln(2\rho\varepsilon^{-1})}{\varepsilon}$

S contains a large independent set I if $F \subseteq I$ and <u>many</u> vertices of $f(F)$ are contained in I.

$\blacksquare$

7) Other nice things that can be proved with containers

- Generalisation : Turán in random graphs.
- Counting : $\#\Delta$-free graphs with m edges $\leq \sum_S \binom{f(S)}{m-|S|}$
- Choosability, list colourings:

Saxton–Thomason 2015 : Let $\mathcal{H}$ be a k-uniform hypergraph with average degree d and $\Delta_2(\mathcal{H}) = 1$. Then

$$\chi_\ell(\mathcal{H}) \geq \left(\frac{1}{(k-1)^2} + o(1)\right) \log_k d \quad \text{as } d \to +\infty$$

- Ramsey for random graphs:
Every 2-colouring of the edges of K_6 has a monochromatic K_3.
Does there exist a K_4-free graph with this property?

Frankl–Rödl 1986 & Rödl–Ruciński 1994:
For every $r \in \mathbb{N}^*$, there exist $C > 0$ s.t. if $p \gg C n^{-1/2}$, then with probability $1-o(1)$, every r-colouring of $G(n,p)$ has a monochromatic triangle.

- Additive combinatorics: thresholds for k-AP / Sidon set / Schur triples free subsets of $[n]_p$.
- Discrete geometry :

Balogh–Solymosi : There exists a set S of n points in $\mathbb{R}^2$ with no 4 points on a line s.t. every subset of S of size $n^{5/6-o(1)}$ contains 3 points on a line.

8) Sketch of proof of the container lemma:

For simplicity, we sketch the proof of the following container lemma:

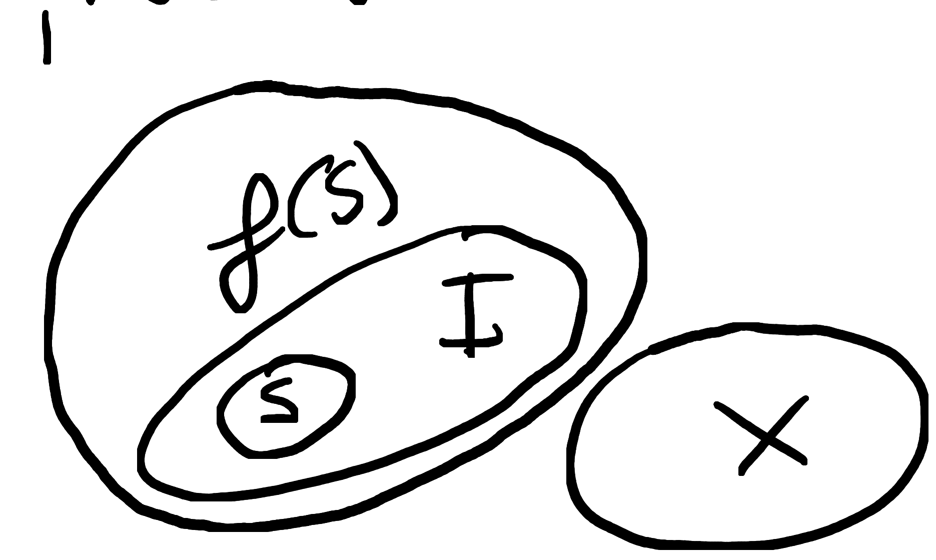
Strong graph container lemma:
For every $c>0$, there exists $\delta>0$ s.t. for every graph G of average degree d with $\Delta(G)\leq cd$, there exists $\mathcal{G}$ and $f: 2^{V(G)} \to \mathcal{G}$ s.t.

(i) For every independent set I of G, there exists a fingerprint S with $S \subseteq I \subseteq f(S)$ and $|S| \leq \frac{2\delta n}{\sqrt{d}}$

(ii) For every $C \in \mathcal{G}$, $|C| \leq (1-\delta) n$.

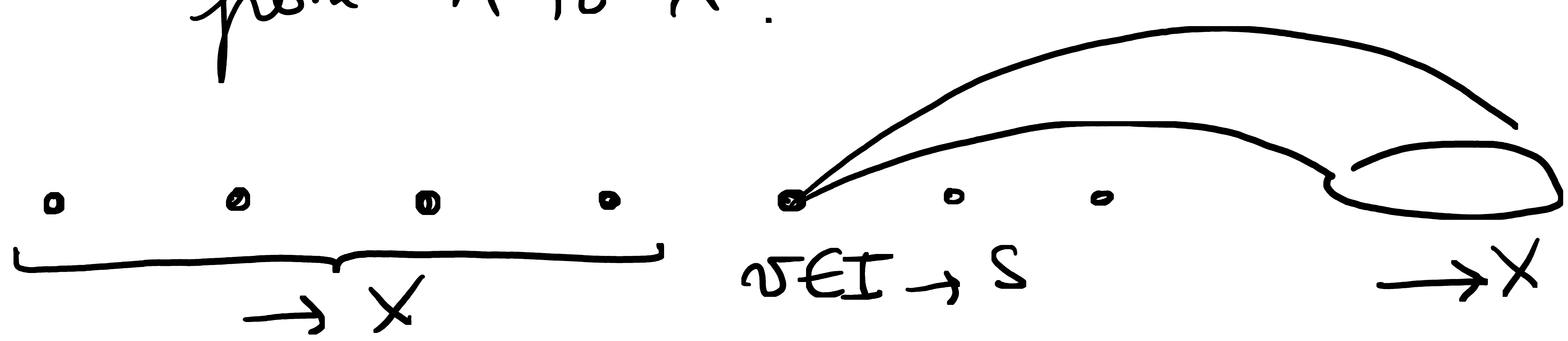
__Sketch of proof:__ Given I an independent set, we build a partition (S, A, X) of $V(G)$.

$$\underbrace{S}_{\text{fingerprint of } I}, \quad \underbrace{A}_{\text{active vertices}}, \quad \underbrace{X}_{V(G)\setminus f(S)}$$



At first $S=\emptyset$, $A=V(G)$, $X=\emptyset$.

While $|X| \leq \delta n$:
1. Let v be the first vertex of $I\cap A$ in the max degree order on $G[A]$.
2. Move v from A to S
3. Move the vertices before v, and the neighbours of v from A to X.



Claim: for every vertex added to S, we add at least $\frac{d}{2}$ vertices to X. $\blacksquare$

References:

- Series of videos and lecture notes of Rob Morris for the minicourse on container method at Discrete math summer school 2021 in Chile
- The method of hypergraph container (J. Balogh, R. Morris, W. Samotij 2018)

Thanks!