

# Extremal Combinatorics Notes

Wilson Pan

February 16, 2026

## Contents

<b>1</b>	<b>Jan 12</b>	<b>2</b>
<b>2</b>	<b>Jan 14</b>	<b>4</b>
<b>3</b>	<b>Jan 16</b>	<b>6</b>
<b>4</b>	<b>Jan 21: Supersaturation and Generalized Turán Problems</b>	<b>9</b>
4.1	Supersaturation . . . . .	9
4.2	Hypergraph Kővári-Sós-Turán . . . . .	9
4.3	The Erdős-Stone Theorem . . . . .	10
4.4	Degenerate Turán Numbers (Bipartite Graphs) . . . . .	10
4.5	Extremal Numbers for Trees . . . . .	11
<b>5</b>	<b>Jan 23</b>	<b>12</b>
5.1	Turan Number of Cycles and Girth Problem . . . . .	12
5.2	Turan Number for Paths . . . . .	12
<b>6</b>	<b>Jan 26: Projective Geometry</b>	<b>13</b>
<b>7</b>	<b>Jan 28</b>	<b>16</b>
<b>8</b>	<b>Regularity Lemmas: Feb 4</b>	<b>18</b>
<b>9</b>	<b>Feb 6</b>	<b>19</b>
<b>10</b>	<b>Feb 13</b>	<b>21</b>
<b>11</b>	<b>Feb 16</b>	<b>23</b>

# 1 Jan 12

**Definition 1.1.** *Turan Numbers (Forbidden Subgraph Problems):*  $F$  is a graph and  $G$  graph is  $F$ -free if  $G$  contains no copy of  $F$  as a subgraph.

We want to maximize the size of  $G$  subject to  $G$  being  $F$ -free. Where  $\text{size} = e(G) = \# \text{ edges in } G$

**Definition 1.2.**  $ex(n, F) = \max\{e(G) | G \text{ is } F\text{-free and } G \text{ is } n\text{-vertex graph}\}$

**Theorem 1.3.** *Crude Turan's Theorem:* If  $G$  is  $n$ -vertex graph,

$$e(G) = \epsilon \binom{n}{2}.$$

then  $G$  contains an independent set of size greater than or equal  $\frac{1}{2} \frac{1}{\epsilon} + \frac{1}{2}$

In particular, if  $e(G) \geq (1 - \epsilon) \binom{n}{2}$  edges, then  $G$  contains a clique on greater than or equal  $\frac{1}{2\epsilon} + \frac{1}{2}$  vertices.

$$ex(n, K_t) \leq \left(1 - \frac{1}{2t-1}\right) \binom{n}{2}.$$

*Proof.* Pick  $k$  vertices in  $G$  at random without replacement. Say we pick  $v_1, \dots, v_k$

For each edge between  $v_i$  and  $v_j$ , delete  $v_i$ .

Let  $X = \#$  deleted vertices

Note:  $X \leq e(G[v_1, \dots, v_k]) =: Y$

$$\mathbb{E}[Y] = \sum_{1 \leq i \leq j \leq k} \mathbb{P}[v_i v_j \in e(G)] = \binom{k}{2} \frac{\epsilon \binom{n}{2}}{\binom{n}{2}} = \epsilon \binom{k}{2}.$$

Thus, there is a choice of  $v_1, \dots, v_k$  such that  $Y \leq \epsilon \binom{k}{2}$ . Fix it. Let  $\alpha(G) =$  size of largest independent set in  $G$  satisfies  $\alpha(G) \geq k - Y \geq k - \epsilon \binom{k}{2}$

$$\left(k - \epsilon \binom{k}{2}\right)' = 1 + \frac{\epsilon}{2} - \epsilon k.$$

Set  $k = \frac{1+\frac{\epsilon}{2}}{2} + \delta$  for some  $|\delta| \leq \frac{1}{2}$ .

$$\begin{aligned} \alpha(G) &\geq \frac{1}{\epsilon} + \frac{1}{2} + \delta - \epsilon \left( \left( \frac{1}{\epsilon} + \frac{1}{2} + \delta \right) \left( \frac{1}{\epsilon} - \frac{1}{2} + \delta \right) \right) / 2 \\ &= \frac{1}{\epsilon} + \frac{1}{2} + \delta - \epsilon \frac{(\frac{1}{\epsilon} + \delta)^2 - \frac{1}{4}}{2} \\ &= \frac{1}{\epsilon} + \frac{1}{2} + \delta - \epsilon \frac{(1/\epsilon)^2 + 2\frac{1}{\epsilon}\delta + \delta^2 - \frac{1}{4}}{2} \\ &= \frac{1}{2\epsilon} + \frac{1}{2} + \frac{\frac{1}{4} - \delta^2}{2} \geq \frac{1}{2\epsilon} + \frac{1}{2} \end{aligned}$$

□

**Remark 1.4.** Recall:  $k$ -uniform hypergraph on vertex set  $V$  is a subset of

$$\binom{V}{k} := \{e \subseteq V : |e| = k\}.$$

**Example 1.5.** In  $k$ -uniform hypergraph on vertex with  $\epsilon \binom{n}{k}$  edges there is an independent set of size  $\geq c_k \frac{1}{\epsilon^{\frac{1}{k-1}}}$

**Theorem 1.6.** Mantel's Theorem:  $ex(n, K_3) = \lfloor \frac{n^2}{4} \rfloor$

More generally, define Turan's graph  $T_{n,r}$  to be the  $n$ -vertex complete  $r$ -partite graph with part sizes as equal as possible.

Note:  $K_{r+1} \not\subseteq T_{n,r}$  because among  $r+1$  vertices some pair is in same part.

**Theorem 1.7.** *Turan's Theorem:  $ex(n, K_{r+1}) = e(T_{n,r})$  and furthermore  $T_{n,r}$  is the only extremizer.*

$$e(T_{n,r}) = \left(1 - \frac{1}{r}\right) \binom{n}{2} + O(n).$$

*Proof.* Let  $G$  be a graph on  $n$  vertices that is  $K_{r+1}$ -free and has the maximum possible number of edges (an extremizer). We proceed in two steps.

Suppose  $G$  is not a complete multipartite graph. Then non-adjacency is not an equivalence relation, implying there exist two non-adjacent vertices  $x, y$  that do not have identical neighborhoods.

Assume  $\deg(x) \leq \deg(y)$ . Delete  $x$  and then clone  $y$  to obtain  $y'$ . The new graph  $G'$  has

$$e(G') = e(G) - \deg(x) + \deg(y) \geq e(G).$$

Also,  $G'$  is  $K_{r+1}$ -free because any copy of  $K_{r+1}$  has at most one of  $y$  or  $y'$  (since they are non-adjacent). If a clique contains  $y'$ , it can be swapped for  $y$  (since  $N(y') \subseteq N(y)$ ), which would imply  $K_{r+1} \subseteq G$ , a contradiction.

Repeating this process transforms  $G$  into a complete multipartite graph without decreasing edges. Thus, the extremizer must be a complete multipartite graph.

Since  $G$  is  $K_{r+1}$ -free, it must be  $k$ -partite with  $k \leq r$ . To maximize edges, we set  $k = r$ . Let the parts have sizes  $n_1, \dots, n_r$ .

$$e(G) = \binom{n}{2} - \sum_{i=1}^r \binom{n_i}{2}$$

To maximize  $e(G)$ , we must minimize the subtracted sum of pairs inside the parts. This occurs when  $n_i$  are as equal as possible ( $|n_i - n_j| \leq 1$ ). If  $n_i \geq n_j + 2$ , moving a vertex from part  $i$  to part  $j$  increases the edge count by  $(n_i - 1) - n_j > 0$ .

Thus, the unique extremizer is the balanced complete  $r$ -partite graph,  $T_{n,r}$ . □

## 2 Jan 14

*Proof.* Second proof of Turan's Theorem with induction on  $r$ .

(Induction Step) Let  $x$  be a vertex of largest degree, let  $A = N(x)$  be its neighborhood.

$$B := V(G) \setminus A.$$

We have  $e(G) = e(A) + e(A, B) + e(B)$  and  $A$  is  $K_{r-1}$ -free so  $e(A) \leq e(T_{|A|, r})$

Each vertex in  $B$  has  $\deg \leq \deg(x) = |A|$

So,

$$e(A, B) + e(B) \leq \sum_{v \in B} \deg(v) \leq |B||A|.$$

So,

$$e(G) \leq e(T_{|A|, r}) + |A||B| \leq e(T_{n, r}).$$

The right hand side is equal to number of edges in complete  $r$ -partite graph with one part of size  $|B|$  and other being as equal as possible.

Equality holds iff  $B$  is independent and  $A$  must be a copy of  $e(T_{|A|, r-1})$ . Hence  $G$  is  $T_{n, r}$  for equality to hold.  $\square$

**Theorem 2.1.** Observe that if  $F$  is a graph then

$$\frac{ex(n, F)}{\binom{n}{2}} \geq \frac{ex(n+1, F)}{\binom{n+1}{2}}.$$

*Proof.* Pick  $n$  vertices in an  $F$ -free  $(n+1)$  vertex graph  $G$  at random without replacement; let  $G'$  be the induced graph on these  $n$  vertices.

$$e(G') \leq ex(n, F).$$

$$\mathbb{E}(e(G')) = \binom{n}{2} \frac{e(G)}{\binom{n+1}{2}} \implies \frac{e(G)}{\binom{n+1}{2}} \leq \frac{ex(n, F)}{\binom{n}{2}}.$$

$\square$

**Definition 2.2.**

$$\pi(F) := \lim_{n \rightarrow \infty} \frac{ex(n, F)}{\binom{n}{2}}.$$

The Turan density of  $F$  is

$$\pi(K_{r+1}) = 1 - \frac{1}{r}.$$

**Theorem 2.3.** Let  $F$  be any graph and let  $\epsilon > 0$ . Let  $G$  be an  $n$ -vertex graph. There is a  $\delta > 0$  such that

$$e(G) \geq (\pi(F) + \epsilon) \binom{n}{2} \text{ and } n \geq n_0(F, \epsilon).$$

then there are at least  $\delta n^{|V(F)|}$  copies of  $F$ .

*Proof.* Let  $n_0$  be large enough such that  $\frac{ex(n_0, F)}{\binom{n_0}{2}} \leq \pi(F) + \frac{\epsilon}{2}$ . Given  $n$ -vertex  $G$  as in theorem. Pick  $n_0$  vertices at random without replacement. Let  $G'$  be the  $n_0$ -vertex graph induced on these chosen vertices.

$$\mathbb{E}[e(G')] = \frac{e(G)}{\binom{n}{2}} \binom{n_0}{2} \geq (\pi(F) + \epsilon) \binom{n_0}{2}.$$

Say  $G'$  is good if

$$e(G') \geq \left(\pi(F) + \frac{\epsilon}{2}\right) \binom{n_0}{2}.$$

$$\begin{aligned} \mathbb{E}(e(G')) &\leq \mathbb{P}(G' \text{ is bad}) \left(\pi(F) + \frac{\epsilon}{2}\right) \binom{n_0}{2} + \mathbb{P}(G' \text{ is good}) \binom{n_0}{2} \\ &\leq \left(\pi(F) + \frac{\epsilon}{2}\right) \binom{n_0}{2} + \mathbb{P}(G' \text{ is good}) \binom{n_0}{2} \end{aligned}$$

Then  $\frac{\epsilon}{2} \binom{n_0}{2} \leq \mathbb{P}(G' \text{ is good}) \binom{n_0}{2}$  so  $\mathbb{P}(G' \text{ is good}) \geq \frac{\epsilon}{2}$   
 Pick  $|V(F)|$  many vertices without replacement inside  $G'$ , call the subgraph they induce  $G''$ . So

$$G \supset G' \supset G''.$$

$$\mathbb{P}(G'' \text{ contains } F) \geq \mathbb{P}(G' \text{ is good}) \cdot \mathbb{P}(G'' \text{ contains } F | G' \text{ is good}) \geq \frac{\epsilon}{2} \frac{1}{\binom{n_0}{|V(F)|}}.$$

Note that  $G''$  is uniformly randomly chosen  $|V(F)|$ -vertex subset of  $G$ . This implies the # copies of  $F$  in  $G \geq \mathbb{P}(G'' \text{ contains } F) \binom{n}{|V(F)|} \geq P(\dots) \frac{n^{|V(F)|} (1-o(1))}{|V(F)|!}$ . We can choose  $\delta = \frac{1}{\binom{n_0}{|V(F)|}}$   $\square$

### 3 Jan 16

**Theorem 3.1** (Erdős-Stone-Simonovits). *Let  $T_{n,r}$  denote the Turán graph. We know that:*

$$ex(n, T_{s,(r+1)}) = \left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2}.$$

*This implies that if  $H$  is any graph with chromatic number  $\chi(H) = r + 1$ , then*

$$ex(n, H) \leq \left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2}.$$

*Conversely, if  $\chi(H) > r$  (i.e.,  $H$  is not  $r$ -partite), then  $T_{n,r}$  is  $H$ -free. Hence,*

$$ex(n, H) \geq e(T_{n,r}) = \left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2}.$$

*Combining these, for any non-bipartite graph  $H$ :*

$$ex(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \binom{n}{2}.$$

*It is important to note that this theorem holds for bipartite graphs as well, where  $\chi(H) = 2$ . However, substituting  $r = \chi(H) - 1 = 1$  gives a coefficient of  $1 - 1/1 = 0$ . This implies that  $ex(n, H) = o(n^2)$ , but it does not provide the precise asymptotic behavior (e.g.,  $n^{2-\epsilon}$ ). Therefore, while valid, the Erdős-Stone-Simonovits theorem is not the primary tool for determining extremal numbers of bipartite graphs; we rely on theorems like Kővári-Sós-Turán for those bounds.*

**Theorem 3.2** (Kővári-Sós-Turán 1954).

$$ex(n, K_{s,t}) \leq \frac{1}{2}(t-1)^{\frac{1}{s}} \cdot n^{2-\frac{1}{s}} + O(n).$$

*Proof.* Let  $G$  be a  $K_{s,t}$ -free graph with  $n$  vertices. We will count  $X$ , the number of copies of  $K_{1,s}$  (stars with  $s$  leaves).



We count this in two ways:

1. Since there is no  $K_{s,t}$ , any set of  $s$  vertices can have at most  $t - 1$  common neighbors. Thus:

$$X \leq \binom{n}{s} (t - 1).$$

2. Let  $d_v$  be the degree of vertex  $v$ . Then:

$$X = \sum_{v \in V(G)} \binom{d_v}{s}.$$

A simple-minded application of Jensen's inequality here fails because the binomial coefficient  $\binom{x}{s}$  is not convex for all  $x$ . To fix this, we define the function:

$$f_s(x) = \begin{cases} \binom{x}{s} & \text{if } x \geq s - 1 \\ 0 & \text{otherwise.} \end{cases}$$

Note that  $f_s(x)$  is convex for all  $x$  because its derivative:

$$\frac{d}{dx}f_s(x) = \begin{cases} \left(\frac{1}{x} + \frac{1}{x-1} + \cdots + \frac{1}{x-s+1}\right) \binom{x}{s} & \text{if } x > s-1 \\ 0 & \text{otherwise} \end{cases}$$

is non-decreasing. Since  $d_v$  is an integer,  $\binom{d_v}{s} = f_s(d_v)$ . We can now validly apply Jensen's inequality:

$$X = \sum_{v \in V} f_s(d_v) \geq n f_s(d_{\text{avg}}) = n f_s\left(\frac{2m}{n}\right),$$

where  $m = e(G)$ .

Combining (1) and (2):

$$n f_s\left(\frac{2m}{n}\right) \leq \binom{n}{s} (t-1).$$

If  $m \leq \frac{1}{2}(s-1)n$ , the bound holds trivially. Otherwise, we approximate  $\binom{x}{s} \approx \frac{x^s}{s!}$ :

$$n \frac{\left(\frac{2m}{n} - s\right)^s}{s!} \leq (t-1) \frac{n^s}{s!}.$$

Rearranging yields:

$$\frac{2m}{n} \leq (t-1)^{\frac{1}{s}} n^{1-\frac{1}{s}} + s \implies m \leq \frac{1}{2}(t-1)^{\frac{1}{s}} n^{2-\frac{1}{s}} + O(n).$$

□

**Example 3.3** (Unit Distance Problem). *How can we place  $n$  points in  $\mathbb{R}^2$  to maximize the number of pairs  $(p, q)$  such that  $\|p - q\| = 1$ ?*

*A strong construction is a  $\sqrt{n} \times \sqrt{n}$  grid. Rather than scaling the grid to unit step size, we scale it such that the distance 1 corresponds to a distance  $d$  in the integer grid that occurs most frequently.*

*The possible distances in an unscaled integer grid are  $\sqrt{a^2 + b^2}$  for  $0 \leq a, b < \sqrt{n}$ . By number theory (Landau-Ramanujan), integers representable as a sum of two squares are those where primes  $p \equiv 3 \pmod{4}$  appear with even exponents.*

*To get a better bound, we look for "highly composite" integers. Specifically, if we choose a number  $d$  composed of a product of many distinct primes  $p \equiv 1 \pmod{4}$ , the number of ways to write  $d$  as a sum of two squares is very large. Erdős used this to show there exists a distance occurring:*

$$n^{1 + \frac{c}{\log \log n}}.$$

*This is significantly larger than the  $n\sqrt{\log n}$  bound derived from the Pigeonhole Principle on generic integers.*

**Theorem 3.4** (Upper Bound for Unit Distances). *The number of unit distances is  $O(n^{3/2})$ .*

*Proof.* Given  $n$  points in  $\mathbb{R}^2$ , construct a graph where vertices are points and  $v \sim u$  if  $\|v - u\| = 1$ . Observe the graph is  $K_{2,3}$ -free. Geometrically, this is because two circles with radius 1 intersect at most 2 times; thus, no 2 vertices can share 3 common neighbors.

By KST with  $s = 2, t = 3$ :

$$e(G) \leq \frac{1}{2} \sqrt{3-1} \cdot n^{2-\frac{1}{2}} + o(n^2) = \frac{1}{\sqrt{2}} n^{\frac{3}{2}} + o(n^2).$$

□

**Definition 3.5.** Define  $K_{s_1, s_2, t}^{(3)}$  to be the complete 3-uniform, 3-partite hypergraph with parts of size  $s_1, s_2, t$ .

**Theorem 3.6** (Generalization of KST to Hypergraphs). *For the 3-uniform case:*

$$ex(n, K_{s_1, s_2, t}^{(3)}) \leq C_{s_1, s_2, t} n^{3 - \frac{1}{s_1 s_2}} + O(n^2).$$

*Proof.* We will count the number of stars  $K_{1,1,s_1}$  (sets of edges sharing a common pair of vertices). Let  $X$  be this number. For any pair  $\{v_1, v_2\}$ , let  $d_{v_1, v_2}$  be the co-degree (number of edges containing  $\{v_1, v_2\}$ ).

1. **Upper Bound on  $X$ :** Fix a set  $U$  of size  $s_1$ . Define an auxiliary graph  $G_U$  on  $V(G)$  where  $v_1 \sim v_2$  in  $G_U$  if  $\{v_1, v_2, u\} \in E(G)$  for all  $u \in U$ . If  $G_U$  contains a copy of  $K_{s_2, t}$ , then combined with  $U$ , we form a  $K_{s_1, s_2, t}^{(3)}$ . Thus,  $G_U$  must be  $K_{s_2, t}$ -free.

$$X \leq \binom{n}{s_1} ex(n, K_{s_2, t}) \leq C \binom{n}{s_1} n^{2 - \frac{1}{s_2}}.$$

2. **Lower Bound on  $X$ :** Summing over pairs:

$$X = \sum_{\{v_1, v_2\} \in \binom{V}{2}} \binom{d_{v_1, v_2}}{s_1} \geq \binom{n}{2} f_{s_1} \left( \frac{3m}{\binom{n}{2}} \right).$$

Combining bounds (assuming  $m$  is large enough):

$$\binom{n}{2} \frac{\left( \frac{3m}{n^2/2} \right)^{s_1}}{s_1!} \lesssim \frac{n^{s_1}}{s_1!} n^{2 - \frac{1}{s_2}}.$$

Simplifying:

$$n^2 \left( \frac{m}{n^2} \right)^{s_1} \lesssim n^{s_1 + 2 - \frac{1}{s_2}} \implies m^{s_1} \lesssim n^{3s_1 - \frac{1}{s_2}}.$$

Taking the  $s_1$ -th root:

$$m \leq C n^{3 - \frac{1}{s_1 s_2}}.$$

□



## 4 Jan 21: Supersaturation and Generalized Turán Problems

### 4.1 Supersaturation

The classical Turán problem asks for the maximum number of edges in a graph avoiding a subgraph  $F$ . Supersaturation asks: what happens if we have just *slightly* more edges than this threshold? It turns out we don't just get one copy of  $F$ , but exponentially many.

**Lemma 4.1** (Supersaturation Lemma). *If  $G$  is a graph on  $n$  vertices with  $e(G) \geq (\pi(F) + \epsilon) \binom{n}{2}$  edges, then  $G$  contains at least  $\delta(\epsilon)n^{|V(F)|}$  copies of  $F$ , where  $\delta(\epsilon) > 0$  is a constant independent of  $n$ .*

**Corollary 4.2.** *For the complete graph  $K_{r+1}$ , we know  $\pi(K_{r+1}) = 1 - \frac{1}{r}$ . Thus:*

$$e(G) \geq \left(1 - \frac{1}{r} + \epsilon\right) \binom{n}{2} \implies G \text{ contains } \delta(\epsilon)n^{r+1} \text{ copies of } K_{r+1}.$$

### 4.2 Hypergraph Kővári-Sós-Turán

We generalize the KST theorem from bipartite graphs to  $r$ -uniform hypergraphs. We are looking for the Turán number of the complete  $r$ -partite hypergraph  $K_{s_1, \dots, s_r}^{(r)}$ . Note that the theorem statement often simplifies the last parameter to  $t$ .

**Theorem 4.3** (Generalized Kővári-Sós-Turán). *Let  $K = K_{s_1, s_2, \dots, s_{r-1}, t}^{(r)}$  be the complete  $r$ -partite  $r$ -uniform hypergraph with part sizes  $s_1, \dots, s_{r-1}, t$ . Then:*

$$ex(n, K) = O_{s,t} \left( n^{r - \frac{1}{s_1 s_2 \dots s_{r-1}}} \right).$$

*Proof.* We proceed by induction on the uniformity  $r$ . The base case  $r = 2$  is the standard graph KST theorem. Assume the bound holds for  $(r-1)$ -uniform hypergraphs.

Let  $G$  be an  $r$ -uniform hypergraph with  $n$  vertices and  $m$  edges. We define a “hypergraph star”  $K_{1,1,\dots,1,s_1}^{(r)}$  as a set of  $s_1$  edges that all share a common core of  $r-1$  vertices.

Let  $X$  be the number of such stars (copies of  $K_{1,1,\dots,1,s_1}^{(r)}$ ) in  $G$ . We double count  $X$ .

**1. Lower Bound (Convexity):** For every subset  $S \in \binom{V(G)}{r-1}$  of size  $r-1$ , let the *co-degree*  $d_S$  be the number of edges in  $G$  containing  $S$ . A set  $S$  contributes  $\binom{d_S}{s_1}$  stars to the count.

$$X = \sum_{S \in \binom{V(G)}{r-1}} \binom{d_S}{s_1}.$$

Using Jensen's Inequality (and the convexity of  $\binom{x}{s_1}$ ), we bound this sum:

$$X \geq \binom{n}{r-1} \binom{\bar{d}}{s_1}, \quad \text{where } \bar{d} = \frac{\sum d_S}{\binom{n}{r-1}} = \frac{rm}{\binom{n}{r-1}}.$$

Approximating for large  $n$ :  $\bar{d} \approx \frac{r!m}{n^{r-1}}$  and  $\binom{x}{k} \approx \frac{x^k}{k!}$ , we get:

$$X \gtrsim n^{r-1} \cdot \left( \frac{m}{n^{r-1}} \right)^{s_1}. \quad (1)$$

**2. Upper Bound (Induction):** Fix a set of vertices  $V_1 = \{v_1, \dots, v_{s_1}\}$ . We ask: how many ways can we complete this into a star?

Define an auxiliary  $(r-1)$ -uniform hypergraph  $G_{V_1}$  on  $V(G)$ . A set  $e'$  of size  $r-1$  is an edge in  $G_{V_1}$  if  $e' \cup \{v_i\} \in E(G)$  for all  $i = 1, \dots, s_1$ . In other words,  $e'$  is in the common neighborhood of all  $v_i$ .

If  $G_{V_1}$  contains a copy of  $K_{s_2, \dots, s_{r-1}, t}^{(r-1)}$ , then combined with  $V_1$ , we would form a copy of the forbidden  $K_{s_1, \dots, t}^{(r)}$  in  $G$ . Thus,  $G_{V_1}$  must be free of  $K_{s_2, \dots, t}^{(r-1)}$ . By the inductive hypothesis:

$$e(G_{V_1}) \leq C \cdot n^{(r-1) - \frac{1}{s_2 \dots s_{r-1}}}.$$

Summing over all choices of  $V_1$  (there are  $\binom{n}{s_1}$  such sets):

$$X \leq \binom{n}{s_1} \cdot \max_{V_1} e(G_{V_1}) \lesssim n^{s_1} \cdot n^{r-1 - \frac{1}{s_2 \dots s_{r-1}}}. \quad (2)$$

**3. Conclusion:** Comparing (1) and (2):

$$n^{r-1} \left( \frac{m}{n^{r-1}} \right)^{s_1} \lesssim n^{s_1 + r - 1 - \frac{1}{s_2 \cdots s_{r-1}}}.$$

Rearranging for  $m$ :

$$m^{s_1} \lesssim n^{(r-1)s_1} \cdot n^{s_1 - \frac{1}{s_2 \cdots s_{r-1}}} = n^{rs_1 - \frac{1}{s_2 \cdots s_{r-1}}}.$$

Taking the  $s_1$ -th root:

$$m \lesssim n^{r - \frac{1}{s_1 s_2 \cdots s_{r-1}}}.$$

□

### 4.3 The Erdős-Stone Theorem

**Definition 4.4** (Blow-up). A **blow-up** of a graph  $H$ , denoted  $H(s)$ , is obtained by replacing each vertex  $v \in V(H)$  with an independent set  $I_v$  of size  $s$ , and replacing each edge  $(u, v) \in E(H)$  with a complete bipartite graph between  $I_u$  and  $I_v$ .

The Turán graph  $T_{n,r}$  is essentially a blow-up of  $K_r$  with parts of size  $n/r$ .

**Theorem 4.5** (Erdős-Stone). If  $e(G) \geq (1 - \frac{1}{r} + \epsilon) \binom{n}{2}$  and  $n$  is sufficiently large, then  $G$  contains  $T_{s,(r+1)}$  (a blow-up of  $K_{r+1}$  with parts of size  $s$ ).

Consequently,  $G$  contains any graph  $H$  with  $\chi(H) = r + 1$  (since any such  $H$  is a subgraph of a large enough blow-up of  $K_{r+1}$ ).

*Proof. Step 1: Find many cliques.* By the Supersaturation Corollary, since the edge density is greater than the Turán threshold for  $K_{r+1}$ ,  $G$  contains many copies of  $K_{r+1}$ :

$$\#K_{r+1} \geq \delta n^{r+1}.$$

**Step 2: Construct Auxiliary Hypergraph.** Define an  $(r + 1)$ -uniform hypergraph  $\mathcal{H}$  with vertex set  $V(G)$ . Let a set of  $r + 1$  vertices form an edge in  $\mathcal{H}$  if and only if they form a  $K_{r+1}$  in the original graph  $G$ .

From Step 1, we know  $e(\mathcal{H}) \geq \delta n^{r+1}$ . Since  $\delta$  is constant, this is a dense hypergraph.

**Step 3: Find a dense structure in the hypergraph.** We apply the **Hypergraph KST Theorem** to  $\mathcal{H}$ . Since  $\mathcal{H}$  is dense (order  $n^{r+1}$ ), it exceeds the KST threshold (which is roughly  $n^{(r+1)-\epsilon}$ ). Therefore,  $\mathcal{H}$  must contain a copy of the complete  $(r + 1)$ -partite hypergraph  $K_{s,s,\dots,s}^{(r+1)}$ .

**Step 4: Map back to  $G$ .** Let the parts of this hypergraph copy be  $U_1, \dots, U_{r+1}$ , each of size  $s$ . The definition of the complete hypergraph implies that for *any* choice of vertices  $u_1 \in U_1, \dots, u_{r+1} \in U_{r+1}$ , the set  $\{u_1, \dots, u_{r+1}\}$  is an edge in  $\mathcal{H}$ .

By definition of  $\mathcal{H}$ , this means  $\{u_1, \dots, u_{r+1}\}$  forms a clique  $K_{r+1}$  in  $G$ . If every possible tuple forms a clique, then every pair of vertices in distinct parts  $U_i, U_j$  must be connected in  $G$ .

Thus,  $G[U_1 \cup \dots \cup U_{r+1}]$  contains the complete multipartite graph with parts of size  $s$ , which is exactly  $T_{s(r+1)}$ . □

### 4.4 Degenerate Turán Numbers (Bipartite Graphs)

When  $H$  is bipartite ( $r = 1$ ), the Erdős-Stone bound gives  $1 - 1/1 = 0$ , which is trivial. We rely on KST bounds  $ex(n, K_{s,t}) = O(n^{2-1/s})$ . We examine the sharpness of these bounds.

**Theorem 4.6** (Tightness of KST).

1. **Case  $s = 1$  (Trees/Forests):**

$$ex(n, K_{1,t}) \leq \frac{1}{2}(t-1)n.$$

This is tight for a graph consisting of disjoint copies of  $K_t$  (or a matching if  $t = 1$ ).

2. **Case  $s = 2$  ( $C_4$ -free graphs):** We know  $ex(n, K_{2,2}) = O(n^{3/2})$ . This is tight.

**Construction (Finite Geometry):** Let  $q$  be a prime power. Construct a bipartite graph  $G$  with parts  $A$  and  $B$ , where  $|A| = |B| = q^2$ .

- Identify vertices in  $A$  and  $B$  with points in the affine plane  $\mathbb{F}_q^2$ .
- Define edge  $(a, b) \sim (x, y)$  if  $ax + by = 1$ , where  $(a, b) \in A$  and  $(x, y) \in B$ .

For a fixed vertex  $v = (a, b) \in A$ , its neighborhood is the set of solutions  $(x, y)$  to the linear equation  $ax + by = 1$ . This describes a line in  $\mathbb{F}_q^2$ , which contains  $q$  points (unless  $(a, b) = (0, 0)$ , which we can discard).

$$e(G) \approx q^2 \cdot q = q^3.$$

Since  $n = 2q^2$ , we have  $q \approx \sqrt{n/2}$ , so  $e(G) \approx (n/2)^{3/2} = \Theta(n^{3/2})$ .

**Why is it  $C_4$ -free?** A  $C_4$  corresponds to two vertices in  $A$  sharing two common neighbors in  $B$ . Geometrically, this means two distinct lines intersect at two distinct points. In a plane, two distinct lines intersect at most at one point. Thus, no  $C_4$  exists.

3. **Case  $s = 3$  ( $K_{3,3}$ -free):** Tightness  $ex(n, K_{3,3}) = \Theta(n^{5/3})$  is known.

**Construction Sketch:** Vertices are points in  $\mathbb{F}_q^3$ . Adjacency is defined by  $(x, y, z) \sim (a, b, c)$  if  $(x-a)^2 + (y-b)^2 + (z-c)^2 = 1$ . Geometrically, the neighborhood of a point is a unit sphere. If  $G$  contained  $K_{3,3}$ , three vertices in one part would share three common neighbors. This would imply that three unit spheres intersect at three points. In Euclidean geometry, three spheres intersect at most at 2 points (a circle intersects a sphere at 2 points). This logic carries over to  $\mathbb{F}_q$  provided  $-1$  is not a square.

4. **Case  $s = 4$  ( $K_{4,4}$ -free):** It is an **open problem** whether the KST bound  $O(n^{2-1/4})$  is tight for  $K_{4,4}$ .

## 4.5 Extremal Numbers for Trees

**Lemma 4.7.** If a graph  $G$  has average degree  $d$ , it contains a subgraph  $H$  with minimum degree  $\delta(H) \geq d/2$ .

*Proof.* Let  $n = |V(G)|$ . We have  $e(G) = \frac{dn}{2}$ . Iteratively delete any vertex with degree strictly less than  $d/2$ . Let  $S$  be the set of removed vertices. The number of edges removed is strictly less than  $|S| \cdot \frac{d}{2}$ . Even if we remove  $n-1$  vertices, the number of edges removed is  $< (n-1) \frac{d}{2} < \frac{dn}{2}$ . Thus, the edge set cannot be emptied. The process must stop at a non-empty subgraph where every remaining vertex has degree  $\geq d/2$ .  $\square$

**Theorem 4.8.** If  $T$  is a tree with  $r$  edges (so  $r+1$  vertices), then:

$$ex(n, T) \leq rn.$$

*Proof.* Suppose  $e(G) > rn$ . Then the average degree of  $G$  is  $d(G) = \frac{2e(G)}{n} > 2r$ . By the Lemma, there exists a subgraph  $H \subseteq G$  with minimum degree  $\delta(H) > r$ . Since degrees are integers,  $\delta(H) \geq r$ .

We embed  $T$  into  $H$  greedily:

1. Order the vertices of  $T$  as  $v_1, v_2, \dots, v_{r+1}$  such that each  $v_i$  (for  $i > 1$ ) has exactly one neighbor among  $\{v_1, \dots, v_{i-1}\}$ . (This is always possible by doing a BFS/DFS from a root).
2. Map  $v_1$  to any vertex  $u_1 \in V(H)$ .
3. Suppose we have embedded  $v_1, \dots, v_{i-1}$  as  $u_1, \dots, u_{i-1}$ . Let  $v_j$  be the parent of  $v_i$  in the tree ( $j < i$ ).
4. We need to map  $v_i$  to a neighbor of  $u_j$  in  $H$ .
5. The vertex  $u_j$  has at least  $r$  neighbors in  $H$ . We have used at most  $i-1$  vertices so far. Since  $i \leq r+1$ , we have used at most  $r$  vertices.
6. We only care that  $u_j$  has neighbors not already used in the current tree embedding. The degree of  $u_j$  is  $\geq r$ . The number of currently used vertices is  $i-1 \leq r$ .
7. The bound requires strict inequality or careful counting. If  $\delta(H) \geq r$ ,  $u_j$  has  $r$  neighbors. We have used  $i-1$  vertices total. If  $i-1 < r$ , there is space. If  $i-1 = r$ , we are placing the last vertex. The neighbors of  $u_j$  could be occupied by the other  $r-1$  vertices of the tree.
8. However, the conjecture is actually  $\frac{1}{2}(r)n$ . The bound  $rn$  is loose enough. If  $e(G) > rn$ ,  $\text{avg deg} > 2r$ ,  $\text{min deg} \geq r+1$ . Then  $u_j$  has  $r+1$  neighbors, but the tree only has  $r+1$  vertices total. At step  $i$ , only  $i-1 \leq r$  vertices are occupied. Thus there is always at least one free neighbor for  $u_j$ .  $\square$

## 5 Jan 23

### 5.1 Turan Number of Cycles and Girth Problem

**Definition 5.1.** If  $\mathcal{F}$  is a family of graphs,  $ex(n, \mathcal{F}) =$  maximum number of edges in an  $n$  vertex graph  $G$  that contains no  $F \in \mathcal{F}$ .

**Lemma 5.2.**  $ex(n, \mathcal{F}) = \left(1 - \frac{1}{r} + o(1)\right) \binom{n}{2}$  if  $r = \min_{F \in \mathcal{F}} \mathcal{X}(F)$

*Proof.*

$$ex(n, \mathcal{F}) \leq \min_{F \in \mathcal{F}} ex(n, F).$$

By Erdos Stone, we're done. □

**Theorem 5.3.** Let  $\mathcal{F} = \{C_3, \dots, C_{2k}\}$  then  $ex(n, \mathcal{F}) \leq n^{1+\frac{1}{k}}$

**Lemma 5.4.**  $G$  is  $\mathcal{F}$  free iff  $G$  has girth  $\geq 2k+1$

**Lemma 5.5.** Every graph  $G$  contains a subgraph  $G'$  that is bipartite with  $e(G') \geq \frac{1}{2}e(G)$ .

*Proof.* Color vertices red or blue uniformly and independent. Then delete the edges that are monochromatic. By independence, for any  $e \in E(G)$

$$\Pr[e \in E(G')] = \frac{1}{2} \implies E[e(G')] = \frac{1}{2}e(G).$$

□

We can also prove this with the greedy algorithm and will achieve a slightly better bound.

*Proof.* To prove theorem 5.3:

Suppose  $G$  has at least  $n^{1+\frac{1}{k}}$ , by lemma 5.2 we may replace  $G$  by a sub-graph of min deg  $d \geq n^{1+\frac{1}{k}}$ .

We can do a depth first search by picking a root vertex  $v$ . Let  $V_i$  be the set of vertices at distance  $i$  from  $v$ . So  $V_0 = \{v\}$ . Observe that  $G[V_0, V_1, \dots, V_k]$  forms a tree with root  $v$  and  $V_i$  being the descendent at level  $i$  else we get a cycle of length at most  $2k$ . So we have

$$|V_0| + |V_1| + \dots + |V_k| \geq 1 + d + (d-1)d + d(d-1)^2 + \dots + d(d-1)^{k-1} = 1 + d^k > n.$$

So, we have a contradiction. □

**Theorem 5.6.**  $ex(n, C_{2k}) \leq c_k n^{1+\frac{1}{k}}$

**Theorem 5.7.**  $ex(n, \{C_4, C_6, \dots, C_{2k}\}) \leq 2n^{1+\frac{1}{k}}$

**Remark 5.8.** The exponent  $1 + \frac{1}{k}$  is sharp if  $k = 2, 3, 5$

### 5.2 Turan Number for Paths

**Theorem 5.9.** Let  $P_k$  be a path with  $k$  edges then

$$ex(n, P_k) \leq \frac{k}{2}n.$$

*Proof.* By induction on  $n$  it is enough to work with connected graph as we can just split into connected components. Delete vertices of degree  $\leq \frac{k}{2}$ . If  $e(G) > \frac{k}{2}n$  then this also holds after deletions. So WLOG  $G$  has minimum degree at least  $\frac{k}{2}$ .

Let  $v_0, v_1, \dots, v_r$  be the longest path in  $G$ . The neighbors of  $v_0, v_r$  are in the path otherwise you can extend the path. If  $v_0 \sim v_r$  then  $v_0 v_1 \dots v_r$  is a cycle and so  $v_i v_{i+1} \dots v_r v_0 \dots v_{i-1}$  is also a path of same length. Thus, the neighbors of  $v_i$  are inside the path. So  $\{v_0, v_1, \dots, v_r\}$  is a connected component of  $G$  and thus  $r \geq k+1$  else  $e(G) \leq \frac{(r-1)n}{2}$ . Otherwise, if  $v_0 \not\sim v_r$  then there exist  $i$  such that  $v_0 \sim v_{i+1}$  and  $v_r \sim v_i$  then we can produce a cycle  $v_0 v_{i+1} v_{i+2} \dots v_r v_i v_{i-1} \dots v_0$  and this is of length  $r$ , so it becomes the previous case. By pigeonhole, such a  $i$  exist as  $\{i : v_0 v_{i+1} \in E(G)\}$  and  $\{i : v_r v_i \in E(G)\}$  must intersect as  $r \leq k$  by  $v_0 v_r \notin E(G)$  □

## 6 Jan 26: Projective Geometry

**Definition 6.1. Projective Plane over a Field.** Let  $\mathbb{F}$  be a field.

1. *Points:* The set of lines through 0 in  $\mathbb{F}^3$ ; equivalently, the 1-dimensional subspaces.
2. *Lines:* The set of planes through 0 in  $\mathbb{F}^3$ ; equivalently, the 2-dimensional subspaces.
3. *Incidence:* A point  $P$  lies on a line  $\ell$  if  $P \subset \ell$ .

Observe:

1. Any two points determine a unique line.
2. Any two lines intersect in a unique point.
3. There exist 4 points no three of which are collinear.

Inside a projective plane over  $\mathbb{F}$ , there is a copy of the affine plane  $\mathbb{F}^2$  embedded as  $\{z = 1\}$ .

We denote this by  $\mathbb{P}^2(\mathbb{F})$ .

We say  $p \sim \ell$  if  $p \in \ell$  for  $p$  a point and  $\ell$  a line of  $\mathbb{P}^2(\mathbb{F}_q)$ .

**Remark 6.2.** There are  $q^3 - 1$  nonzero vectors in  $\mathbb{F}_q^3 \setminus \{0\}$ . We identify points of  $\mathbb{P}^2(\mathbb{F}_q)$  with equivalence classes of nonzero vectors under

$$(x, y, z) \sim (\lambda x, \lambda y, \lambda z) \quad \text{for } \lambda \in \mathbb{F}_q^\times.$$

There are  $q - 1$  nonzero choices for  $\lambda$ , so the number of points in  $\mathbb{P}^2(\mathbb{F}_q)$  is

$$\frac{q^3 - 1}{q - 1} = q^2 + q + 1.$$

By a similar argument, the number of lines is also  $q^2 + q + 1$ .

Each line contains exactly  $q + 1$  points and dually, each point lies on exactly  $q + 1$  lines. There is a natural bijection (called a polarity) between points and lines, preserving incidence: given a point  $p$ , its corresponding line is the orthogonal complement  $p^\perp$ .

**Theorem 6.3. Construction of a  $C_4$ -free graph:**

We construct an improved  $C_4$ -free graph as follows:

- The vertices are the points of  $\mathbb{P}^2(\mathbb{F}_q)$ .
- Points  $p_1$  and  $p_2$  are adjacent if and only if  $p_1 \in p_2^\perp$  (i.e.,  $p_1$  lies on the line associated to  $p_2$ 's orthogonal complement).

For any pair  $p_1, p_2$ , their common neighborhood is

$$p_1^\perp \cap p_2^\perp,$$

which is a single point. Thus, the graph is  $C_4$ -free (contains no 4-cycles). This gives  $q^2 + q + 1$  vertices with average degree  $\approx q$ .

Note: There exist  $p$  such that  $p \subseteq p^\perp$ .

**Definition 6.4. General Projective Planes.** If  $P$  (points) and  $L$  (lines) are sets with an incidence relation such that:

1. For any two points, there is a unique line through them.
2. Any two lines meet at a unique point.
3. There exist four points, no three collinear.



*Example: The Fano plane (order 2) with 7 points and 7 lines. Each line contains 3 points, and each point lies on 3 lines.*

**Definition 6.5. Order of a Projective Plane.** *The order of a projective plane is the number of points per line minus one (so  $q$  if each line has  $q + 1$  points).*

**Remark 6.6.** *Are there planes of order which are not prime powers?*

**Theorem 6.7. Construction of  $C_{2k}$ -free graphs for  $k = 2, 3, 5$ :**

*Let  $q$  be a prime power and let  $\mathbb{F}_q$  be the finite field with  $q$  elements. Define a graph  $G$  on the vertex set  $\mathbb{F}_q^k$  (the set of  $k$ -tuples over  $\mathbb{F}_q$ ) as follows:*

*For  $\mathbf{a} = (a_0, a_1, \dots, a_{k-1})$  and  $\mathbf{b} = (b_0, b_1, \dots, b_{k-1})$  in  $\mathbb{F}_q^k$ , declare  $\mathbf{a} \sim \mathbf{b}$  if and only if for all  $j = 0, 1, \dots, k-2$ ,*

$$b_j = a_j + a_{j+1} \cdot b_{k-1}$$

*where indices are taken modulo  $k$  (so  $a_k = a_0$ ).*

*Then  $G$  is  $C_{2k}$ -free, i.e., it contains no cycle of length  $2k$ .*

*Proof.* Suppose for contradiction that  $G$  contains a cycle of length  $2k$ :

$$\mathbf{a}^1, \mathbf{b}^1, \mathbf{a}^2, \mathbf{b}^2, \dots, \mathbf{a}^k, \mathbf{b}^k,$$

with  $\mathbf{a}^i \sim \mathbf{b}^i \sim \mathbf{a}^{i+1}$  for all  $i$ , indices modulo  $k$ .

**Step 1: System of equations.** For each  $i = 1, \dots, k$  and  $j = 0, \dots, k-2$ , adjacency gives:

$$b_j^i = a_j^i + a_{j+1}^i \cdot b_{k-1}^i \quad (\text{from } \mathbf{a}^i \sim \mathbf{b}^i)$$

$$b_j^i = a_j^{i+1} + a_{j+1}^{i+1} \cdot b_{k-1}^i \quad (\text{from } \mathbf{b}^i \sim \mathbf{a}^{i+1})$$

Subtract:

$$(a_j^i - a_j^{i+1}) + (a_{j+1}^i - a_{j+1}^{i+1}) \cdot b_{k-1}^i = 0$$

or

$$(a_j^{i+1} - a_j^i) + (a_{j+1}^{i+1} - a_{j+1}^i) \cdot b_{k-1}^i = 0$$

Summing over  $i$ :

$$\sum_{i=1}^k (a_j^{i+1} - a_j^i) + \sum_{i=1}^k (a_{j+1}^{i+1} - a_{j+1}^i) b_{k-1}^i = 0$$

The first sum telescopes to zero (since indices are cyclic), so for each  $j = 0, \dots, k-2$ ,

$$\sum_{i=1}^k (a_{j+1}^{i+1} - a_{j+1}^i) b_{k-1}^i = 0 \quad (3)$$

**Step 2: Specialize  $j = k - 2$ .** Let  $\Delta_i = a_{k-1}^{i+1} - a_{k-1}^i$  and  $x_i = b_{k-1}^i$ . Then:

$$\sum_{i=1}^k \Delta_i x_i = 0$$

and

$$\sum_{i=1}^k \Delta_i = 0$$

by telescoping.

**Step 3: Linear algebra.** The system (3) for  $j = 0, \dots, k - 2$  can be written as a  $(k - 1) \times k$  Vandermonde-type matrix  $V$  applied to the vector  $(\Delta_1, \dots, \Delta_k)$ :

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_k \\ x_1^2 & x_2^2 & \cdots & x_k^2 \\ \vdots & \vdots & & \vdots \\ x_1^{k-2} & x_2^{k-2} & \cdots & x_k^{k-2} \end{pmatrix}$$

The equations become  $V \cdot (\Delta_1, \dots, \Delta_k)^\top = 0$ . Since  $\sum_{i=1}^k \Delta_i = 0$ , the vector is in the kernel of the row  $(1, 1, \dots, 1)$ .

**Step 4: Key property of the  $x_i$ .**

**Claim 6.8.** For each  $i$ ,  $x_i \neq x_{i+1}$  (indices modulo  $k$ ).

*Proof.* If  $x_i = x_{i+1}$ , then  $b_{k-1}^i = b_{k-1}^{i+1}$ . By the adjacency relations, this forces  $\mathbf{b}^i = \mathbf{b}^{i+1}$ , so  $\mathbf{a}^{i+1} = \mathbf{a}^{i+2}$ , contradicting that the cycle is proper.  $\square$

**Step 5: Contradiction for  $k = 2, 3, 5$ .** Since not all  $\Delta_i$  are zero (otherwise the cycle is trivial), and  $\sum_{i=1}^k \Delta_i = 0$ , equation  $\sum_{i=1}^k \Delta_i x_i = 0$  forces a linear dependence among the  $x_i$  (not all distinct). Thus, for these small  $k$  we can classify possibilities:

**Claim 6.9.** For  $k = 2, 3, 5$ , if  $x_i = x_j$  for some  $i \neq j$ , then  $x_i = x_{i+1}$  for some  $i$ , contradicting the above claim.

- **Case  $k = 2$ :** Only possible if  $x_1 = x_2$ , violating  $x_1 \neq x_2$ .
- **Case  $k = 3$ :** If  $x_1 = x_2$ ,  $x_2 = x_3$ , or  $x_3 = x_1$ , we violate the claim.
- **Case  $k = 5$ :** If  $x_i = x_j$  for  $i \neq j$ , by relabeling, could assume  $x_1 = x_j$  for  $j = 2, 3, 4, 5$ :
  - $j = 2$  or  $5$  gives  $x_1 = x_2$  or  $x_5 = x_1$ , directly a violation.
  - $j = 3$ :  $x_1 = x_3$ . The system  $\sum_{i=1}^5 \Delta_i = 0$ ,  $\sum_{i=1}^5 \Delta_i x_i = 0$  with  $x_1 = x_3$  and  $x_i \neq x_{i+1}$  for any  $i$  only allows further coincidences that eventually force  $x_i = x_{i+1}$ .
  - $j = 4$ : similarly,  $x_1 = x_4$  eventually forces a pair  $x_i = x_{i+1}$ .
- **Case  $k = 4$ :** The pattern  $x_1 = x_3 \neq x_2 = x_4$  is possible with no consecutive equal  $x_i$ , which is why this construction fails for  $k = 4$ .

Therefore, for  $k = 2, 3, 5$ , no proper  $2k$ -cycle exists in  $G$ .  $\square$

## 7 Jan 28

Recall:  $ex(n, K_{s,t}) \leq c_{s,t} n^{2-\frac{1}{s}} + o(n)$ . We will see that  $ex(n, K_{s,t}) = \theta \left( n^{2-\frac{1}{s}} \right)$  for  $t \geq t_0(s)$

**Theorem 7.1.** *Construction of the extremal graph for  $ex(n, K_{s,t})$ .*

*Proof.* We will first consider a random graph by picking edges with probability  $p = n^{-\frac{1}{s}}$ . Let  $x^{(1)}, \dots, x^{(s)}$  be  $s$  vertices chosen uniformly at random without replacement on the left. Let  $N(x^{(1)}, \dots, x^{(s)})$  that are the common neighbors with  $d(x^{(1)}, \dots, x^{(s)}) = |N(\dots)|$ . Let our random variable be  $d(x^{(1)}, \dots, x^{(s)})$  and is distributed according to

$$\text{Binomial}(n, p^s) = \text{Binomial}\left(n, \frac{1}{n}\right) \sim \text{Poisson}(1).$$

So  $\mathbb{E}[d(x^{(1)}, \dots, x^{(s)})] = 1$  and  $\mathbb{E}[d(\dots)^k] = o_k(1)$ . With probability  $\frac{1}{n^\epsilon}$  there is at least  $\epsilon \log n$  neighbors. Yet there are  $\binom{n}{s}$  such  $s$  tuples  $(x^{(1)}, \dots, x^{(s)})$ . and then  $\frac{n}{s}$  disjoint. This does not work.

Consider another construction of  $K_{s,t}$ -free graph with  $\Omega(n^{2-\frac{1}{s}})$  edges.

For a bipartite graph with  $\mathbb{F}_p^s$  where  $n = 2p^s$  on each side we say  $x \sim y$  if  $f(x, y) = 0$ . Pick equation  $f$  uniformly at random among all polynomials in  $s, t$  of degree at most  $d$  in each of  $x$  and  $y$ . It would be

$$f = \sum_{|\alpha|, |\beta| \leq d} c_{\alpha\beta} x_1^{\alpha_1} x_2^{\alpha_2} \dots x_s^{\alpha_s} y_1^{\beta_1} y_2^{\beta_2} \dots y_s^{\beta_s}. \quad (c_{\alpha\beta} \in \mathbb{F}_p, |\alpha| = \alpha_1 + \dots + \alpha_d)$$

Part I: This random graph is locally (neighborhood look "similar") similar to uniform random graph.

Part II: Globally very different/more discrete.  $\square$

**Lemma 7.2.** *For any fixed  $x, y \in \mathbb{F}_p^s$*

$$\Pr[x \sim y] = \frac{1}{p}.$$

*Proof.* We have

$$f = c_{0,0} + \sum_{\alpha, \beta \neq 0} c_{\alpha\beta} x^\alpha y^\beta.$$

For any choice of  $c_{\alpha, \beta}$  fir  $(\alpha, \beta) \neq (0, 0)$  there is a unique choice of  $c_{0,0}$  such that  $f(x, y) = 0$ .

$$\mathbb{E}[e(G)] = \frac{1}{p} n^2 = c_s n^{2-\frac{1}{s}}.$$

$\square$

Let  $x^{(1)}, \dots, x^{(s)}$  be  $s$  vertices on say left,  $N(x^{(1)}, \dots, x^{(s)})$  be the common neighbors with of  $x^{(1)}, \dots, x^{(s)}$  and this is same as  $\{y \in \mathbb{F}_p^s : f(x^{(1)}, y) = 0, \dots, f(x^{(s)}, y) = 0\}$ .

**Lemma 7.3.** *If  $g$  is a random polynomial of degree of  $D$  in  $m$  variables and  $z^{(1)}, \dots, z^{(t)} \in \mathbb{F}_p^m$  then the random variables  $g(z^{(1)}), \dots, g(z^{(t)})$  are independent if  $D > t$  and  $p > t^2$ .*

*Warm up:* When  $m = 1$ , write  $g(x) = \underbrace{a_0 + a_1 x + \dots + a_{t-1} x^{t-1}}_{g_{\text{small}}(x)} + \underbrace{\dots + a_D x^D}_{g_{\text{large}}(x)}$ .

For any choice of  $g_{\text{large}}(x)$  and any values  $b_1, \dots, b_t$  there is a unique polynomial  $g_{\text{small}}$  such that  $g(z^{(i)}) = g_{\text{small}}(z^{(i)}) + g_{\text{large}}(z^{(i)}) = b_i$  for  $i = 1, \dots, t$  if and only if  $g_{\text{small}}(z^{(i)}) = b_i - g_{\text{large}}(z^{(i)})$  by lagrange interpolation.

For the general case, if  $z_1^{(1)}, \dots, z_1^{(t)}, \dots, z_m^{(1)}, \dots, z_m^{(t)}$  are distinct, then  $g(x_1, \dots, x_m) = \bar{g}(x_1) + \text{other terms}$  and use  $m = 1 =$  case. Pick a invertible linear transform  $f : \mathbb{F}_p^s \rightarrow \mathbb{F}_p^s$  such that  $Tz^{(1)}, \dots, Tz^{(t)}$  are distinct. To pick such a  $T$ , let  $T(x_1, \dots, x_s) = (x_1 + S(x_2, \dots, x_s), x_2, \dots, x_s)$ . For  $z^{(i)}, z^{(j)}$  distinct, the space of  $S$  for which  $T(z^{(i)}) = T(z^{(j)})$  is of co-dimension 1, so a random  $S$  has probability  $\frac{1}{p}$  of satisfying this. So  $\Pr[\exists i, j \text{ such that } T(z^{(i)}) = T(z^{(j)})] \leq \frac{1}{p} \binom{t}{2} < 1$ .



We can now conclude from

$$\Pr[f(x^{(1)}, y) = 0, \dots, f(x^{(s)}, y) = 0] = \frac{1}{p^s}.$$

If  $d > s$  and  $p > s^2$  then

$$\mathbb{E}[N(x^{(1)}, \dots, x^{(s)})] = 1.$$

Furthermore, if  $\alpha$  is a positive integer and  $p \cdot s < d$  then

$$\mathbb{E} \left[ \left| N(x^{(1)}, \dots, x^{(s)}) \right|^p \right] = \mathbb{E} \left[ \left( \sum_{y \in \mathbb{F}_p^s} R(y) \right)^\alpha \right] = \mathbb{E} \left[ \sum_{y^{(1)}, \dots, y^{(\alpha)}} R(y^{(1)}) \dots R(y^{(\alpha)}) \right].$$

Where  $R(y) = \begin{cases} 1 & \text{if } y \in N(x^{(1)}, \dots, x^{(s)}) \\ 0 & \text{otherwise} \end{cases}$

$$= \sum_{y^{(1)}, \dots, y^{(\alpha)}} p^{-s \cdot \# \text{distinct } y\text{'s}} \leq \sum_k \sum_{y^{(1)}, \dots, y^{(k)} \in \mathbb{F}_p^s} p^{-sk} = \sum_k p^{sk} p^{-sk} \cdot \# \text{ways to partition a objection into } k \text{ bins.}$$

Last equality from having to choose  $y^{(i)}$  such that  $k$  are distinct. This is in  $O(1)$ . We want to assign values to  $y^{(i)}$  such that exactly  $k$  are distinct. Equivalently, choose  $k$  distinct values in  $\mathbb{F}_p^s$ . Then choose which of the  $y^{(i)}$  go to which bin.

$$\Pr[|N(x^{(1)}, \dots, x^{(s)})| \geq T] \leq \frac{\mathbb{E} \left[ \left| N(x^{(1)}, \dots, x^{(s)}) \right|^\alpha \right]}{T^\alpha} = \frac{C_{\alpha, s}}{T^\alpha}.$$

The set  $N(x^{(1)}, \dots, x^{(s)})$  is contained in  $\{y \in \mathbb{F}_p^s : f(x^{(1)}, y) = 0, \dots, f(x^{(s)}, y) = 0\}$ . A vector subspace  $V$  of  $\mathbb{F}_p^s$  of dimension  $D$  has  $p^D$  elements.

**Theorem 7.4.** *Lang-Weil Theorem: If  $g_1, \dots, g_r$  are polynomials of degree  $\leq D$  in  $s$  variables over  $\mathbb{F}_p$ , then the set  $\{y \in \mathbb{F}_p^s : g_1(y) = \dots = g_r(y) = 0\}$  has either  $O_{D, r}(1)$  points, or more than  $\frac{p}{2}$  points.*

So  $\Pr[\text{some } s \text{ vertices have neighborhood of size } \geq O_{s, \alpha}(1)] \leq C_{s, \alpha} \left(\frac{p}{2}\right)^{-\alpha} \left(\frac{p^s}{s}\right)$ . Choose  $\alpha = s^2 + 1$ . To ensure this probability goes to 0, we need to choose  $d = s^2 + 1$ .

## 8 Regularity Lemmas: Feb 4

**Definition 8.1.** Binary word is sequence  $w_1, w_2, \dots, w_n \in \{0, 1\}$  and write  $w = w_1 w_2 \dots, w_n$ .  
The density of  $w$  is

$$d(w) = \frac{\#1's \text{ in } w}{n}.$$

**Definition 8.2.** Word  $w$  is  $\epsilon$ -regular if for any subword  $w'$  of  $w$  of length  $\geq \epsilon \cdot \text{len}(w)$  satisfies

$$|d(w) - d(w')| \leq \epsilon.$$

**Lemma 8.3.** By Chernoff bound,

$$\Pr[|d(w') - d(w)| \geq \epsilon] \leq e^{-c_\epsilon n}.$$

## 9 Feb 6

**Definition 9.1.** Partition  $\mathcal{P}'$  is a refinement of partition  $\mathcal{P}$  if each  $w' \in \mathcal{P}'$  is a sub-word of some word in  $\mathcal{P}$

**Theorem 9.2.** (Regularity Lemma for Binary Words)

For all  $\epsilon > 0$ , every word  $w \in \{0, 1\}^n$  can be partitioned into  $\leq M(\epsilon)$  many sub-words  $w^{(1)}, \dots, w^{(m)}$  (i.e.  $w = w^{(1)} \dots w^{(m)}$ ). So that the total length of  $w^{(i)}$ 's that are  $\epsilon$ -irregular is at most  $\epsilon n$ .

*Proof.* For a partition  $\mathcal{P}$  of  $w$  into sub-words  $\mathcal{P} = \{w^{(1)}, \dots, w^{(m)}\}$  we define  $f(\mathcal{P}) = \sum_{w' \in \mathcal{P}} d(w)^2 |w'|/|w|$ . Alternatively, we pick a symbol of  $w$  at random, say  $i \in [n]$  uniformly, then pick another symbol at random from the same part as  $i$ 's, say  $j$  is picked.

$$\Pr[w_i = w_j] = \sum_{w' \in \mathcal{P}} \frac{|w'|}{|w|} (d(w')^2 + (1 - d(w'))^2) = 2f(\mathcal{P}) + \sum_{w'} \frac{|w'|}{|w|} - 2 \sum_{w'} d(w') \frac{|w'|}{|w|} = 2f(\mathcal{P}) + 1 - 2d(w).$$

**Lemma 9.3.** If  $\mathcal{P}'$  is a refinement of  $\mathcal{P}$  then

$$f(\mathcal{P}') = f(\mathcal{P}) + \sum_{w' \in \mathcal{P}'} (d(w') - d(\text{parent}(w')))^2 \frac{|w'|}{|w|}.$$

*Proof.* Expanding the second term we have

$$\sum_{w' \in \mathcal{P}'} d(w') \frac{|w'|}{|w|} - 2d(w')d(\text{parent}(w')) + d(\text{parent}(w'))^2 \frac{|w'|}{|w|} = \sum_1 + \sum_2 + \sum_3.$$

We have  $\sum_1 = f(\mathcal{P}')$  and

$$\begin{aligned} \sum_2 &= \sum_{\tilde{w} \in \mathcal{P}} \sum_{\substack{w' \in \mathcal{P}' \\ \text{parent}(w') = \tilde{w}}} \left( -2d(w')d(\tilde{w}) \frac{|w'|}{|w|} \right) \\ &= \sum_{\tilde{w} \in \mathcal{P}} -2d(\tilde{w}) \frac{1}{|w|} \sum_{\substack{w' \in \mathcal{P}' \\ \text{parent}(w') = \tilde{w}}} d(w')|w'| \\ &= \sum_{\tilde{w} \in \mathcal{P}} -2d(\tilde{w}) \frac{1}{|w|} \cdot d(\tilde{w})|\tilde{w}| \\ &= \sum_{\tilde{w} \in \mathcal{P}} -2d(\tilde{w})^2 \frac{|\tilde{w}|}{|w|} \\ &= -2f(\mathcal{P}) \end{aligned}$$

Then for the third sum

$$\begin{aligned} \sum_3 &= \sum_{\tilde{w} \in \mathcal{P}} \sum_{\substack{w' \in \mathcal{P}' \\ \text{parent}(w') = \tilde{w}}} d(\tilde{w})^2 \frac{|w'|}{|w|} \\ &= \sum_{\tilde{w} \in \mathcal{P}} d(\tilde{w})^2 \frac{|\tilde{w}|}{|w|} = f(\mathcal{P}) \end{aligned}$$

So we have  $\sum_1 + \sum_2 + \sum_3 = f(\mathcal{P}') - f(\mathcal{P})$  □

Start with the trivial partition (i.e.  $\mathcal{P} = \{w\}$ ) and repeat while  $\sum_{w' \in \mathcal{P}} |w'|$  is  $\epsilon$ -irregular  $|w'| \leq \epsilon |w|$ .

For  $w' \in \mathcal{P}$  that is  $\epsilon$ -irregular, find a way to write as  $w' = w^{(1)}w^{(2)}w^{(3)}$  where  $|d(w^{(2)}) - d(w^{(1)})| \geq \epsilon$  and  $|w^{(2)}| \geq \epsilon |w'|$ . Replace  $w'$  with the three words  $w^{(1)}, w^{(2)}, w^{(3)}$  to obtain a new partition of  $w$ , call it  $\mathcal{P}'$ . Repeat with  $\mathcal{P}$  replaced by  $\mathcal{P}'$ .

To analyze this algorithm, if we drop some words we have

$$f(\mathcal{P}') \geq f(\mathcal{P}) + \sum_{\substack{w' \in \mathcal{P}' \\ w' \text{ is } \epsilon\text{-irregular}}} \sum_{w^{(2)}} (d(w^{(2)}) - d(w')) \frac{|w^{(2)}|}{|w|} \geq f(\mathcal{P}) + \sum_{\substack{w' \in \mathcal{P}' \\ w' \text{ is } \epsilon\text{-irregular}}} \epsilon^2 \frac{|w'|}{|w|} = f(\mathcal{P}) + \epsilon^4.$$

Note:  $f(\mathcal{P}) \leq 1$  for any partition  $\mathcal{P}$ .

So the number of steps in the algorithm is  $\leq \frac{1}{\epsilon^4}$ . The theorem holds with  $M(\epsilon) = 3^{1/\epsilon^4}$   $\square$

**Example 9.4.** Define a twin in a word  $w \in \{0,1\}^n$  is a pair of subsequences  $x, y \in w$  that are disjoint (no symbol in both  $x$  and  $y$ ) and are equal as word.

We define

$$t(w) = \text{the length of longest pair twins in } w \text{ and } t(n) = \min_{w \in \{0,1\}^n} t(w).$$

We have the trivial bounds  $\frac{n}{4} \leq t(n) \leq \frac{n}{2}$ .

**Theorem 9.5.**

$$t(n) \geq \left( \frac{1}{2} - o(1) \right) n.$$

*Proof.* Pick small  $\epsilon > 0$  and cut  $w$  into  $\epsilon$ -regular words  $t \leq \epsilon n$  junk symbols. Enough  $t(w) \geq \left( \frac{1}{2} - c\epsilon \right) |w|$  for  $\epsilon$ -regular  $w$ .

Cut  $w$  into  $\frac{1}{\epsilon}$  equally long sub-words  $w^{(1)}, \dots, w^{(m)}$  where  $m = \frac{1}{\epsilon}$ .

The first twin is 0's from  $w^{(1)}$  + 1's from  $w^{(2)}$  + 0's from  $w^{(3)}$ ...

The second twin is 1's from  $w^{(1)}$  + 0's from  $w^{(2)}$  + 1's from  $w^{(3)}$ ...

Choose twins then

$$|\# \text{ 0's in } w^{(i)}| - |\# \text{ 1's in } w^{(2)}| \leq \epsilon |w^{(1)}|.$$

So, # symbols not in either of the twins is less than or equal to

$$\epsilon |w| (1's \text{ in } w^{(1)}) + \epsilon |w| (1's \text{ in } w^{(2)}) + \sum_i 4\epsilon |w^{(i)}|.$$

$\square$

## 10 Feb 13

**Definition 10.1.** For disjoint sets  $U, V \subset V(G)$ , the density between  $U$  and  $V$  is

$$d(U, V) = \frac{e(U, V)}{|U||V|}.$$

**Definition 10.2.** A pair  $(U, V)$  of disjoint subsets of  $V(G)$  is  $\epsilon$ -regular if  $\forall U' \subseteq U, V' \subseteq V$  such that  $|U'| \geq \epsilon|U|$  and  $|V'| \geq \epsilon|V|$  we have

$$|d(U, V) - d(U', V')| \leq \epsilon.$$

**Definition 10.3.** We say a partition  $V(G) = V_1 \cup \dots \cup V_k \cup J$  is  $\epsilon$ -regular partition when

1.  $|J| \leq \epsilon|V(G)|$
2.  $|V_1| = |V_2| = \dots = |V_k|$
3. All except  $\leq \epsilon k^2$  pairs  $(V_i, V_j)$  are  $\epsilon$ -regular

**Theorem 10.4.** (Szemerédi's Regularity Lemma)  $\forall \epsilon > 0, m$  there exist a constant  $M = M(\epsilon, m)$  such that every graph admits an  $\epsilon$ -regular partition  $V(G) = V_1 \cup \dots \cup V_k \cup J$  where the number of parts  $m \leq k \leq M$ .

**Remark 10.5.** # edges inside parts + # edges adjacent to  $J$  + # edges in  $\epsilon$ -irregular pairs is at most

$$k \binom{n/k}{2} + (\epsilon n) \cdot n + \epsilon k^2 (n/k)^2 \leq \frac{1}{k} n^2 + \epsilon n^2 + \epsilon n^2.$$

*Proof.* For a partition  $\mathcal{P}$  of  $V := V(G)$ , we define

$$f(\mathcal{P}) = \sum_{U, W \in \mathcal{P}, U \neq W} \underbrace{d(U, W)^2 \frac{|U||W|}{|V|^2}}_{f(U, W)}.$$

**Lemma 10.6.** Suppose  $A, B \subset V$  are disjoint and we have partitions  $A = A_1 \cup \dots \cup A_k$  and  $B = B_1 \cup \dots \cup B_l$  then

$$\sum_{i, j} f(A_i, B_j) = f(A, B) + \sum_{i, j} (d(A, B) - d(A_i, B_j))^2 \frac{|A_i||B_j|}{|V|^2}.$$

*Proof is exactly the same as the proof of the regularity lemma for binary words.*

**Algorithm:** To find an  $\epsilon$ -regular partition

1. Start with any equipartition of  $V$  into  $m$  parts.
2. We will start with partitions of the form  $V = V_1 \cup V_2 \cup \dots \cup V_k \cup J_1 \cup \dots \cup J_\ell$
3. For each  $\epsilon$ -irregular pair  $(V_i, V_j)$  there exist partitions  $V_i = V_{i,1}^{(i,j)} \cup V_{i,2}^{(i,j)}$  and  $V_j = V_{j,1}^{(i,j)} \cup V_{j,2}^{(i,j)}$  such that  $|d(V_i, V_j) - d(V_{i,1}^{(i,j)}, V_{j,1}^{(i,j)})| \geq \epsilon$  and  $|V_{i,1}^{(i,j)}| \geq \epsilon|V_i|$  and  $|V_{j,1}^{(i,j)}| \geq \epsilon|V_j|$   
Note: We write is as such since there may be many  $\epsilon$ -irregular pairs  $(V_i, V_j)$  and we want to keep track of them.
4. For each  $i$  consider all partitions of  $V_i$  as  $V_i = V_{i,1} \cup V_{i,2}$  and take the common refinement. Two  $V, V'$  are in the same part of the refinement iff  $V, V'$  are in the same part in each of the partitions.  
Observe: Each  $V_i$  is cut into  $\leq 2^k$  parts.

Let  $\mathcal{P}$  be the old partition and  $\mathcal{P}_{\text{new}}$  be the new partition.

$$\begin{aligned}
f(\mathcal{P}_{\text{new}}) &= \sum_{\substack{A, B \in \mathcal{P}_{\text{new}} \\ A \neq B}} f(A, B) \\
&\geq \sum_i f(J_i, ?) + \sum_{i, j} \sum_{\substack{A, B \\ \text{parent}(A)=V_i \\ \text{parent}(B)=V_j}} f(A, B) \\
&\geq f(\text{Junk}) + \sum_{i \neq j} \left( f(V_i, V_j) + \epsilon^2 \frac{|V_i||V_j|}{|V|^2} \right) \\
&= f(\mathcal{P}) + \epsilon^4 \left( \frac{|V| - |\text{Junk}|}{|V|} \right)^2 \\
&\geq f(\mathcal{P}) + \frac{\epsilon^4}{4}
\end{aligned}$$

5. Cut each of the parts in  $\mathcal{P}_{\text{new}}$  into equally big parts so total # parts is  $\leq k \cdot 4^k$  and some left over where # of leftovers is  $\leq k2^k$  and the size of them  $\leq k2^k \cdot \frac{|V|}{k4^k} \leq \frac{|V|}{2^k}$

So the total junk size is  $\leq \sum_{k \geq m} \frac{|V|}{2^k}$

The total number of iterations is at most  $\frac{4}{\epsilon^4}$ , and let  $g(k) = k \cdot 4^k$ . The total number of parts at the end is at most

$$M \leq g^{(\lceil 1/\epsilon^5 \rceil)}(m)$$

where  $g^{(t)}$  denotes  $g$  composed with itself  $t$  times, starting from  $m$ ; that is, we iterate  $g$  at most  $\frac{1}{\epsilon^5}$  times as  $\epsilon k^2 \cdot \frac{\epsilon^4}{k^2} = \epsilon^5$ .

So using  $g(k) \leq 4^{4^k}$  we have

$$M \leq 4^{4^{4^{\dots 4^m}}} \left\} \frac{2}{\epsilon^5} \text{ times.} \right.$$

□

## 11 Feb 16

**Lemma 11.1 (Triangle Counting Lemma).** *Let  $X, Y, Z \subset V(G)$  be disjoint with  $d(X, Y), d(X, Z), d(Y, Z) \geq 2\epsilon$ , and suppose  $(X, Y), (X, Z), (Y, Z)$  are  $\epsilon$ -regular. Then the number of  $XYZ$ -triangles is at least*

$$(1 - 2\epsilon)(d(X, Y) - \epsilon)(d(X, Z) - \epsilon)(d(Y, Z) - \epsilon)$$

*Proof.* Observe the following:

1.  $Y_{\text{bad}} = \{x \in X : \deg_Y(x) < (d(X, Y) - \epsilon)|Y|\}$
2.  $Z_{\text{bad}} = \{x \in Z : \deg_Z(x) < (d(X, Z) - \epsilon)|Z|\}$

Then  $|Y_{\text{bad}}|, |Z_{\text{bad}}| \leq \epsilon|X|$ .

Consider  $Y_{\text{bad}}$ ,

$$e(Y_{\text{bad}}, Y) = \sum_{x \in Y_{\text{bad}}} \deg_Y(x) < |Y_{\text{bad}}|(d(X, Y) - \epsilon)|Y|.$$

By  $\epsilon$ -regularity we have  $d(X, Y) - \epsilon$  is within  $\epsilon$  of  $d(X, Y)$  so a contradiction if  $|Y_{\text{bad}}| > \epsilon|X|$ .

To prove the Triangle Counting Lemma, let  $X' := X \setminus (Y_{\text{bad}} \cup Z_{\text{bad}})$  then  $|X'| \geq (1 - 2\epsilon)|X|$  by observation.

Let  $x \in X'$  then  $N_Y(x) := N(x) \cap Y$  and  $N_Z(x) := N(x) \cap Z$ .

The number of triangles containing  $x$  is

$$e(N_Y(x), N_Z(x)) \geq |N_Y(x)||N_Z(x)|(d(Y, Z) - \epsilon). \quad (*)$$

By  $\epsilon$ -regularity of  $(Y, Z)$  and definition of bad. Additionally

$$(*) \geq (d(X, Y) - \epsilon)(d(X, Z) - \epsilon)(d(Y, Z) - \epsilon).$$

So

$$\#\Delta = \sum_{x \in X'} (*) \geq (*) (1 - 2\epsilon)|X|.$$

□