Math 554

Hemanshu Kaul

kaul@iit.edu

Entropy and Counting

Lemma (Jensen's Inequality)

Let f be a continuous concave function. Then for any reals $x_1, \ldots, x_n$,

$$\frac{\sum_{i=1}^{n} f(x_i)}{n} \leq f\left(\frac{\sum x_i}{n}\right),$$

or more generally,

$$\frac{\sum a_i f(x_i)}{\sum a_i} \leq f\left(\frac{\sum a_i x_i}{\sum a_i}\right) \quad \text{for } a_i > 0 \ \forall i.$$

$$\boxed{\text{Or, } \int_{\Omega} f \circ g \, d\mu \leq f\left(\int_{\Omega} g \, d\mu\right) \qquad \mathbb{E}[f(X)] \leq f(\mathbb{E}[X])}$$

Property (Uniform Bound)

$H(X) \leq \log |\mathrm{Range}(X)|$, where $\mathrm{Range}(X)$ is the set of values X takes with positive probability.

$H(X) = \log |\mathrm{Range}(X)|$ iff X is uniform on its range.

Suppose $\mathcal{C}$ is a set whose size we want to estimate. If X selects each element of $\mathcal{C}$ uniformly at random then $H(X) = \log |\mathcal{C}|$, i.e. $|\mathcal{C}| = 2^{H(X)}$. So estimate $H(X)$ to find $|\mathcal{C}|$.

We will work with vectors of random variables,

$X = (X_1, \dots, X_n)$, X is the joint r.v. of $X_1, \dots, X_n$.

taking values in $\text{Range}(X_1) \times \dots\dots \times \text{Range}(X_n)$

with joint probability distribution of $X_1, \dots, X_n$.

Property (Subadditivity) $H(X_1, \dots, X_n) \leq \sum_{i=1}^{n} H(X_i)$

We will work with vectors of random variables,

$X = (X_1, \dots, X_n)$, X is the joint r.v. of $X_1, \dots, X_n$.

taking values in $\mathrm{Range}(X_1) \times \dots \times \mathrm{Range}(X_n)$

with joint probability distribution of $X_1, \dots, X_n$.

Property (Subadditivity) $H(X_1, \dots, X_n) \leq \sum_{i=1}^{n} H(X_i)$

Proof $n=2$ $H(X,Y) \leq H(X) + H(Y)$.

$H(X) + H(Y) - H(X,Y)$

$$= \sum_x \mathbb{P}[X=x] \log \frac{1}{\mathbb{P}[X=x]} + \sum_y \mathbb{P}[Y=y] \log \frac{1}{\mathbb{P}[Y=y]} - \sum_x \sum_y \mathbb{P}[X=x, Y=y] \log \frac{1}{\mathbb{P}[X=x, Y=y]}$$

$$= \sum_x \sum_y \mathbb{P}[X=x, Y=y] \log \frac{1}{\mathbb{P}[X=x]} + \sum_y \sum_x \mathbb{P}[Y=y, X=x] \log \frac{1}{\mathbb{P}[Y=y]} - \sum_x \sum_y \mathbb{P}[X=x, Y=y] \log \frac{1}{\mathbb{P}[X=x, Y=y]}$$

$$= \sum_x \sum_y p(x,y) \left[\log \frac{1}{p(x)} + \log \frac{1}{p(y)} - \log \frac{1}{p(x,y)} \right]$$

$$= \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = \sum_x \sum_y p(x)p(y) \left[\frac{p(x,y)}{p(x)p(y)} \log \frac{p(x,y)}{p(x)p(y)} \right]$$

$$\geq f\left(\sum_x \sum_y p(x)p(y) \frac{p(x,y)}{p(x)p(y)} \right) = f(1) = 0,$$ by Jensen to $f(z) = z \log z$ convex function.

Observations

- $H(X) \leq H(X, Y)$ (by monotonicity of $\log z$)
- $H(X, Y) = H(X) + H(Y)$ if X & Y are independent

What if X & Y are not independent?

Observations

- $H(X) \leq H(X,Y)$ (by monotonicity of $\log z$)
- $H(X,Y) = H(X) + H(Y)$ if X & Y are independent

What if X & Y are not independent?

Conditional Entropy

If A is an event then define $H(X \mid A) = \sum_x \mathbb{P}[X = x \mid A] \log \frac{1}{\mathbb{P}[X = x \mid A]}$

$$= \sum_x p(x \mid A) \log \frac{1}{p(x \mid A)}$$

This naturally extends to

$$H(X \mid Y) = \mathbb{E}_y(H(X \mid Y = y)) = \sum_y H(X \mid Y = y)\, \mathbb{P}[Y = y]$$

Observations

- $H(X) \leq H(X,Y)$ (by monotonicity of $\log z$)
- $H(X,Y) = H(X) + H(Y)$ if X & Y are independent

What if X & Y are not independent?

Conditional Entropy

If A is an event then define $H(X|A) = \sum_{x} \mathbb{P}[X=x|A] \log \frac{1}{\mathbb{P}[X=x|A]}$

$$= \sum_{x} p(x|A) \log \frac{1}{p(x|A)}$$

This naturally extends to

$$H(X|Y) = \mathbb{E}_y(H(X|Y=y)) = \sum_{y} H(X|Y=y)\,\mathbb{P}[Y=y]$$

Observations

- $H[X|X] = 0$
- $H[X|Y] = 0$ if X is determined by Y $(X = g(Y))$
- $H[X|Y] = H[X]$ if X & Y are independent.

Property (Chain Rule)

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i \mid X_1, \ldots, X_{i-1})$$

Proof Try it for $n=2$ $\quad H(X,Y) = H(X) + H(Y|X)$

same algebra as proof for subadditivity but Jensen is not needed

Entropy of a random vector can be found by revealing its components one by one (recall, conditional probability & Bayes rule)

Property (Chain Rule)

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i \mid X_1, \ldots, X_{i-1})$$

Proof Try it for $n=2$ $\quad H(X,Y) = H(X) + H(Y|X)$

same algebra as proof for subadditivity but Jensen is not needed.

Entropy of a random vector can be found by revealing its components one by one (recall, conditional probability & Bayes rule)

Property (Dropping Conditioning) $\quad H(X|Y) \leq H(X)$ ← interpretation?

$H(X|Y,Z) \leq H(X|Z)$

Proof By chain rule & subadditivity, $H(X|Y) = H(X,Y) - H(Y) \leq H(X)$

which also gives $H(X|Y, Z=z) \leq H(X|Z=z), \ldots$ ∎

The non-neg quantity $\underline{I(X;Y) = H(X) + H(Y) - H(X,Y)}$ is called <u>mutual information</u>, amount of common information between X & Y.

The conditional versions of all basic properties are also true (can be proved in the same manner)

- $H(X|A) \le \log |\mathrm{Range}(X|A)|$ for r.v. X & event A where $\mathrm{Range}(X|A)$ is the set of values X takes on with positive probability, given A has occured.

- $H(X_1, \ldots, X_n | Z) = \sum_{i=1}^{n} H(X_i | X_1, \ldots, X_{i-1}, Z)$

- $H(X_1, \ldots, X_n | Z) \le \sum_{i=1}^{n} H(X_i | Z)$

An Application

Consider n distinct points in $\mathbb{R}^3$ that have

n_1 distinct projections on xy-plane

n_2 distinct projections on xz-plane

n_3 distinct projections on zy-plane

Show that $n \leq (n_1 n_2 n_3)^{1/2}$

An Application

Consider n distinct points in $\mathbb{R}^3$ that have

n_1 distinct projections on xy-plane

n_2 distinct projections on xz-plane

n_3 distinct projections on zy-plane

Show that $n \leq (n_1 n_2 n_3)^{1/2}$

Let (X, Y, Z) be picked uniformly at random from these n points

so, $\log n = H(X,Y,Z) = H(X) + H(Y|X) + H(Z|X,Y)$

we know, $\log n_1 \geq H(X,Y) = H(X) + H(Y|X)$ ← why?

$\log n_2 \geq H(X,Z) = H(X) + H(Z|X) = H(Z) + H(X|Z)$

$\log n_3 \geq H(Y,Z) = H(Y) + H(Z|Y)$

An Application

Consider n distinct points in $\mathbb{R}^3$ that have

n_1 distinct projections on xy-plane

n_2 distinct projections on xz-plane

n_3 distinct projections on zy-plane

Show that $n \leq (n_1 n_2 n_3)^{1/2}$

Let (X, Y, Z) be picked uniformly at random from these n points

so, $\log n = H(X,Y,Z) = H(X) + H(Y|X) + H(Z|X,Y)$

we know, $\log n_1 \geq H(X,Y) = H(X) + H(Y|X)$

$\log n_2 \geq H(X,Z) = H(X) + H(Z|X) = H(Z) + H(X|Z)$

$\log n_3 \geq H(Y,Z) = H(Y) + H(Z|Y)$

$n \leq \sqrt{n_1 n_2 n_3} \iff 2\log n \leq \log n_1 + \log n_2 + \log n_3 \iff 2\log n - (\log n_1 + \log n_2 + \log n_3) \leq 0$

Plug-in & do the algebra

An Application

Consider n distinct points in $\mathbb{R}^3$ that have

n_1 distinct projections on xy-plane $(z=0)$
n_2 distinct projections on xz-plane $(y=0)$
n_3 distinct projections on zy-plane $(x=0)$

Then $n \leq (n_1 n_2 n_3)^{1/2}$

This can be generalized to d-dimensions (HW!) and then using scaling & limiting arguments gives

Loomis-Whitney Isoperimetric Ineq. (1949)

Let B be a measurable body in $\mathbb{R}^d$. Then $\mathrm{vol}(B) \leq \prod_{i=1}^{d} \mathrm{vol}(B_i)^{1/d-1}$ where B_i is the projection onto the hyperplane $x_i = 0$.

This bound is tight, e.g. for the cube.

Binary Entropy and Binomial coefficients

are closely related.

Stirling's approximation to $n!$ can be used to show

Fix $\alpha \in (0,1)$, then $\dfrac{2^{H(\alpha)n}}{n+1} \leq \binom{n}{\alpha n} \leq 2^{H(\alpha)n}$ where $H(\alpha)$ is the binary Entropy function

which gives $\binom{n}{\alpha n} \sim \dfrac{2^{H(\alpha)n}}{\sqrt{2\pi n \alpha(1-\alpha)}}$

Binary Entropy and Binomial coefficients

are closely related.

<u>Theorem</u> Fix $\alpha \in (0, \frac{1}{2}]$. Then $\sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i} \leq 2^{H(\alpha) n}$

<u>Comment</u> The above inequality can also be interpreted probabilistically $\longrightarrow$ $X \sim \text{Binomial}(n, \frac{1}{2}) \Rightarrow \mathbb{P}[X \leq \lfloor \alpha n \rfloor] \leq 2^{H(\alpha) n}$

which can be proved by MGF technique for proving Chernoff-type bounds to get

$$\sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i} \leq \frac{(1+x)^n}{x^{\lfloor \alpha n \rfloor}} \quad \forall x \in [0,1]$$

and taking infimum of $\frac{(1+x)^n}{x^{\lfloor \alpha n \rfloor}}$ over $x \in [0,1]$

gives $2^{H(\alpha) n}$.

To generalize this to Binomial(n, p) we need the notion of "relative entropy" or Kullback–Leibler divergence.

Binary Entropy and Binomial coefficients

are closely related.

Theorem Fix $\alpha \in (0, \frac{1}{2}]$. Then $\sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i} \leq 2^{H(\alpha) n}$

Proof Let $\mathcal{C}$ be the set of all subsets of $[n]$ of size at most αn, so $|\mathcal{C}| = \sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i}$ — ① ← Combinatorial Interpretation

Let X be uniformly chosen member of $\mathcal{C}$, then

$H(X) = \log|\text{Range}(X)| = \log|\mathcal{C}|$, i.e. $|\mathcal{C}| = 2^{H(X)}$ — ②

Binary Entropy and Binomial coefficients

are closely related.

Theorem Fix $\alpha \in (0, \frac{1}{2}]$. Then $\sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i} \leq 2^{H(\alpha)n}$

Proof Let $\mathcal{C}$ be the set of all subsets of $[n]$ of size at most αn, so $|\mathcal{C}| = \sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i}$ — ①

Let X be uniformly chosen member of $\mathcal{C}$, then

$H(X) = \log |\text{Range}(X)| = \log |\mathcal{C}|$, i.e. $|\mathcal{C}| = 2^{H(X)}$. — ②

We view X as $(X_1, \ldots, X_n)$ where X_i is the indicator r.v. for picking the element i in X, $1 \leq i \leq n$.

$\therefore H(X) \leq H(X_1) + H(X_2) + \cdots + H(X_n) = n H(X_1)$ — ③

Claim $H(X_1) \leq H(\alpha)$ (①, ②, ③ with this claim will finish the proof.)

n

Binary Entropy and Binomial coefficients

are closely related.

Theorem Fix $\alpha \in (0, \frac{1}{2}]$. Then $\sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i} \leq 2^{H(\alpha) n}$

Proof Let $\mathcal{C}$ be the set of all subsets of $[n]$ of size at most αn, so $|\mathcal{C}| = \sum_{i=0}^{\lfloor \alpha n \rfloor} \binom{n}{i}$ — ①

Let X be uniformly chosen member of $\mathcal{C}$, then

$H(X) = \log |\mathrm{Range}(X)| = \log |\mathcal{C}|$, i.e. $|\mathcal{C}| = 2^{H(X)}$. — ②

We view X as $(X_1, \ldots, X_n)$ where X_i is the indicator r.v. for picking the element i in X, $1 \leq i \leq n$.

$\therefore H(X) \leq H(X_1) + H(X_2) + \cdots + H(X_n) = n H(X_1)$ — ③

Claim $H(X_1) \leq H(\alpha)$ (①, ②, ③ with this claim will finish the proof.)

Note $H(X_1) = H(p)$ where $p = \mathbb{P}[1 \in X]$ and $H(p)$ is the binary entropy fn.

Since $p \leq \alpha < \frac{1}{2}$ ($|X| = \lfloor \alpha n \rfloor \Rightarrow p = \mathbb{P}[i \in X] = \alpha$; & $|X| < \lfloor \alpha n \rfloor \Rightarrow p < \alpha$)

we have $H(p) \leq H(\alpha)$.

A generalization of Subadditivity

Shearer's Lemma (1978 / 1986)

Let $\mathcal{F} \subseteq 2^{[n]}$, family of subsets of $[n]$, such that each element of $[n]$ appears in at least k members of $\mathcal{F}$. Then, for r.v. $X = (X_1, \dots, X_n)$,

system of coordinates

$$H(X) \leq \frac{1}{k} \sum_{F \in \mathcal{F}} H(X_F)$$

, where $X_F = (X_j : j \in F)$

"X_F = projection of X onto coordinates given by F"

$\mathcal{F} = \{\{1\}, \{2\}, \dots, \{n\}\}$ gives usual subadditivity of $H(X)$

$\mathcal{F} = \{[n]\setminus\{1\}, [n]\setminus\{2\}, \dots, [n]\setminus\{n\}\}$ gives Han's inequality (1978).

A generalization of Subadditivity

Shearer's Lemma (1978 / 1986) Let $\mathcal{F} \subseteq 2^{[n]}$, family of subsets of $[n]$, such that each element of $[n]$ appears in at least k members of $\mathcal{F}$. Then, for r.v. $X = (X_1, \ldots, X_n)$,

$$H(X) \leq \frac{1}{k} \sum_{F \in \mathcal{F}} H(X_F), \quad \text{where } X_F = (X_j : j \in F)$$

Proof $H(X) = \sum_{j \in 1}^{n} H(X_j \mid (X_\ell : \ell < j))$, chain rule

For each $F \in \mathcal{F}$, $H(X_F) = \sum_{j \in F} H(X_j \mid (X_\ell : \ell \in F \,\&\, \ell < j))$, chain rule

Adding up all the equations for all $F \in \mathcal{F}$, we get since $j \in [n]$ appears in at least k F in $\mathcal{F}$, the resulting equation has k terms of the form $H(X_j \mid \text{——})$ on the RHS.

$$\therefore \sum_{F \in \mathcal{F}} H(X_F) \geq k \sum_{j=1}^{n} H(X_j \mid (X_\ell ; \ell < j)) = k\, H(X)$$

A generalization of Subadditivity

Shearer's Lemma (1978 / 1986) Let $\mathcal{F} \subseteq 2^{[n]}$, family of subsets of $[n]$, such that each element of $[n]$ appears in at least k members of $\mathcal{F}$. Then, for r.v. $X = (X_1, \ldots, X_n)$,

$$H(X) \leq \frac{1}{k} \sum_{F \in \mathcal{F}} H(X_F) \quad , \text{ where } X_F = (X_j : j \in F)$$

Shearer's Lemma (combinatorial Version) Let Ω be a finite set. Let $\mathcal{S} = \{S_1, \ldots, S_m\}$ be a family of subsets of Ω s.t. each element of Ω appears in at least k members of $\mathcal{S}$. Let $\mathcal{F} \subseteq 2^{\Omega}$, be any family of subsets of Ω. Then

$$|\mathcal{F}| \leq \left(\prod_{i=1}^{m} |\mathcal{F}_i| \right)^{1/k} \quad \text{where } \mathcal{F}_i = \{F \cap S_i \mid F \in \mathcal{F}\}$$

Projection / Trace of $\mathcal{F}$ onto S_i

Clever choice of "coordinate system" / "projections" reduces a global problem ($|\mathcal{F}|$ or $H(X)$) to easier local problems ($|\mathcal{F}_i|$ or $H(X_F)$)

A generalization of Subadditivity

Shearer's Lemma (1978 / 1986) Let $\mathcal{F} \subseteq 2^{[n]}$, family of subsets of $[n]$, such that each element of $[n]$ appears in at least k members of $\mathcal{F}$. Then, for r.v. $X = (X_1, \ldots, X_n)$,

$$H(X) \leq \frac{1}{k} \sum_{F \in \mathcal{F}} H(X_F) \quad , \text{ where } X_F = (X_j : j \in F)$$

Shearer's Lemma (combinatorial Version) Let Ω be a finite set. Let $\mathcal{S} = \{S_1, \ldots, S_m\}$ be a family of subsets of Ω s.t. each element of Ω appears in at least k members of $\mathcal{S}$. Let $\mathcal{F} \subseteq 2^{\Omega}$, be any family of subsets of Ω. Then

$$|\mathcal{F}| \leq \left(\prod^{m} |\mathcal{F}_i|\right)^{1/k} \quad \text{where } \mathcal{F}_i = \{F \cap S_i \mid F \in \mathcal{F}\}$$

Proof WLOG let $\Omega = [n]$. $X = (X_1, \ldots, X_n)$ be an element of $\mathcal{F}$ chosen uniformly at random, where X_i is indicator r.v. for "$i \in X$".
Then $\log|\mathcal{F}| = H(X) \leq \frac{1}{k}\sum_{i=1}^{m} H(X_{S_i}) \leq \frac{1}{k}\sum_{i=1}^{m} \log|\text{Range}(X_{S_i})| = \frac{1}{k}\sum_{i=1}^{m} \log|\mathcal{F}_i|$)

By Shearer

An easy Application

Proposition Let $\mathcal{F} \subseteq 2^{[n]}$ s.t. $\forall F, F' \in \mathcal{F}$, $\exists\, i \in [n]$ s.t. $\{i, i+1\} \subseteq F \cap F'$

Then $|\mathcal{F}| \leq 2^{n-2}$

Sharpness?

An easy Application

Proposition Let $\mathcal{F} \subseteq 2^{[n]}$ s.t. $\forall F, F' \in \mathcal{F}$, $\exists\, i \in [n]$ s.t. $\{i, i+1\} \subseteq F \cap F'$

Then $|\mathcal{F}| \leq 2^{n-2}$

Sharp, e.g. 

Proof Let S_0 = set of all even numbers in $[n]$
S_1 = set of all odd numbers in $[n]$ $\Big\}$ choice of projection coordinates

$\mathcal{F}_i = \{F \cap S_i \mid F \in \mathcal{F}\}$, projection of $\mathcal{F}$ onto S_i, $i = 0, 1$.

By Combinatorial Shearer, $|\mathcal{F}| \leq \left(\underbrace{|\mathcal{F}_0|}_{?}\,\underbrace{|\mathcal{F}_1|}_{?}\right)^{1/\underbrace{k}_{?}} \leq ?$

An easy Application

Proposition Let $\mathcal{F} \subseteq 2^{[n]}$ s.t. $\forall F, F' \in \mathcal{F}$, $\exists\, i \in [n]$ s.t. $\{i, i+1\} \subseteq F \cap F'$

Then $|\mathcal{F}| \leq 2^{n-2}$

Proof Let S_0 = set of all even numbers in $[n]$
S_1 = set of all odd numbers in $[n]$ } choice of projection coordinates

$\mathcal{F}_i = \{F \cap S_i \mid F \in \mathcal{F}\}$, projection of $\mathcal{F}$ onto S_i, $i = 0, 1$.

By Combinatorial Shearer, $|\mathcal{F}| \leq (\underbrace{|\mathcal{F}_0|}_{?}\, \underbrace{|\mathcal{F}_1|}_{?})^{1/\underset{?}{k}} \leq ?$

Claim: $\mathcal{F}_i$ is an intersecting family

For any $G, G' \in \mathcal{F}_i$, $G \cap G' = (F \cap S_i) \cap (F' \cap S_i) = (F \cap F') \cap S_i$

Since $F \cap F'$ contains 2 consecutive integers, $(F \cap F') \cap S_i \neq \emptyset$
(one even & one odd)

$\therefore |\mathcal{F}_i| \leq 2^{?}$

An easy Application

Proposition Let $\mathcal{F} \subseteq 2^{[n]}$ s.t. $\forall F, F' \in \mathcal{F}$, $\exists\, i \in [n]$ s.t. $\{i, i+1\} \subseteq F \cap F'$

Then $|\mathcal{F}| \leq 2^{n-2}$

Proof Let S_0 = set of all even numbers in $[n]$
S_1 = set of all odd numbers in $[n]$ } choice of projection coordinates

$\mathcal{F}_i = \{F \cap S_i \mid F \in \mathcal{F}\}$, projection of $\mathcal{F}$ onto S_i, $i = 0, 1$.

By Combinatorial Shearer, $|\mathcal{F}| \leq (\underbrace{|\mathcal{F}_0|}_{?}\,\underbrace{|\mathcal{F}_1|}_{?})^{1/k ?} \leq ?$

Claim: $\mathcal{F}_i$ is an intersecting family

For any $G, G' \in \mathcal{F}_i$, $G \cap G' = (F \cap S_i) \cap (F' \cap S_i) = (F \cap F') \cap S_i$

Since $F \cap F'$ contains 2 consecutive integers, $(F \cap F') \cap S_i \neq \emptyset$
(one even & one odd)

$\therefore |\mathcal{F}_i| \leq 2^{|S_i|-1}$ (Family of intersecting sets from S has size $\leq 2^{|S|-1}$)

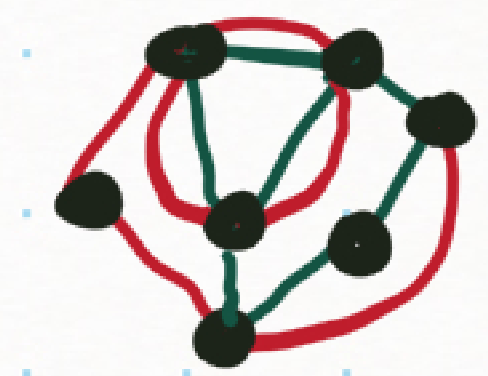
Note $k = 1$, every element of $\Omega = [n]$ appears in 1 of S_0 or S_1.

$$\therefore |\mathcal{F}| \leq (|\mathcal{F}_0|\,|\mathcal{F}_1|)^{1/k} = (2^{|S_0|-1})(2^{|S_1|-1}) = 2^{|S_0|+|S_1|-2} = 2^{n-2}.$$

Intersecting Family of Graphs

Theorem (Chung, Graham, Frankl, Shearer, 1986)

Let $\mathcal{G}$ be a family of labeled graphs on vertex set $[n]$ such that for all $G, G' \in \mathcal{G}$, $G \cap G'$ contains a K_3



Then $|\mathcal{G}| < \frac{1}{4} 2^{\binom{n}{2}}$

Conjecture (Simonovits & Sós 1976) $|\mathcal{G}| \leq \frac{1}{8} 2^{\binom{n}{2}}$

sharp?

Proved by Ellis, Filmus & Friedgut (2012) using Discrete Fourier Analysis.

Recall

Let $\mathcal{F}$ be a family of subsets of $[n]$

Without further info, best bound is $|\mathcal{F}| \leq 2^n$

If $\mathcal{F}$ is intersecting then $|\mathcal{F}| \leq \frac{1}{2} 2^n$ ($\because$ we can only pick one of A or $\bar{A}$ ....)

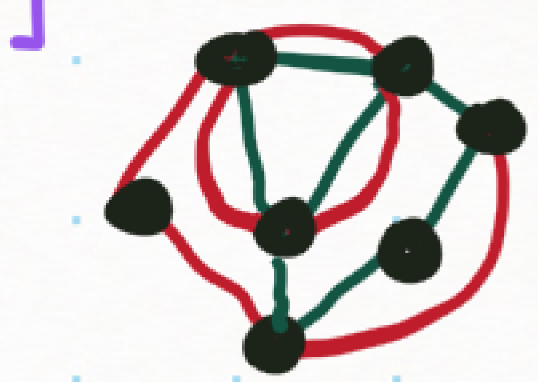
What about $\mathcal{G}$ family of graphs on vertex set $[n]$? $|\mathcal{G}| \leq 2^{\binom{n}{2}}$

If $\mathcal{G}$ is edge-intersecting then $|\mathcal{G}| \leq \frac{1}{2} 2^{\binom{n}{2}}$. (why?) What if $\mathcal{G}$ is K_3-intersecting

Intersecting Family of Graphs

Theorem (Chung, Graham, Frankl, Shearer, 1986)

Let $\mathcal{G}$ be a family of labeled graphs on vertex set $[n]$ such that for all $G, G' \in \mathcal{G}$, $G \cap G'$ contains a K_3

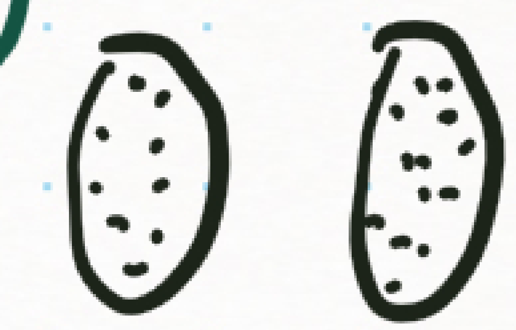


Then $|\mathcal{G}| < \frac{1}{4} 2^{\binom{n}{2}}$

Proof. Since the vertex set $[n]$ is fixed for all graphs in $\mathcal{G}$, we identify each $G \in \mathcal{G}$ with its edge set, that is $G \subseteq \Omega = [\binom{n}{2}]$

For simplicity, assume n is even (n odd case is similar with use of $\lfloor\ \rfloor$ & $\lceil\ \rceil$)

Let $S = \{S_i : 1 \leq i \leq m = ?\ \}$ where each $S_i =$ all possible edges in U and all possible edges in $\overline{U}$
$= E(K_U) \cup E(K_{\overline{U}})$
where $U \subseteq [n]$ and $|U| = \frac{n}{2}$

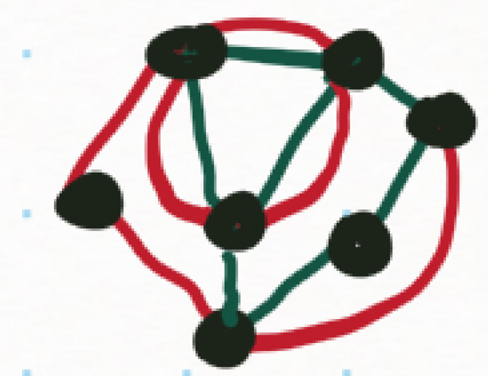


U $\overline{U}$

Intersecting Family of Graphs

Theorem (Chung, Graham, Frankl, Shearer, 1986)

Let $\mathcal{G}$ be a family of labeled graphs on vertex set $[n]$ such that for all $G, G' \in \mathcal{G}$, $G \cap G'$ contains a K_3



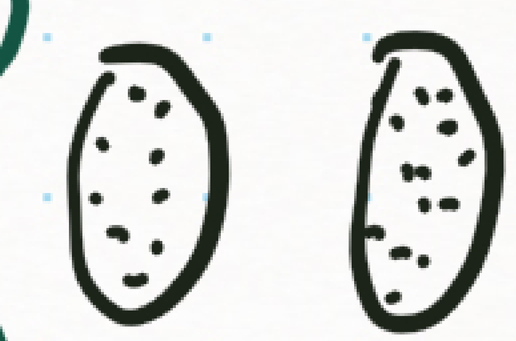
Then $|\mathcal{G}| < \frac{1}{4} 2^{\binom{n}{2}}$

Proof. Since the vertex set $[n]$ is fixed for all graphs in $\mathcal{G}$, we identify each $G \in \mathcal{G}$ with its edgeset, that is $G \subseteq \Omega = [\binom{n}{2}]$

For simplicity, assume n is even (n odd case is similar with use of $\lfloor\,\rfloor$ & $\lceil\,\rceil$)

Let $S = \{S_i : 1 \leq i \leq m = \frac{1}{2}\binom{n}{n/2}\}$ where each S_i = all possible edges in U and all possible edges in $\overline{U}$ $= E(K_U) \cup E(K_{\overline{U}})$

for $U \subseteq [n]$ and $|U| = \frac{n}{2}$



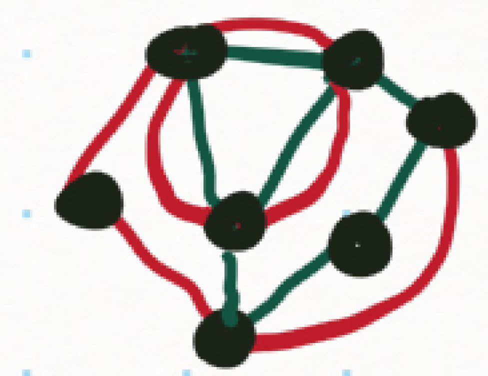
U $\quad$ $\overline{U}$

Each edge in $\Omega = E(K_n)$ is contained in $k = \binom{n-2}{n/2}$ of S_i's.

Intersecting Family of Graphs

Theorem (Chung, Graham, Frankl, Shearer, 1986)

Let $\mathcal{G}$ be a family of labeled graphs on vertex set $[n]$ such that for all $G, G' \in \mathcal{G}$, $G \cap G'$ contains a K_3



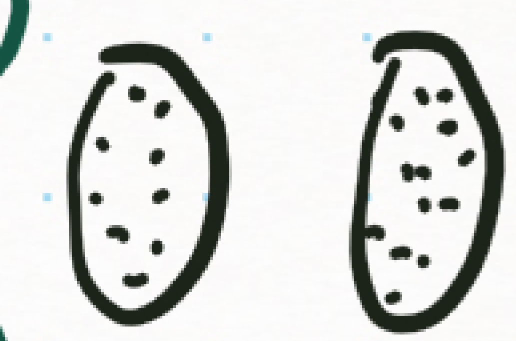
Then $|\mathcal{G}| < \frac{1}{4} 2^{\binom{n}{2}}$

Proof. Since the vertex set $[n]$ is fixed for all graphs in $\mathcal{G}$, we identify each $G \in \mathcal{G}$ with its edge set, that is $G \subseteq \Omega = [\binom{n}{2}]$

For simplicity, assume n is even (n odd case is similar with use of $\lfloor \rfloor$ & $\lceil \rceil$)

Let $S = \{S_i : 1 \leq i \leq m = \frac{1}{2}\binom{n}{n/2}\}$ where each S_i = all possible edges in U and all possible edges in $\overline{U}$ $= E(K_U) \cup E(K_{\overline{U}})$

for $U \subseteq [n]$ and $|U| = \frac{n}{2}$



U $\overline{U}$

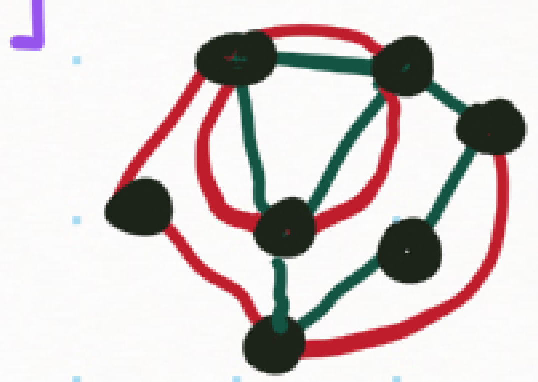
Each edge in $\Omega = E(K_n)$ is contained in $k = \binom{n-2}{n/2}$ of S_i's.

By Combinatorial Shearer, $|\mathcal{G}| \leq \left(\prod_{i=1}^{m} |\mathcal{G}_i|\right)^{1/k}$ where $\mathcal{G}_i = \{G \cap S_i : G \in \mathcal{G}\}$

Intersecting Family of Graphs

Theorem (Chung, Graham, Frankl, Shearer, 1986)

Let $\mathcal{G}$ be a family of labeled graphs on vertex set $[n]$ such that for all $G, G' \in \mathcal{G}$, $G \cap G'$ contains a K_3



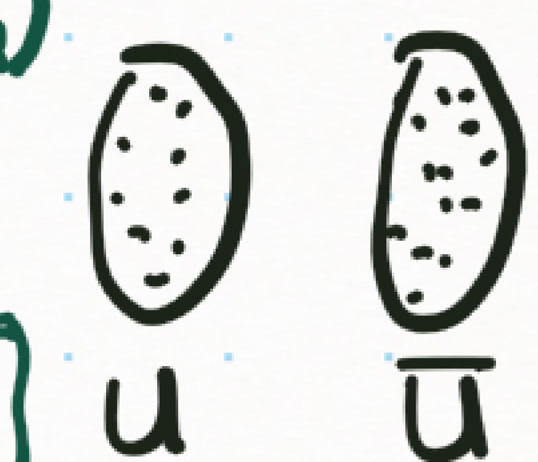
Then $|\mathcal{G}| < \frac{1}{4} 2^{\binom{n}{2}}$

Proof Since the vertex set $[n]$ is fixed for all graphs in $\mathcal{G}$, we identify each $G \in \mathcal{G}$ with its edgeset, that is $G \subseteq \Omega = [\binom{n}{2}]$

For simplicity, assume n is even (n odd case is similar with use of $\lfloor\, \rfloor$ & $\lceil\, \rceil$)

Let $S = \{S_i : 1 \leq i \leq m = \frac{1}{2}\binom{n}{n/2}\}$ where each S_i = all possible edges in U and all possible edges in $\bar{U}$
$= E(K_U) \cup E(K_{\bar{U}})$

for $U \subseteq [n]$ and $|U| = \frac{n}{2}$



Each edge in $\Omega = E(K_n)$ is contained in $k = \binom{n-2}{n/2}$ of S_i's.

By Combinatorial Shearer, $|\mathcal{G}| \leq \left(\prod_{i=1}^{m} |\mathcal{G}_i|\right)^{1/k}$ where $\mathcal{G}_i = \{G \cap S_i : G \in \mathcal{G}\}$

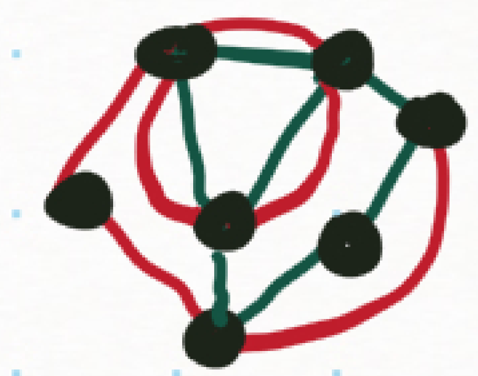
Let $G \cap S_i$ & $G' \cap S_i \in \mathcal{G}_i$ then $(G \cap S_i) \cap (G' \cap S_i) = (G \cap G') \cap S_i \neq \emptyset$

$G \cap G'$ contains K_3 whose at least one edge must belong to S_i [otherwise $K_3 \subseteq$ bipartite

Intersecting Family of Graphs

Theorem (Chung, Graham, Frankl, Shearer, 1986)

Let $\mathcal{G}$ be a family of labeled graphs on vertex set $[n]$ such that for all $G, G' \in \mathcal{G}$, $G \cap G'$ contains a K_3



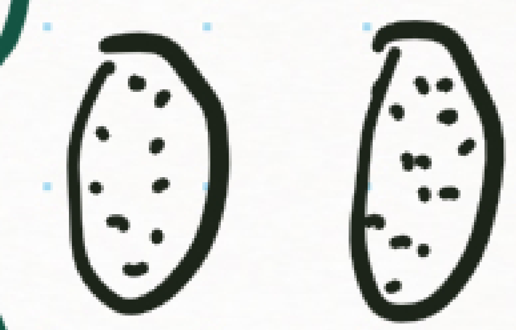
Then $|\mathcal{G}| < \frac{1}{4} 2^{\binom{n}{2}}$

Proof. Since the vertex set $[n]$ is fixed for all graphs in $\mathcal{G}$, we identify each $G \in \mathcal{G}$ with its edgeset, that is $G \subseteq \Omega = [\binom{n}{2}]$

For simplicity, assume n is even (n odd case is similar with use of $\lfloor \rfloor$ & $\lceil \rceil$)

Let $S = \{S_i : 1 \le i \le m = \frac{1}{2}\binom{n}{n/2}\}$ where each $S_i =$ all possible edges in U and all possible edges in $\overline{U}$
$= E(K_U) \cup E(K_{\overline{U}})$

for $U \subseteq [n]$ and $|U| = \frac{n}{2}$



U $\overline{U}$

Each edge in $\Omega = E(K_n)$ is contained in $k = \binom{n-2}{n/2}$ of S_i's.

By Combinatorial Shearer, $|\mathcal{G}| \le \left(\prod_{i=1}^{m} |\mathcal{G}_i|\right)^{1/k}$ where $\mathcal{G}_i = \{G \cap S_i : G \in \mathcal{G}\}$

$|\mathcal{G}_i| \le 2^{|S_i|-1} = 2^{s-1}$, since $\mathcal{G}_i$ is an intersecting family on S_i, & $|S_i| = s = 2\binom{n/2}{2}$

$\therefore |\mathcal{G}| \le 2^{(s-1)m/k}$ and $\frac{(s-1)m}{k} = \cdots \cdots = \binom{n}{2} - \frac{n(n-1)}{n(\frac{n}{2}-1)} < \binom{n}{2} - 2$ ∎

Counting the number of "copies" of hypergraph H in G

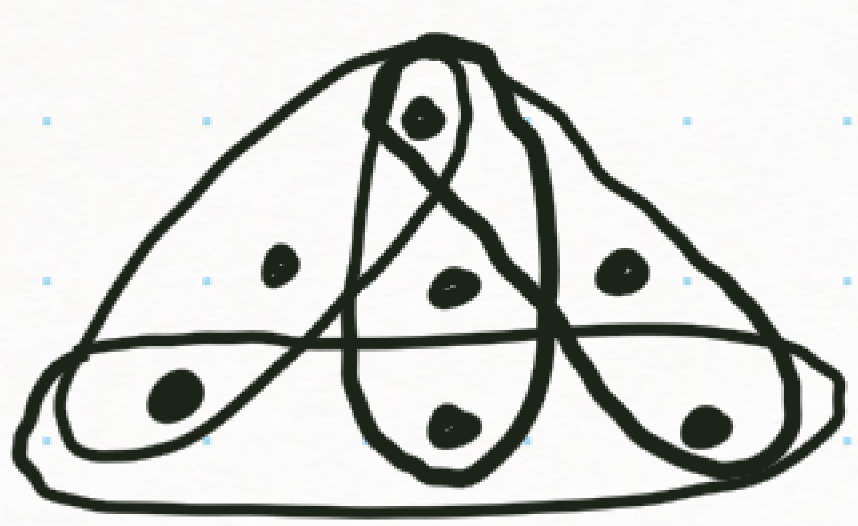
To understand this problem, we need to specify what "copy" means:

	Graphs	Hypergraphs
• Allow removal of both vertices & edges	Subgraph	Partial hypergraph
• Allow removal of only vertices	Induced subgraph	Two possibilities

What to do with edges whose at least one vertex has been removed?

"shrink" the edge but keep it.

Remove any such edge

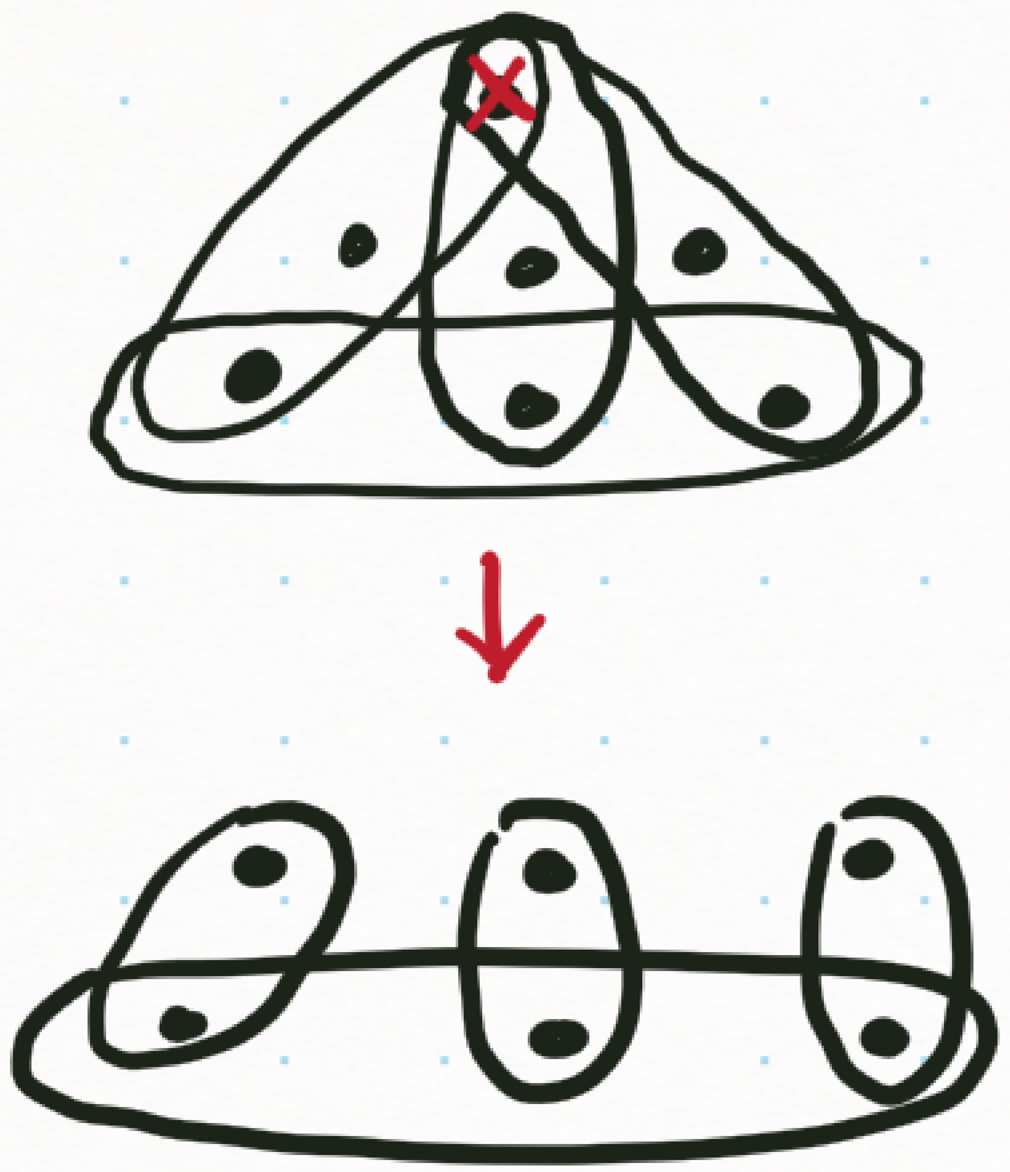


Confusingly, both versions are called subhypergraph in the literature.

Counting the number of "copies" of hypergraph H in G

To understand this problem, we need to specify what "copy" means:

	Graphs	Hypergraphs
• Allow removal of both vertices & edges	Subgraph	Partial hypergraph
• Allow removal of only vertices	Induced subgraph	Two possibilities what to do with edges whose at least one vertex has been removed?



"shrink" the edge but keep it.

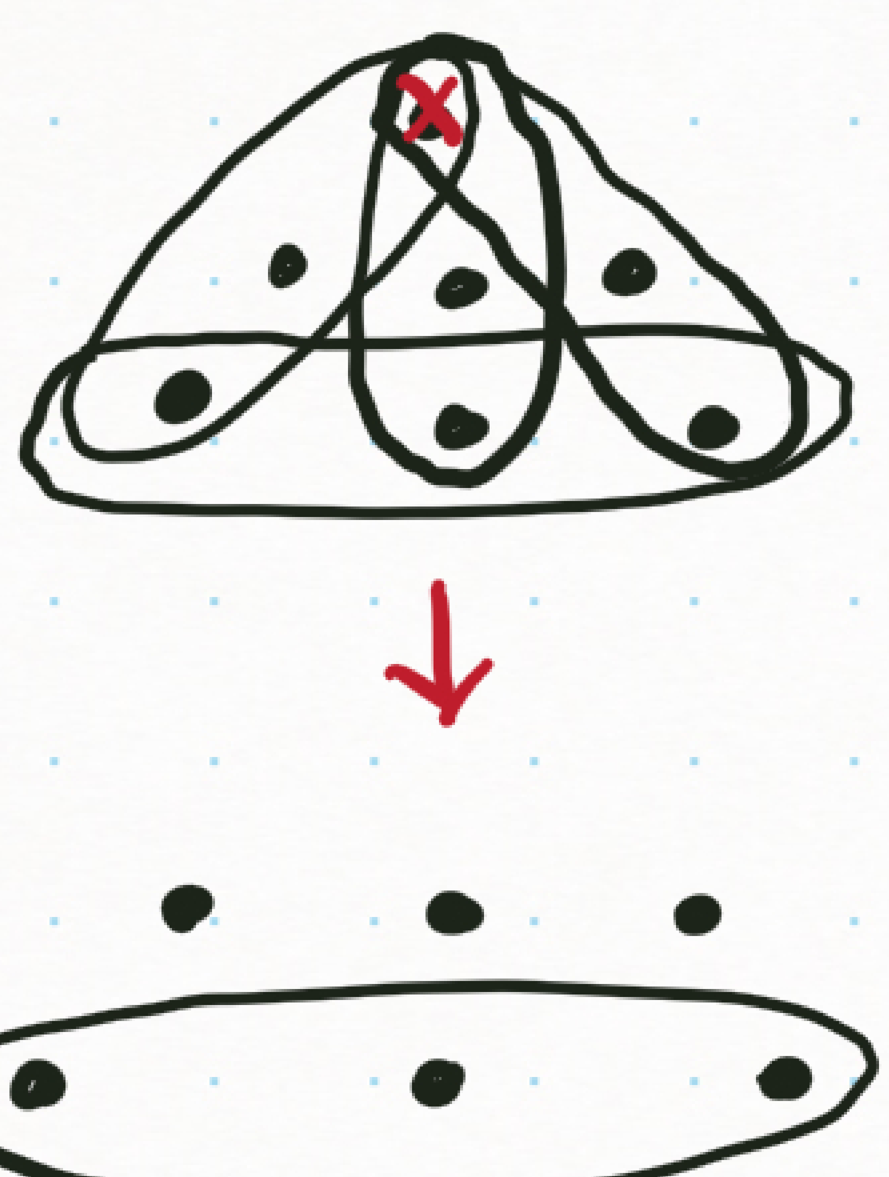
Remove any such edge

Confusingly, both versions are called subhypergraph in the literature.

Counting the number of "copies" of hypergraph H in G

To understand this problem, we need to specify what "copy" means:

	Graphs	Hypergraphs
• Allow removal of both vertices & edges	Subgraph	Partial hypergraph
• Allow removal of only vertices	Induced subgraph	Two possibilities what to do with edges whose at least one vertex has been removed?



"shrink" the edge but keep it.

Remove any such edge

Confusingly, both versions are called subhypergraph in the literature.

Counting the number of "copies" of hypergraph H in G

To understand this problem, we need to specify what "copy" means:

	Graphs	Hypergraphs
• Allow removal of both vertices & edges	Subgraph	Partial hypergraph
• Allow removal of only vertices	Induced subgraph	Two possibilities: what to do with edges whose at least one vertex has been removed?

"shrink" the edge but keep it.

★ Remove any such edge

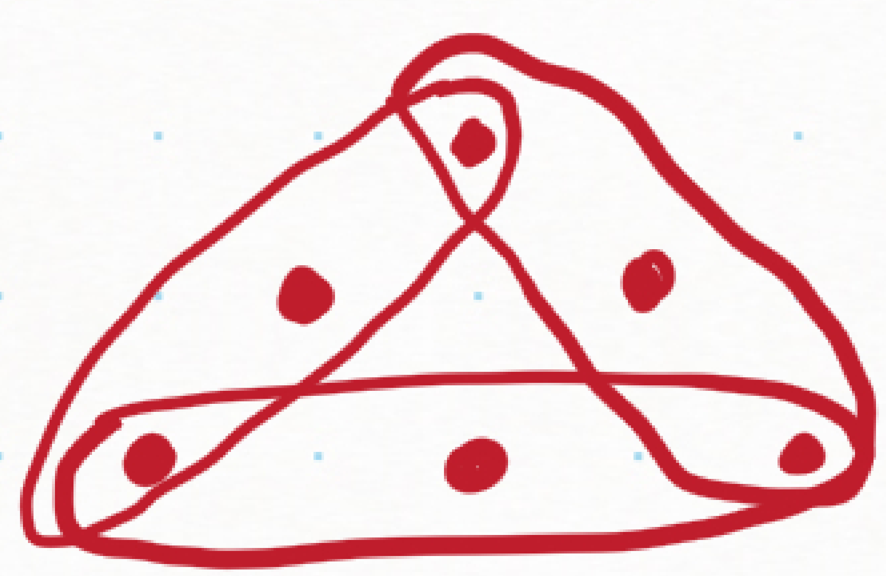
Defn $H=(V',E')$ is a subhypergraph of $G=(V,E)$ induced by $V' \subseteq V$ if $E' = \{e \in E : e \subseteq V'\}$

is the definition we will use.

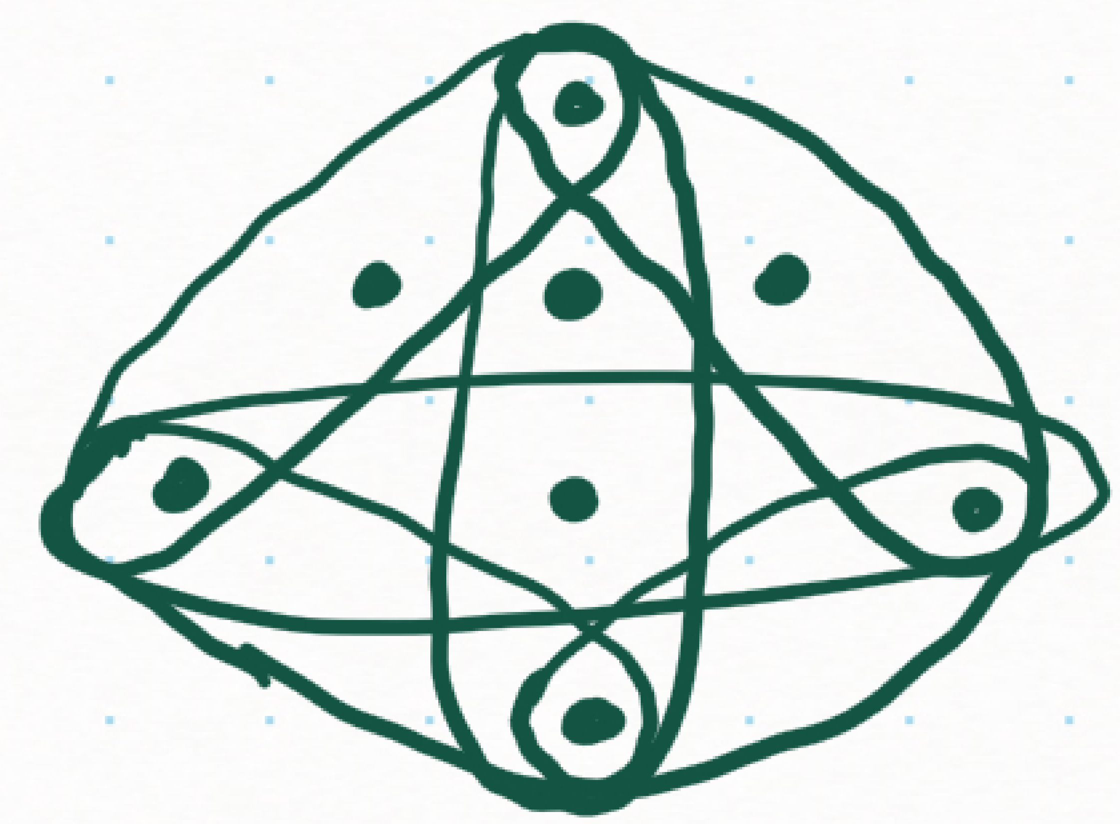
Confusingly, both versions are called subhypergraph in the literature.

Defn A hypergraph homomorphism $\varphi : V(H) \to V(G)$ is a map which preserves edges, i.e. each edge of H maps to an edge of G.

A copy of H in G corresponds to an injective homomorphism from H to G, i.e., a 1-to-1 map $\sigma : V(H) \to V(G)$ s.t. $\sigma(e) \in E(G)\ \forall e \in E(H)$.



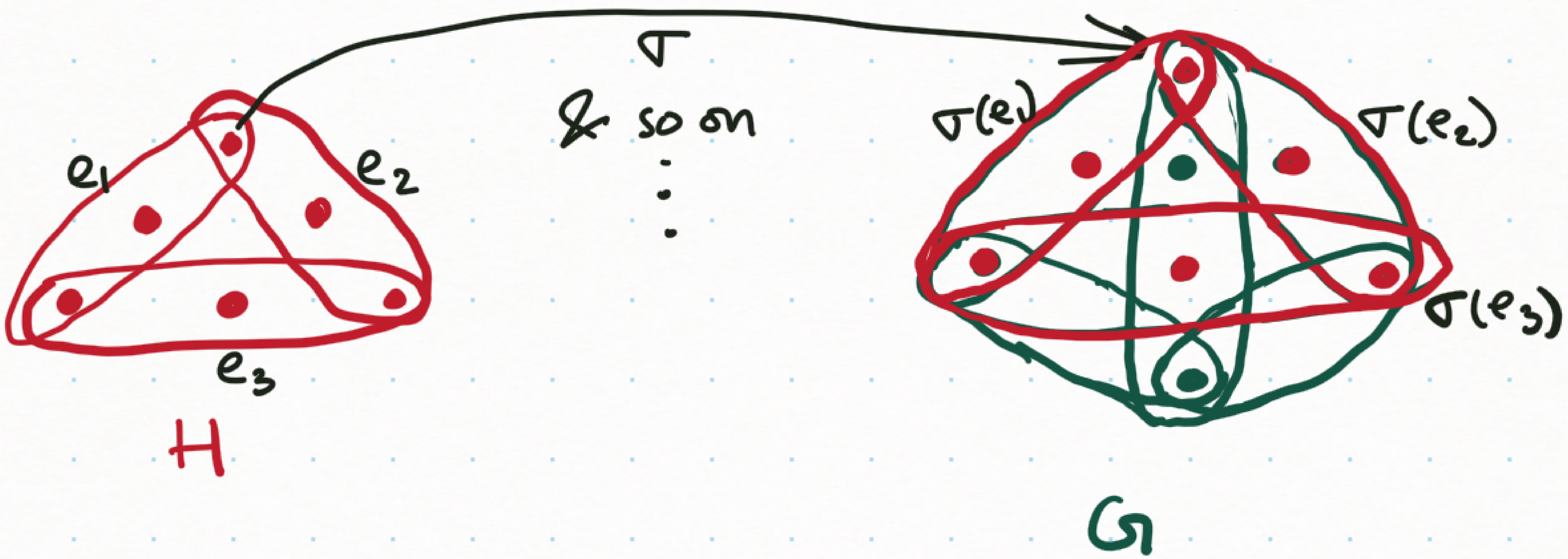
H



G

Defn A <u>hypergraph homomorphism</u> $\varphi : V(H) \to V(G)$ is a map which preserves edges, i.e. each edge of H maps to an edge of G.

A <u>copy of H in G</u> corresponds to an injective homomorphism from H to G, i.e., a 1-to-1 map $\sigma : V(H) \to V(G)$ s.t. $\sigma(e) \in E(G)$ $\forall e \in E(H)$.



There is a copy of H in G.

<u>Defn</u> A <u>hypergraph homomorphism</u> $\varphi : V(H) \to V(G)$ is a map which preserves edges, i.e. each edge of H maps to an edge of G.

A <u>copy of H in G</u> corresponds to an injective homomorphism from H to G, i.e., a 1-to-1 map $\sigma : V(H) \to V(G)$ s.t. $\sigma(e) \in E(G)$ $\forall e \in E(H)$.

H

G

Is there a copy of this H in G?

Defn A hypergraph homomorphism $\varphi : V(H) \to V(G)$ is a map which preserves edges, i.e. each edge of H maps to an edge of G.

A copy of H in G corresponds to an injective homomorphism from H to G, i.e., a 1-to-1 map $\sigma : V(H) \to V(G)$ s.t. $\sigma(e) \in E(G)$ $\forall e \in E(H)$.

Defn $N(H, \ell)$ = max # copies of a fixed hypergraph H that can appear in a hypergraph G with ℓ edges.

H fixed, small
ℓ large

Theorem [Friedgut & Kahn 1998]

For a hypergraph H with fractional cover number $\rho^*(H)$, $\exists\, c_1, c_2$ s.t. $c_1 \ell^{\rho^*(H)} \leq N(H, \ell) \leq c_2 \ell^{\rho^*(H)}$ for all ℓ.

Linear Programming Duality

$$\left.\begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \leq b \\ & x \geq 0 \end{array}\right\} ① \quad \overset{\text{Dual}}{\longleftrightarrow} \quad \left.\begin{array}{ll} \min & b^T y \\ \text{s.t.} & A^T y \geq c \\ & y \geq 0 \end{array}\right\} ②$$

where
$x \in \mathbb{R}^n$
$c \in \mathbb{R}^n$
$b \in \mathbb{R}^m$
$A_{m \times n}$

Weak Duality says $c^T x$ for any feasible solution x of ① is always at most $b^T y$ for any feasible solution y of ②.

Strong Duality of Linear Programming says
If x^* is an optimal solution of ① and y^* is an optimal soln. of ②
then $c^T x^* = b^T y^*$.

A Binary Linear Program is a linear program whose variables are restricted to be 0 or 1.

Replacing $x \in \{0, 1\}$ by $x \in [0, 1]$ gives the Linear Program relaxation.

Independence number and cover number of hypergraphs

Defn An independent set of a hypergraph H is a subset of $V(H)$ that meets each edge at most once.

$\alpha(H) =$ max size of an ind. set of H = independence number of H

$$= \max_{\varphi} \sum_{v \in V(H)} \varphi(v)$$

where the max is taken over all functions $\varphi : V(H) \to \{0,1\}$ s.t. $\sum_{v \in e} \varphi(v) \leq 1 \ \forall e \in E(H)$

Independence number and cover numbers of hypergraphs

Defn An <u>independent set</u> of a hypergraph H is a subset of $V(H)$ that meets each edge at most once.

$\alpha(H)$ = max size of an ind. set of H = independence number of H

$$= \max_{\varphi} \sum_{v \in V(H)} \varphi(v)$$

where the max is taken over all functions

$$\boxed{\varphi : V(H) \to \{0,1\} \text{ s.t. } \sum_{v \in e} \varphi(v) \leq 1 \ \forall e \in E(H)}$$

<u>Independence function of H.</u>

Defn A <u>fractional independent set</u>, or <u>fractional independence function of H</u> is $\varphi : V(H) \to [0,1]$ s.t. $\sum_{v \in e} \varphi(v) \leq 1 \ \forall e \in E(H)$

$\alpha^*(H)$ = fractional independence number of H

$$= \max_{\varphi} \sum_{v \in V(H)} \varphi(v)$$

where the max is taken over all fractional independence functions of H.

Note: $\alpha(H) \leq \alpha^*(H)$

Independence number and cover numbers of hypergraphs

Defn An (edge) cover of a hypergraph H is a set of edges whose union gives all of $V(H)$, that is each vertex is contained in at least one chosen edge.

$\rho(H)$ = min size of a cover of H = (edge) cover number of H

$= \min\limits_{\psi} \sum\limits_{e \in E(H)} \psi(e)$ where the min is taken over all functions $\boxed{\psi : E(H) \to \{0,1\} \text{ s.t. } \sum\limits_{e : v \in e} \psi(e) \geq 1 \ \forall v \in V(H)}$

cover function of H

Defn A fractional (edge) cover, or fractional cover function, of H is $\psi : E(H) \to [0,1]$ s.t. $\sum\limits_{e : v \in e} \psi(e) \geq 1 \ \forall v \in V(H)$

$\rho^*(H)$ = fractional cover number of H

$= \min\limits_{\psi} \sum\limits_{e \in E(H)} \psi(e)$ where the min is taken over all fractional cover functions of H.

Note: $\rho^*(H) \leq \rho(H)$

Recall, $\alpha(H) \leq \rho(H)$ (an edge cover has to use a different edge for each vertex of an independent set)

By Strong Duality, $\boxed{\alpha^*(H) = \rho^*(H)}$

$$\left.\begin{aligned} \alpha^*(H) = \max_{\varphi} \ & \sum_{v \in V(H)} \varphi(v) \\ \text{s.t.} \ & \sum_{v \in e} \varphi(v) \leq 1 \quad \forall e \in E(H) \\ & 0 \leq \varphi(v) \leq 1 \quad \forall v \in V(H) \end{aligned}\right] \xleftrightarrow{\text{Dual}} \left[\begin{aligned} \rho^*(H) = \min_{\psi} \ & \sum_{e \in E(H)} \psi(e) \\ \text{s.t.} \ & \sum_{e \,:\, v \in e} \psi(e) \geq 1 \quad \forall v \in V(H) \\ & 0 \leq \psi(e) \leq 1 \quad \forall e \in E(H) \end{aligned}\right.$$

Proof of $N(H,\ell) \leq C_2 \ell^{\rho^*(H)}$ Let G be any hypergraph w. ℓ edges

Assume each edge of H has at most k vertices.

Let $N = N(H,G) = \#$ copies of H in G.

Let $\Sigma = \{\sigma : V(H) \to V(G)$, 1-1 homomorphisms$\}$, so $N = |\Sigma|$.

We think of $\sigma \in \Sigma$ as $(\sigma(v) : v \in H)$, i.e. $(\sigma(v_1), \sigma(v_2), \dots, \sigma(v_n))$.

Let $\psi : E(H) \to [0,1]$ be an optimal fractional cover of H,

that is $\sum_{e \in E(H)} \psi(e) = \rho^*(H)$.

We may assume $\rho^*(H) \in \mathbb{Q}$, so $\rho^*(H) = \frac{s}{t}$ s.t. $\psi(e) = \frac{w(e)}{t}$ $\forall e$

where $w(e) \in \mathbb{Z}$ and t as common denominator of all $\psi(e)$

Proof of $N(H,\ell) \le c_2 \ell^{\rho^*(H)}$

Let G be any hypergraph w. ℓ edges

Assume each edge of H has at most k vertices.

Let $N = N(H,G) = \#$ copies of H in G.

Let $\Sigma = \{\sigma : V(H) \to V(G)$, 1-1 homomorphisms$\}$, so $N = |\Sigma|$.

We think of $\sigma \in \Sigma$ as $(\sigma(v) : v \in H)$, i.e. $(\sigma(v_1), \sigma(v_2), \dots, \sigma(v_n))$.

Let $\psi : E(H) \to [0,1]$ be an optimal fractional cover of H,

that is $\sum_{e \in E(H)} \psi(e) = \rho^*(H)$.

We may assume $\rho^*(H) \in \mathbb{Q}$, so $\rho^*(H) = \frac{s}{t}$ s.t. $\psi(e) = \frac{w(e)}{t}$ $\forall e$

where $w(e) \in \mathbb{Z}$ and t as common denominator of all $\psi(e)$

Let $\{e_1, e_2, \dots, e_m\}$ be a list (a multiset) of edges of H where each edge of H appears exactly $w(e)$ times.

Since $\sum_{e : v \in e} \psi(e) \ge 1$ $\forall v$ $\Rightarrow \sum_{e : v \in e} w(e) \ge t$ $\forall v$, each vertex appears in at least t edges in this list.

Let σ be chosen uniformly at random from Σ

so, $\log|\Sigma| = H(\sigma) = H(\sigma(v_1), \ldots, \sigma(v_n)) \leq \sum_{i=1}^{n} H(\sigma(v_i))$

$\hookrightarrow$ trivial upper bd.

<u>Apply Shearer's lemma</u>

What should we project σ onto?

$$H(\sigma) \leq \frac{1}{k} \sum_{F \in \mathcal{F}} H(\sigma_F)$$

$\mathcal{F} = ?$ & $k = ?$

for the $\mathcal{F}$.

$\mathcal{F} = \{e_1, \ldots, e_m\}$

works with $k = t$.

Think! Go back 1 page.

Let σ be chosen uniformly at random from Σ

so, $\log|\Sigma| = H(\sigma) = H(\sigma(v_1), \dots, \sigma(v_n))$

By Shearer $\leq \frac{1}{t} \sum_{i=1}^{m} H(\sigma_{e_i})$, where $\sigma_{e_i} = (\sigma(v) : v \in e_i)$

Let σ be chosen uniformly at random from Σ

so, $\log|\Sigma| = H(\sigma) = H(\sigma(v_1), \ldots, \sigma(v_n))$

By Shearer $\leq \frac{1}{t} \sum_{i=1}^{m} H(\sigma_{e_i})$, where $\sigma_{e_i} = (\sigma(v) : v \in e_i)$

$\leq \frac{1}{t} \sum_{i=1}^{m} (\log (k! * l))$ since $H(\sigma_{e_i}) = H(\sigma(v) : v \in e_i)$

$\leq \log(k! * l)$

\# choices of mapping vertices in each edge from H to G

\# choices for edges

$(-,-,-,-,-) \xrightarrow{\sigma} (-,-,-,-,-)$

$e \in H$ $\sigma(e) \in G$

Let σ be chosen uniformly at random from Σ

so, $\log|\Sigma| = H(\sigma) = H(\sigma(v_1), \dots, \sigma(v_n))$

By Shearer $\leq \frac{1}{t} \sum_{i=1}^{m} H(\sigma_{e_i})$, where $\sigma_{e_i} = (\sigma(v) : v \in e_i)$

$\leq \frac{1}{t} \sum_{i=1}^{m} (\log(k! * l))$ since $H(\sigma_{e_i}) = H(\sigma(v) : v \in e_i)$

$\leq \log(\underset{\substack{\uparrow \\ \text{\# choices of mapping vertices in each edge from H to G}}}{k!} * \underset{\substack{\uparrow \\ \text{\# choices for edges}}}{l})$

$(, , , ,) \xrightarrow{\sigma} (, , , ,)$
$e \in H \qquad \sigma(e) \in G$

$\leq \frac{1}{t} \log(k! * l) \sum_{i=1}^{m} 1$

$= \frac{1}{t} \log(k! * l) \sum_{e \in E(H)} \omega(e)$ since $\{e_1, \dots, e_m\}$ contains each edge $\omega(e)$ times

Let σ be chosen uniformly at random from Σ

so, $\log|\Sigma| = H(\sigma) = H(\sigma(v_1), \ldots, \sigma(v_n))$

By Shearer $\displaystyle \leq \frac{1}{t} \sum_{i=1}^{m} H(\sigma_{e_i})$, where $\sigma_{e_i} = (\sigma(v) : v \in e_i)$

$\displaystyle \leq \frac{1}{t} \sum_{i=1}^{m} \left(\log(k! * \ell)\right)$ since $H(\sigma_{e_i}) = H(\sigma(v) : v \in e_i)$

$\leq \log(k! * \ell)$

\# choices of mapping vertices in each edge from H to G

\# choices for edges

$(\, , \, , \, , \,) \xrightarrow{\sigma} (\, , \, , \, , \,)$

$e \in H$ $\quad$ $\sigma(e) \in G$

$\displaystyle \leq \frac{1}{t} \log(k! * \ell) \sum_{i=1}^{m} 1$

$\displaystyle = \frac{1}{t} \log(k! * \ell) \sum_{e \in E(H)} w(e)$ since $\{e_1, \ldots, e_m\}$ contains each edge $w(e)$ times

$\displaystyle = \log(k! * \ell) \sum_{e \in E(H)} \psi(e)$ since $\psi(e) = \frac{w(e)}{t}$ $\forall e$

$= \log(k! \cdot \ell)\, \rho^*(H)$

Let σ be chosen uniformly at random from Σ

so, $\log|\Sigma| = H(\sigma) = H(\sigma(v_1), \ldots, \sigma(v_n))$

By Shearer $\leq \frac{1}{t} \sum_{i=1}^{m} H(\sigma_{e_i})$, where $\sigma_{e_i} = (\sigma(v) : v \in e_i)$

$\leq \frac{1}{t} \sum_{i=1}^{m} (\log(k! * l))$ since $H(\sigma_{e_i}) = H(\sigma(v) : v \in e_i)$

$\leq \log(k! * l)$

#choices of mapping vertices in each edge from H to G

#choices for edges

$(\,,\,,\,,\,)_{e \in H} \xrightarrow{\sigma} (\,,\,,\,,\,)_{\sigma(e) \in G}$

$\leq \frac{1}{t} \log(k! * l) \sum_{i=1}^{m} 1$

$= \frac{1}{t} \log(k! * l) \sum_{e \in E(H)} w(e)$ since $\{e_1, \ldots, e_m\}$ contains each edge $w(e)$ times

$= \log(k! * l) \sum_{e \in E(H)} \psi(e)$ since $\psi(e) = \frac{w(e)}{t} \; \forall e$

$= \log(k!\, l)\, \rho^*(H)$

$\therefore |\Sigma| \leq (k!\, l)^{\rho^*(H)} = \underbrace{(k!)^{\rho^*(H)}}_{c_2}\, l^{\rho^*(H)} = c_2\, l^{\rho^*(H)}$

where c_2 depends only on H.