Maximal Inequalities on the Cube

Alexandra Kolla (UIUC)

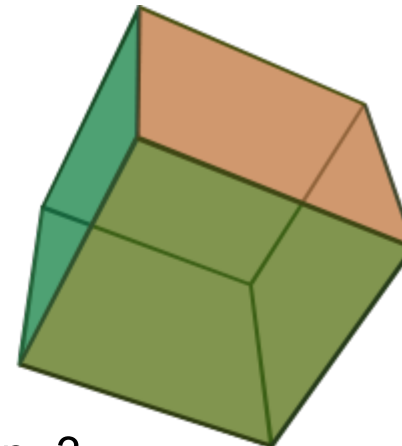
joint with

Aram W. Harrow, MIT

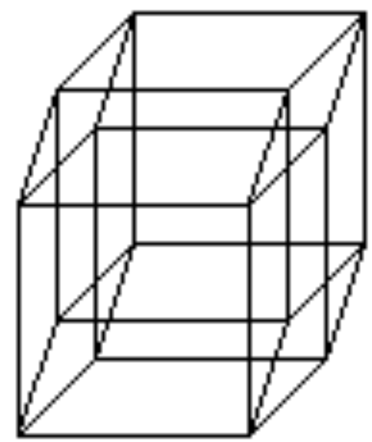
Leonard Schulman, Caltech

Talk outline

- Two motivations.
- A combinatorial problem about the geometry of the n-dimensional hypercube H^n.
- Connection to a problem in Analysis
- How to solve it (sketch).



n=3



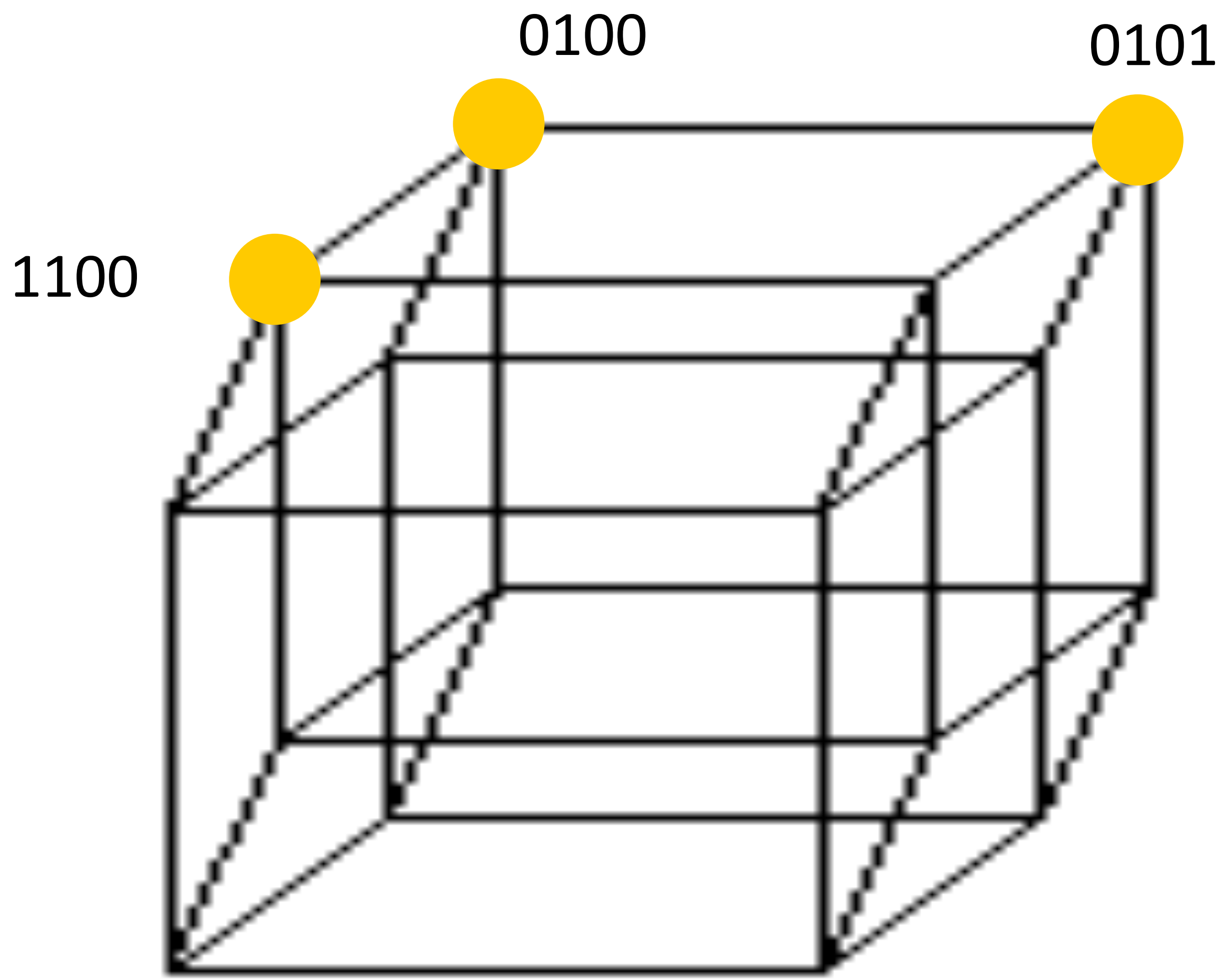
n=4

Motivation 1: An Election Interpretation

- Demographic and personal characteristics influence one's political preferences.
- Categories are binary (almost): male/female, married/single, urban/suburban etc.
- We can add positions on issues: pro-life/pro-choice, gun-rights/gun-control etc. And also other seemingly irrelevant attributes.
- Possible combination of values of n characteristics correspond to the vertices of n-dimensional hypercube.

An Election Interpretation

- Meet our voters:
- (sex, marital status, urban?, religious?)

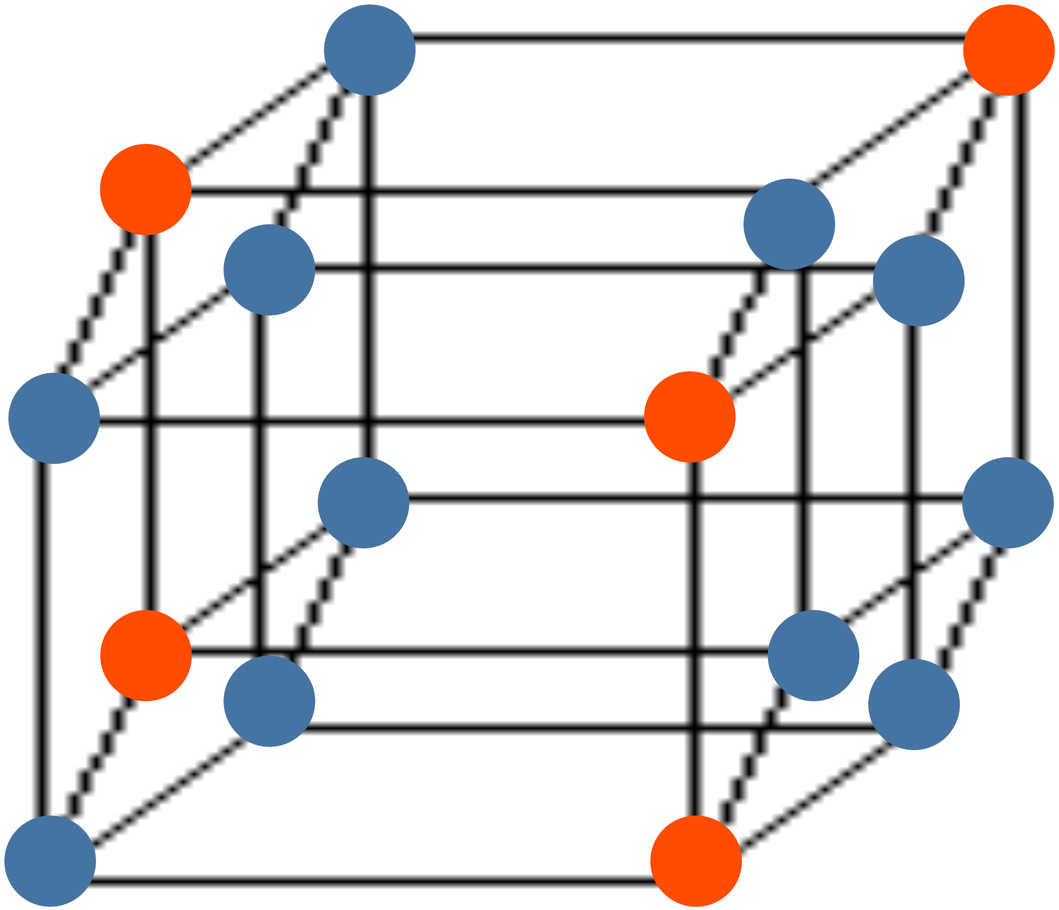


An Election Interpretation

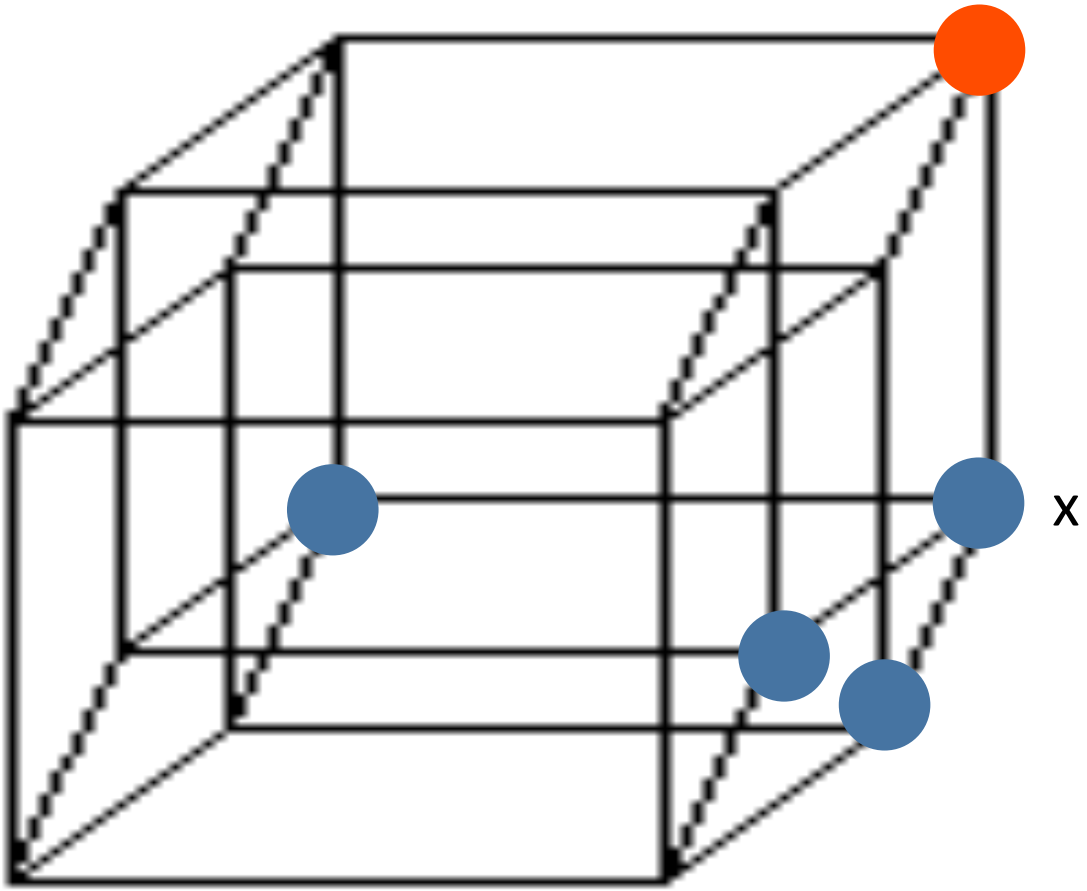
- A combination of x characteristics is “typical” for a party if you vary some of them (any number k of them) you still find mostly people who vote for that party.

- Do typical voters exist?

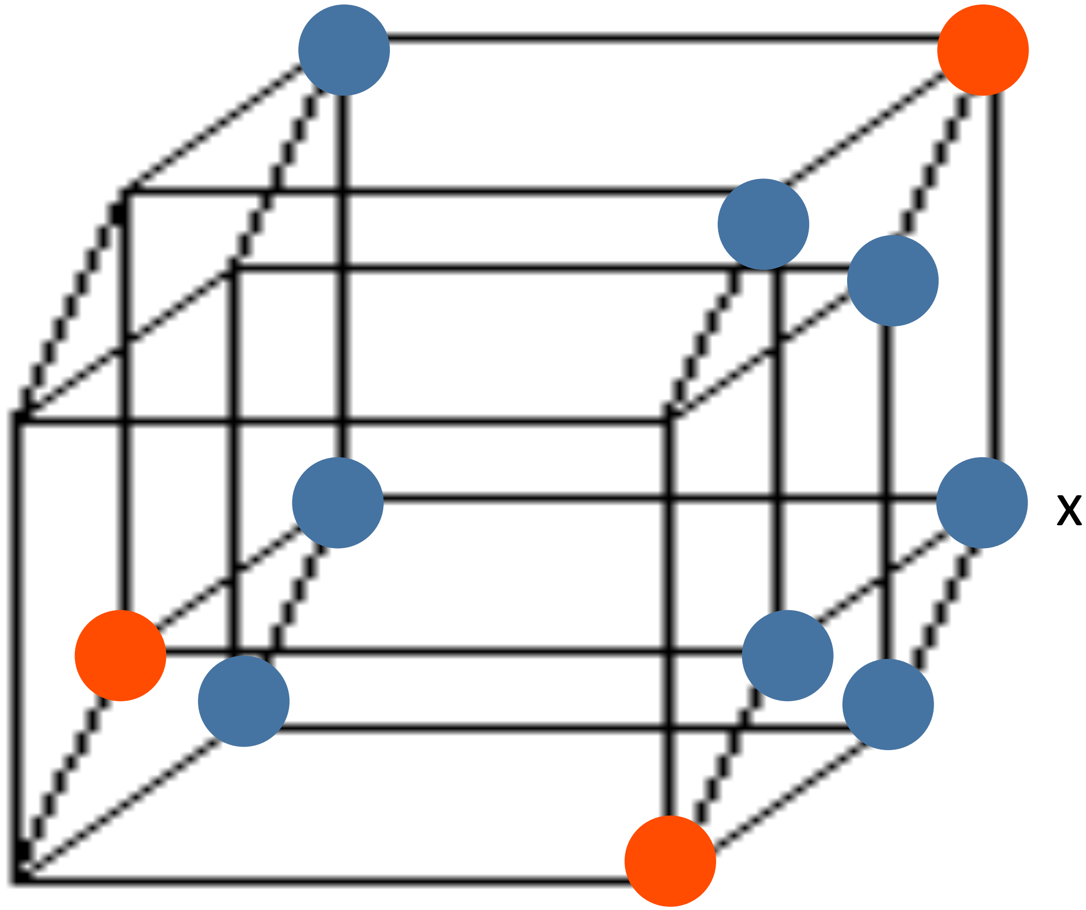
An Election Interpretation



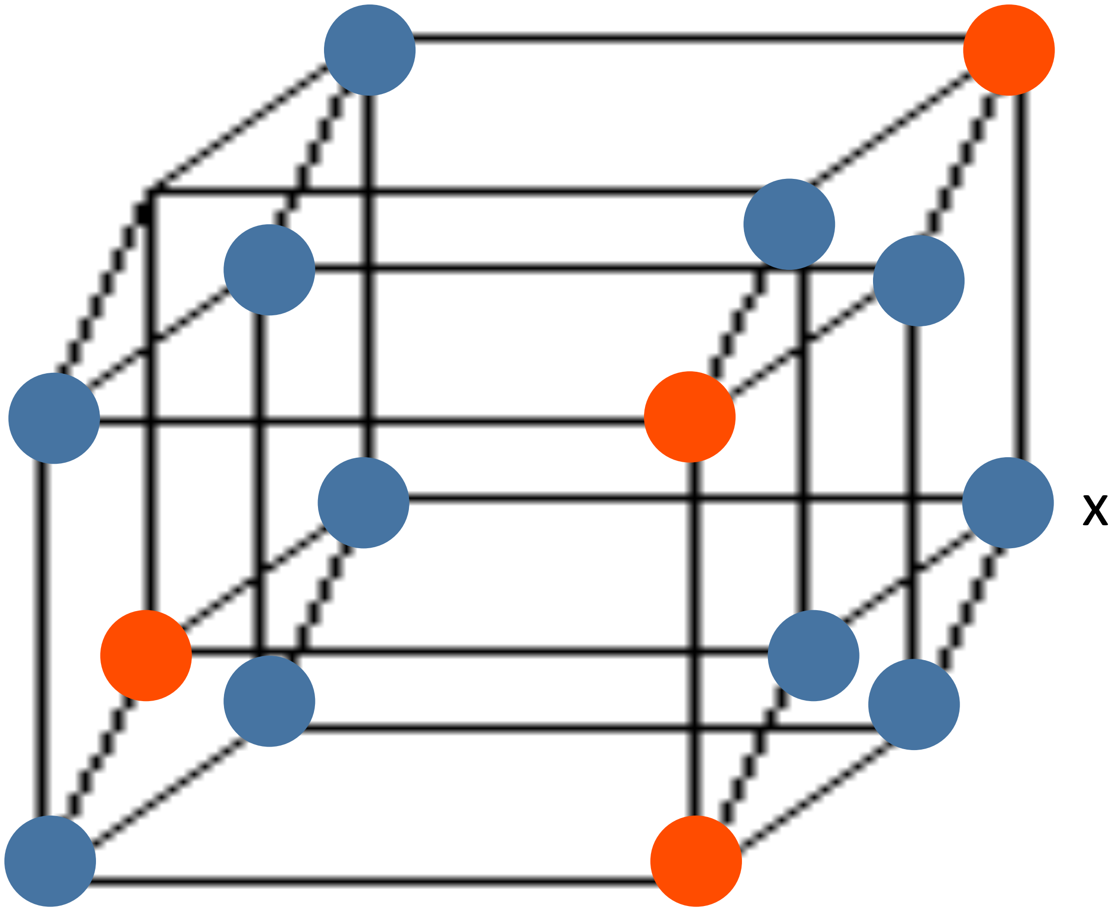
An Election Interpretation



An Election Interpretation



An Election Interpretation



An Election Interpretation

- "If a party wins with large enough landslide, then it has typical voters".

Motivation 2: UGC or "Can We Hope for Better Approximation Algorithms in P?"

Unique Games Conjecture (UGC) captures **exact** inapproximability of many more problems

Problem	Best Approximation Algorithm Known	UGC-Hardness
MaxCut	0.878[GW94]	0.878 [KKMO07]
Vertex Cover	2	$2-\varepsilon$ [KR06]
Max k-CSP	$\Omega(k/2^k)$[CMM07]	$O(k/2^k)$[ST,AM,GR]

Unique Games = Unique Label Cover Problem

Given: set of constraints

Linear Equations mod k :

$x_i - x_j = c_{ij} \bmod k$

GOAL

k =“alphabet” size

Find labeling that satisfies **maximum number of constraints.**

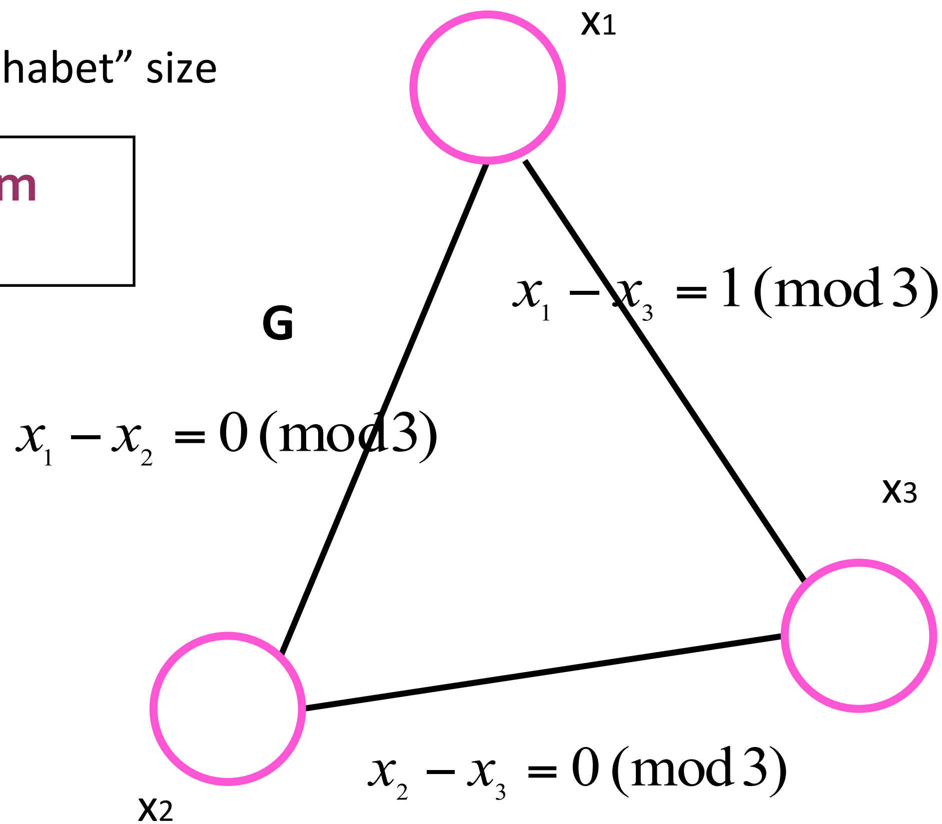
EXAMPLE

$$x_1 - x_2 = 0 \pmod{3}$$

$$x_2 - x_3 = 0 \pmod{3}$$

$$x_1 - x_3 = 1 \pmod{3}$$

The constraint graph



Unique Games , an Example

Given: set of constraints

Linear Equations mod k :

$x_i - x_j = c_{ij}$ mod k

GOAL

k ="alphabet" size

Find labeling that satisfies **maximum number of constraints.**

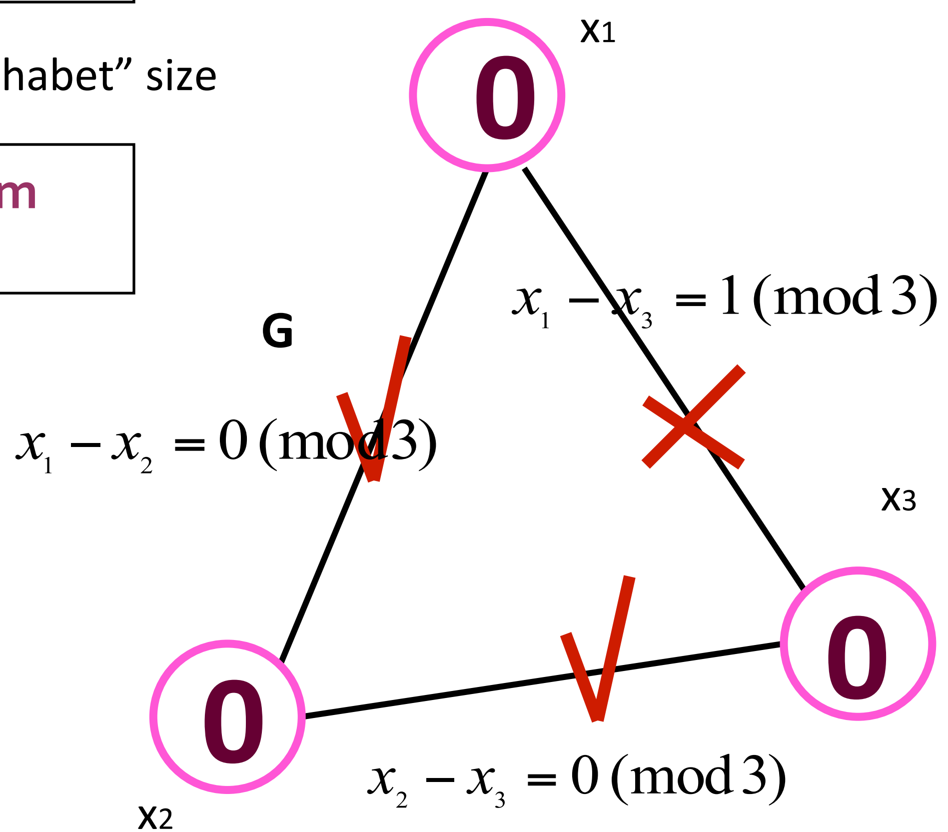
EXAMPLE

$$x_1 - x_2 = 0 \pmod 3$$ ✓

$$x_2 - x_3 = 0 \pmod 3$$ ✓

$$x_1 - x_3 = 1 \pmod 3$$ ✗

The constraint graph



Satisfy 2/3 constraints

Unique Games Conjecture

- **[Khot'02]** For every $\varepsilon, \delta > 0$ there is a (large enough) $k = k(\varepsilon, \delta)$ such that given an instance of Unique Games with alphabet size k it is NP-hard to distinguish between the two cases:

  (1) $OPT > 1 - \varepsilon$

  (2) $OPT < \delta$

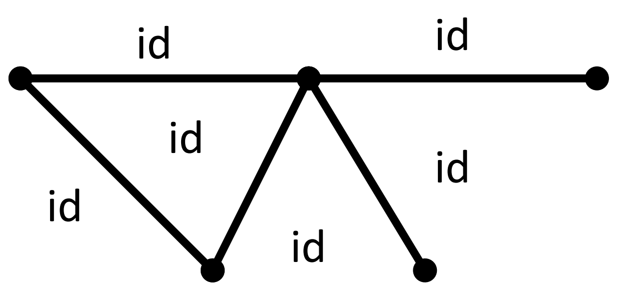
Unique Games Conjecture

- UGC: given a UG instance (graph and set of constraints over alphabet of size k) with the guarantee that it is 99% satisfiable, it is NP-hard to find an assignment that satisfies more than 1% of the constraints.

Really embarrassing not to know, since solving systems of linear equations (exactly) is very easy!

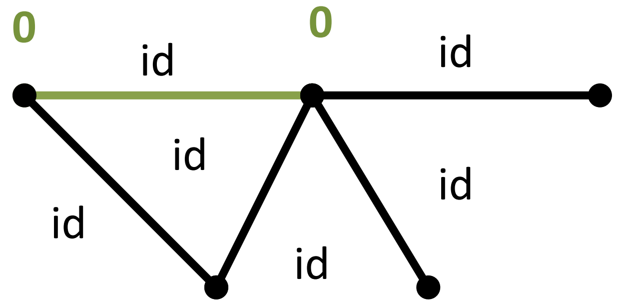
Fully Satisfiable Instances

- Can do it with a propagation algorithm: start with good value, follow constraints across edges.



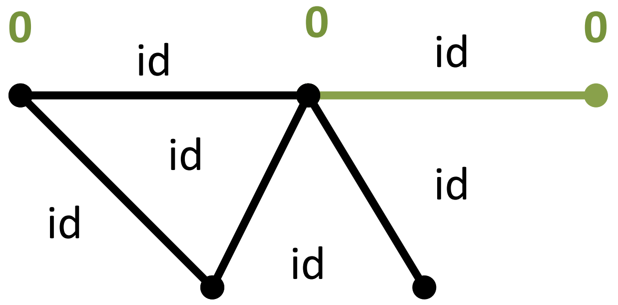
Fully Satisfiable Instances

- Can do it with a propagation algorithm: start with good value, follow constraints across edges.



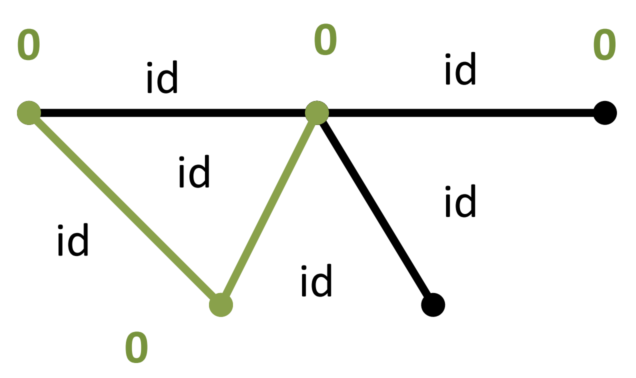
Fully Satisfiable Instances

- Can do it with a propagation algorithm: start with good value, follow constraints across edges.



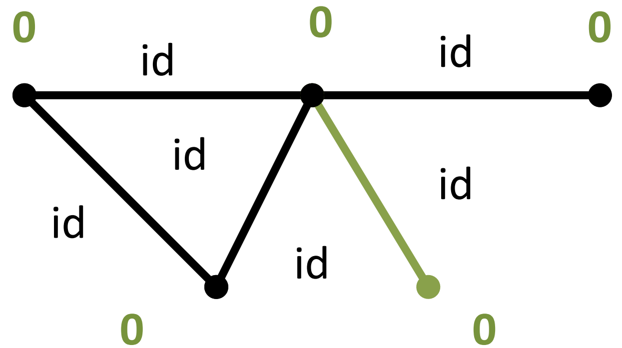
Fully Satisfiable Instances

- Can do it with a propagation algorithm: start with good value, follow constraints across edges.



Fully Satisfiable Instances

- Can do it with a propagation algorithm: start with good value, follow constraints across edges.



Unique Games Conjecture

- Really embarrassing not to know since solving systems of linear equations is easy.
- Can do it with a propagation algorithm: start with good value, follow constraints across edges.
- Sharp boundary comes from taking an easy problem and changing it a bit, makes it hard.

Where to begin if we want to refute UGC?

- Several attempts in recent years to refute or prove UGC.
- Lot of progress but still no consensus.

Plan of attack: start ruling out cases.

Easy Instances

- Classify graphs according to their "spectral profile" (eigenvalues)
- Expanders [AKKTSV'08,KT'08],
- Local expanders, graphs with relatively few large eigenvalues [AIMS'09,SR'09,K'10,ABS'10]

Easy Distributions

- Find distributions that are hard?
  - Random Instances : NO! Follows from expander result.
  - Quasi-Random Instances? [KMM'10] NO!

Where to begin if we want to refute UGC?

Plan of attack: start ruling out cases.

Easy Instances

- Classify graphs according to their "spectral profile" (eigenvalues)
- Expanders [AKKTSV'08,KT'08],
- Local expanders, graphs with relatively few large eigenvalues [AIMS'09,SR'09,K'10,ABS'10]
- Easy when very many large eigenvalues as well [separators, ABS'10]

UGC and the Spectrum of General Graphs

- How "easy" the graph is, depends on the number of large eigenvalues of the adjacency matrix.

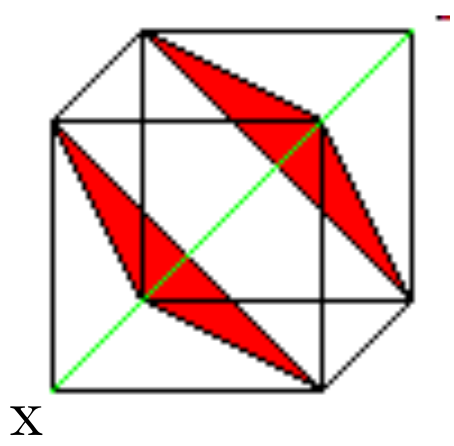
- Can solve previously "hardest" cases, where all other techniques failed.

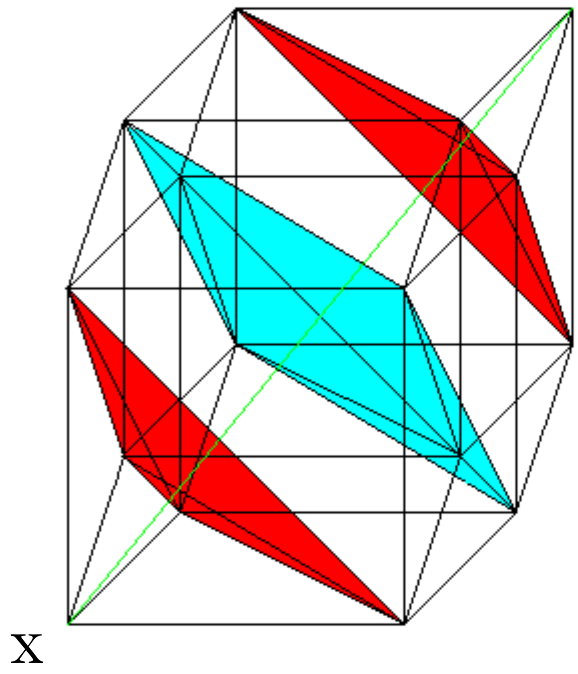
- Essentially only one class of graphs left, largely reflected by the Boolean Hypercube!!

Spheres in H^n

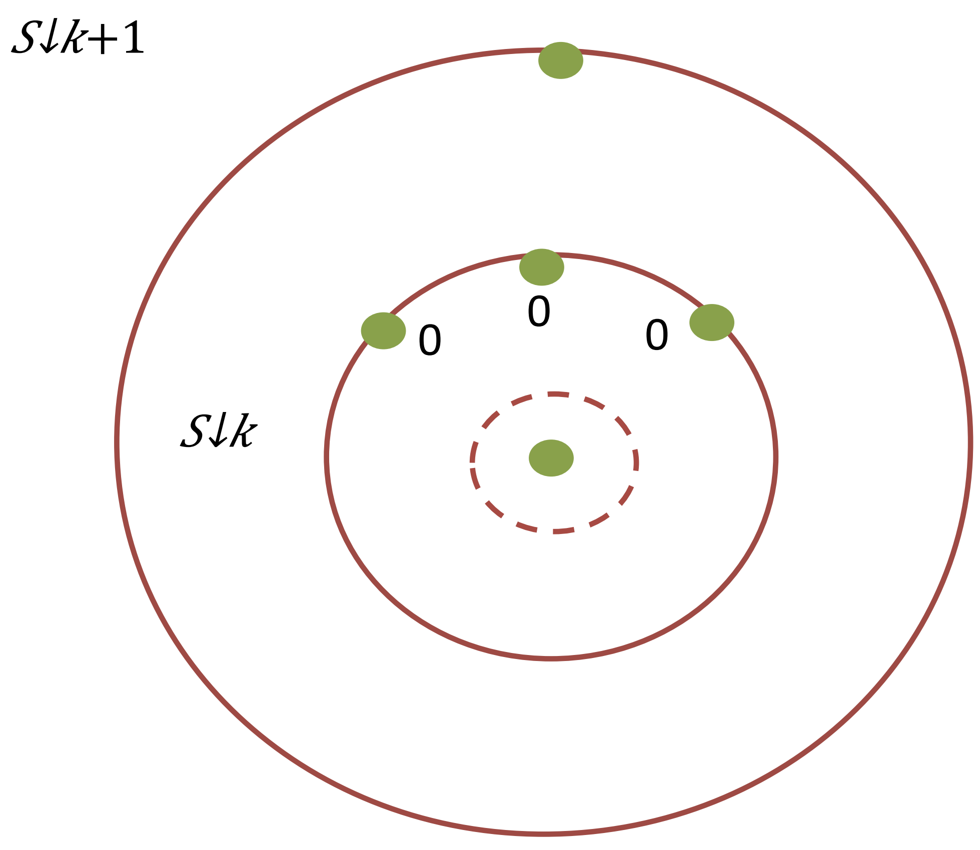
d : Hamming distance

$S(x,r)=\{y: d(x,y)=r\}$

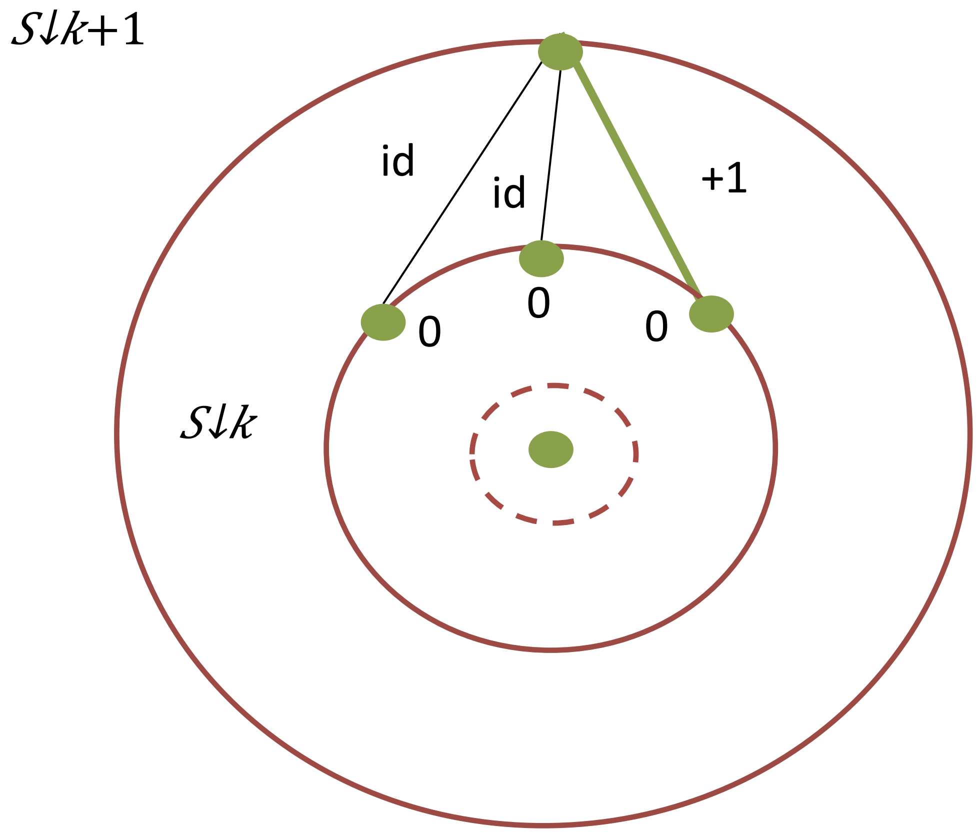




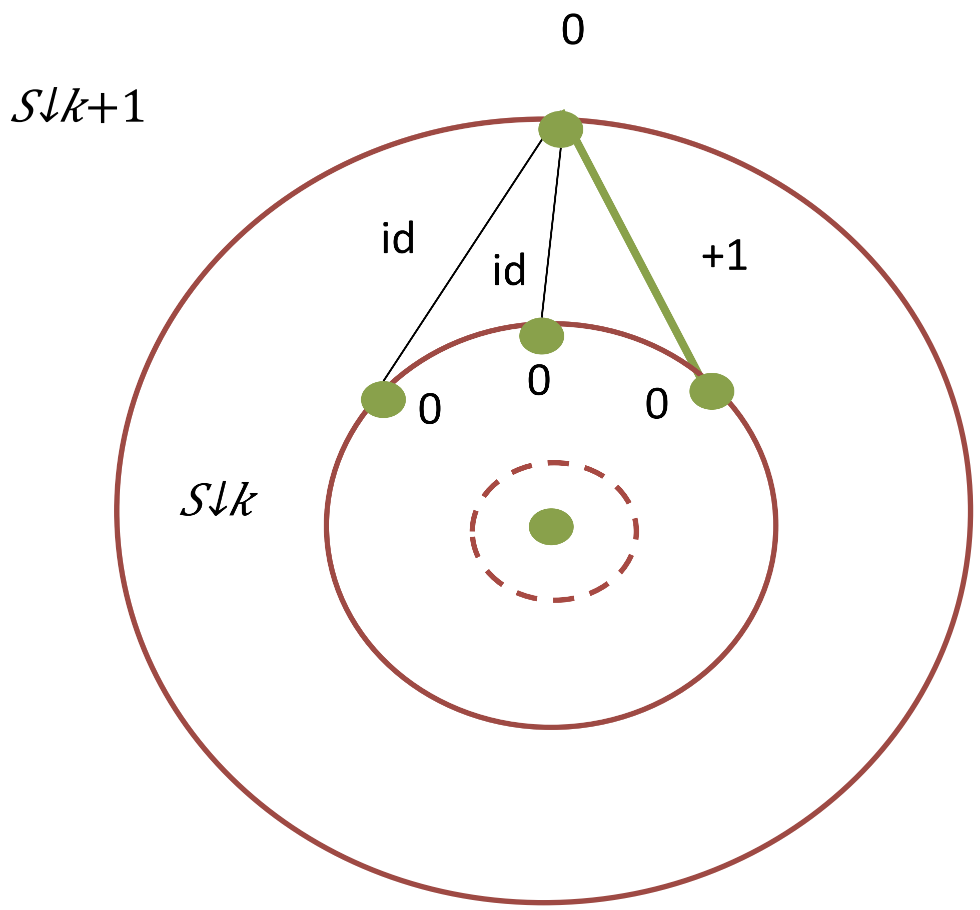
An Algorithm



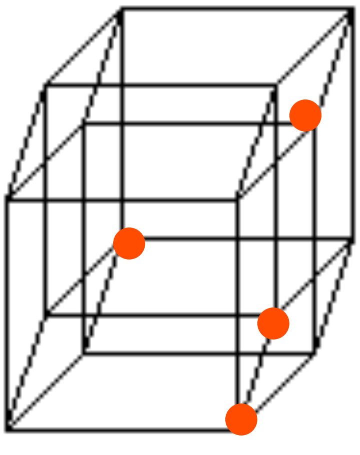
An Algorithm



An Algorithm



The Adversary



The adversary can “spoil” any ε fraction of the vertices of H^n, making them **bad**:

$$B \subseteq H^n$$

$$|B| = \varepsilon\, 2^n$$

Fix a threshold **λ > ε**.

Say sphere **S(x,r) is bad** if fraction >λ of it is bad:

$$|S(x,r) \cap B| > \lambda |S(x,r)|$$

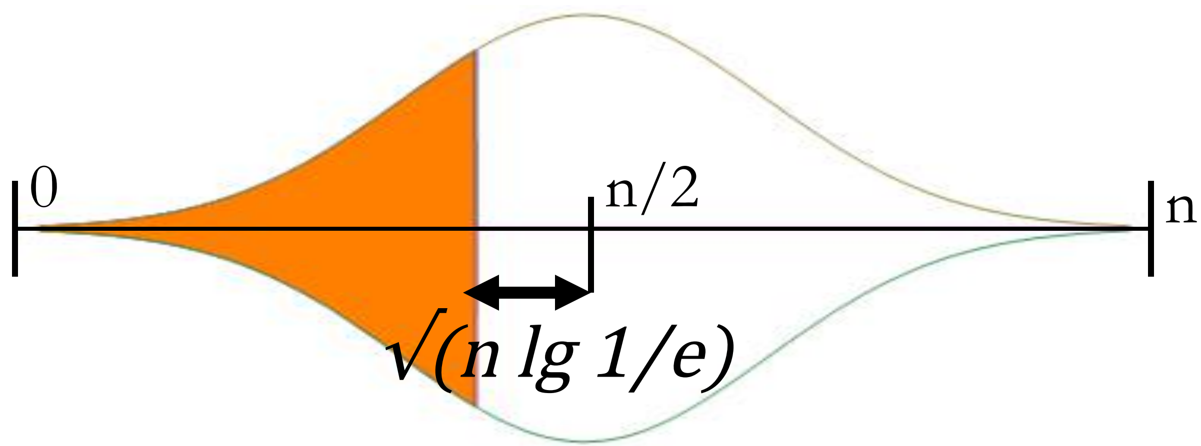
Say point **x is ruined** if there is some $0 \le r \le n$ for which S(x,r) is bad.

Consider ε,λ as small constants.

The Question

Can the adversary ruin *all* vertices x?

First attempt: spoil a metric ball.



- Ruins only the bad set plus a boundary zone of width approx $\sqrt{(n \lg 1/e)}$

The Question

Second attempt: spoil a subcube.

B = all vertices of form

{00000000***********************}

(brace under "00000000": lg $1/\varepsilon$ coordinates)

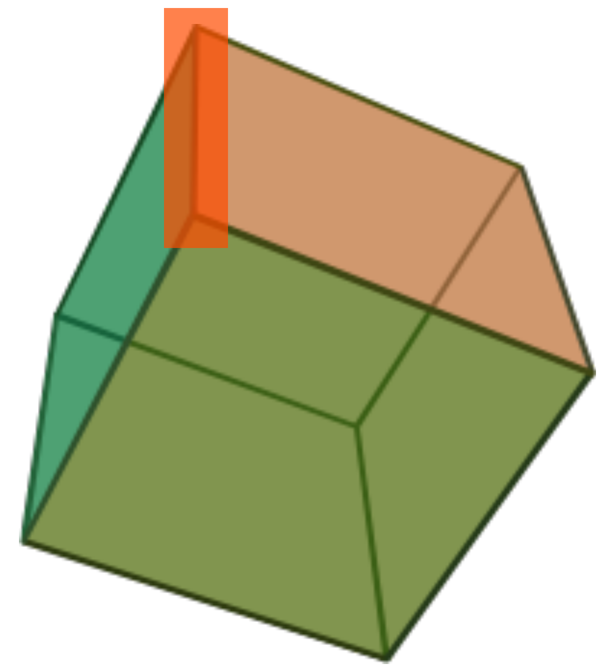
Ruins only "parallel" subcubes within distance approx

$$(\log 1/\lambda)/(\log \lg 1/\varepsilon)$$

$$\ll \lg 1/\varepsilon$$

of the bad subcube.

E.g.,
{01000100***********************}.



The Conjecture

We couldn't find any worse examples than these.

So naturally, we applied the method of *mathematician's induction:*

Conjecture: Nobody else can, either.

More precisely:

Conjecture: For all $\lambda < 1$ there is an $\varepsilon > 0$ s.t. for all n and for all $|B| < \varepsilon 2^n$, $|\text{Ruined set}| < 2^n$ ($\lambda = \sqrt{\epsilon}$ works).

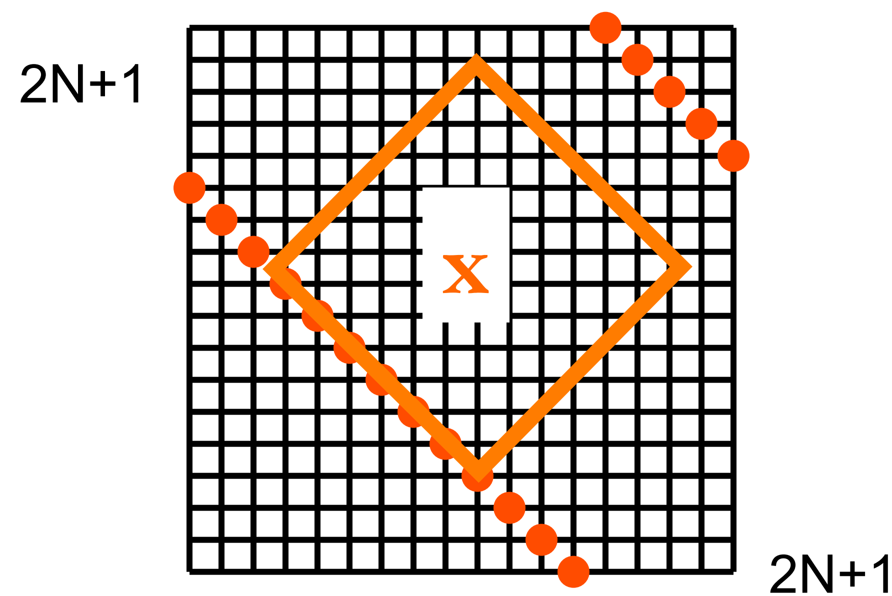
Dimension-independent.

This **theorem** is our main result.

The Challenge

What makes the problem hard:
1) Theorem is is false for closely related graphs

2D Torus (roughly NxN vertices)
$|B| = \{l: \sum l_i = 0 \; mod \; N$
$\sum l_i \leq 2N + 1\}$
$|B| = O(N) = \sqrt{(\# \text{ vertices})}$
(We have $|V|$ about $O(N^2)$)



For *any* vertex x, ¼ of the sphere
(of some radius) is contained in B.
Ruined set = entire torus.
Spectacular failure because $|B| <<$ any constant fraction.

cont.: what makes this problem hard?

(2) that the problem naïvely calls for a union bound over radii, but the union bound fails:

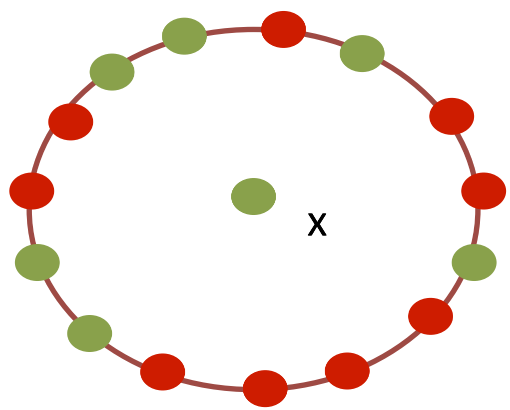
(a) Use Markov inequality:

$|R_r| = |\{x \text{ ruined by its sphere of radius } r\}| < (\varepsilon/\lambda)2^n$

(b) Use union bound: $\Sigma_r |R_r| < (\varepsilon/\lambda)\, n\, 2^n > 2^n$

A useless bound.

$S\downarrow r\uparrow$



cont.: what makes this problem hard?

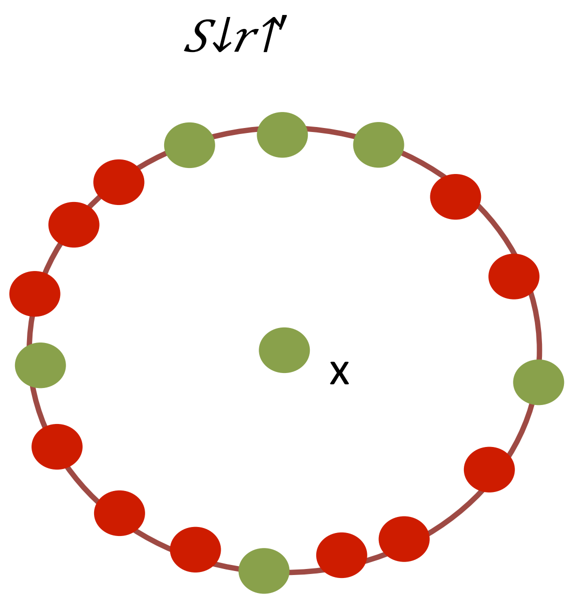
(2) that the problem naïvely calls for a union bound over radii, but the union bound fails:

(a) Use Markov inequality:

$|R_r| = |\{x \text{ ruined by its sphere of radius } r\}| < (\varepsilon/\lambda)2^n$

(b) Use union bound: $\Sigma_r |R_r| < (\varepsilon/\lambda)\, n\, 2^n > 2^n$

A useless bound.



cont.: what makes this problem hard?

(2) that the problem naïvely calls for a union bound over radii, but the union bound fails:

(a) Use Markov inequality:

$|R_r| = |\{x \text{ ruined by its sphere of radius } r\}| < (\varepsilon/\lambda)2^n$

(b) Use union bound: $\Sigma_r|R_r| < (\varepsilon/\lambda)\ n\ 2^n > 2^n$

A useless bound.

The problem is in step (b), the union bound.

Consider the subcube example:

$B \subseteq R_r$ for all $0 \leq r \leq n/(2 \lg 1/\varepsilon)$

$\Sigma_r|R_r| > \varepsilon\ 2^n\, n/(2 \lg 1/\varepsilon) > 2^n$.

The union bound is off because these sets R_r are almost identical. *Need to show this is always what happens.*

Convert to a problem in Analysis

$L_2(\mathsf{H}^n)$ = real-valued functions on the hypercube, with norm

$\|f\| = \sqrt{\Sigma_x |f(x)|^2}$.

Represent B by its indicator function:

f(x)=1 if x $\in$ B, f(x)=0 otherwise.

$$|B| = \|f\|\uparrow 2 .$$

More generally for any f and $\lambda > 0$, have Markov inequality:

$$|\{x: f(x)>\lambda\}| < \|f\|^2/\lambda^2.$$

Convert to a problem in Analysis

Now consider any **operator**

$$S: L_2(\mathsf{H}^n) \to L_2(\mathsf{H}^n)$$

If S has bounded operator norm, $A < \infty$:

$$\|Sf\|\uparrow < A \cdot \|f\|\uparrow \quad \text{for all } f$$

Then $$|\{x: (Sf)(x) > \lambda\}| < A^2 \|f\|\uparrow 2 / \lambda^2.$$

A problem in Analysis

Let $\mathsf{S} = \{S_r\}$ be the collection of all spherical mean operators.

$$(S_r f)(x) = \frac{\Sigma_{y \in S(x,r)} f(y)}{|S(x,r)|}$$

In order to talk about the union bound, introduce the **maximal operator M:**

$$M_{\mathsf{S}}: L_2(\mathsf{H}^n) \rightarrow L_2(\mathsf{H}^n)$$

$$(M_{\mathsf{S}} f)(x) = \max_r (S_r f)(x)$$

Connection to our problem:

$$\text{Ruined set} = \cup R_r = \{x: (M_{\mathsf{S}} f)(x) > \lambda\}$$

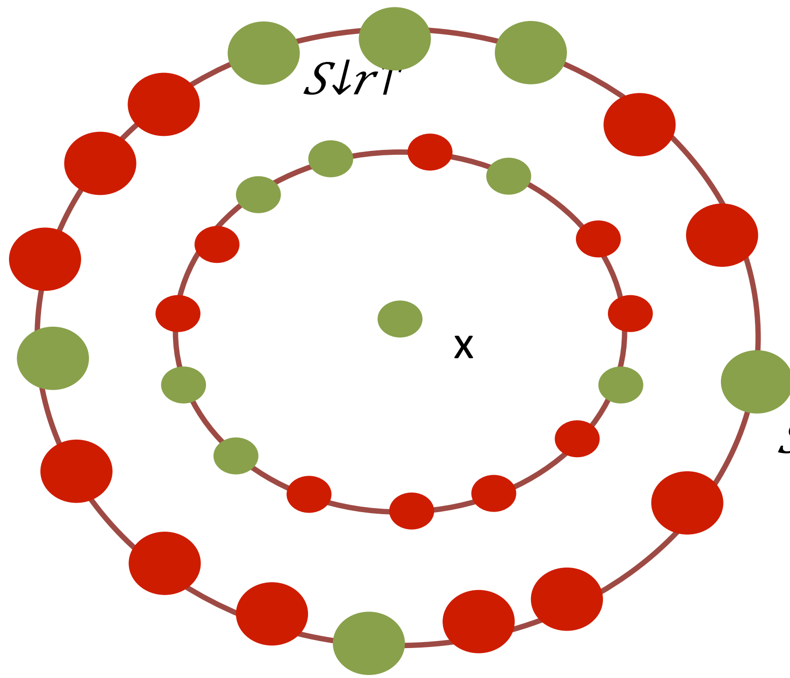
M_{S} is a *sublinear* operator.

Maximal Operator

$(M_S f)(x) = \max_r (S_r f)(x)$

S_r

$S_r f(x) = 11/17 = 0.64$



$S_{r'} f(x) = 9/15 = 0.6$

$M_S f(x) = 0.64$

Maximal Inequalities

Our conjecture will follow from showing:

(*) Theorem: M_S has bounded operator norm, $A < \infty$:

$$\|M_S f\| < A\|f\| \quad \text{for all } f$$

because then

$$|\text{Ruined set}| < A^2 \|f\|^2 / \lambda^2 = A^2 \varepsilon\, 2^n / \lambda^2$$

Taking $\lambda = 2A\sqrt{\varepsilon}$, we'll have

$$|\text{Ruined set}| < 2^{n-2}.$$

The statement (*) is called a maximal inequality.

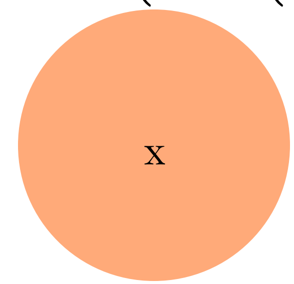
Maximal Inequalities: a little history

Hardy and Littlewood studied means operators for *balls* in Euclidean space E^n:

$$\mathrm{Ball}(r)=\{y:\ //y//\uparrow < r\}$$

Ball mean operator: $(B_r f)(x) = \left(\int_{\mathrm{Ball}(r)} f(x+y)\, dy\right) / \mathrm{Vol}(\mathrm{Ball}(r))$

Maximal operator for balls:



$$(M_{\mathsf{B}} f)(x) = \sup_r\ (B_r f)(x)$$

$$M_{\mathsf{B}}: L_2(\mathsf{E}^n) \to L_2(\mathsf{E}^n)$$

Hardy-Littlewood "weak type" inequality + Marcinkiewicz give:

(*) $$//M\downarrow B\ //\uparrow < A(n) < \infty.$$

We can't use this: wrong metric space, balls rather than spheres, bound not dimension independent.

Maximal Inequalities: a little history

It *would* be sufficient to have a similar result for spherical means in $\mathbf{R}^n$ with L_1 metric --- but as we already saw earlier (discrete version), this is false.

Something is special about $\mathbf{H}^n$ that does not hold for general L_1 metrics.

But other tools developed in the history of the subject are essential ingredients of our proof. Key contributors: Zygmund, Hopf, Kakutani, Yosida, Dunford, Schwartz, Garsia, Stein, Strömberg, Bourgain, Carbery, Naor, Tao...

Spherical-Mean Maximal Inequality: method

- Two main steps.

- Each step we obtain a maximal inequality for one class of operators based on comparison with another more tractable class.

Step 1 of Proof

- Step 1:

"Senate operators" of **S** are the stochastic operators:

$$\mathrm{Sen}(\mathbf{S})_r = (1/(r+1)) \Sigma_{0 \leq k \leq r} S_k$$

$$(M_{\mathrm{Sen}(\mathbf{S})} f)(x) = \max_r (\mathrm{Sen}(\mathbf{S})_r f)(x)$$

We use Stein's comparison method:

$$\|M{\downarrow}\mathbf{S}\| < O(\|M{\downarrow}Sen(\mathbf{S})\| + \|R{\downarrow}\mathbf{S}\|)$$

- Rs error term that we need to bound.

Step 2 of Proof

- Step 2:

"Noise operators" $N=\{N_t\}_{t\geq 0}$

$$N_t = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} S_k$$

Where $p=(1-e^{-t})/2$

- $N_t f(x)$ is the expectation $E[f(y)]$, where y is obtained by running n independent Poisson processes with parameter 1 from time 0 to t and flipping the i-th bit iff there are odd number of events in the i-th process.
- Equivalent to Poisson clocked random walk on cube.

Step 2 of Proof

- Step 2:

"Noise operators" $\boldsymbol{N} = \{N_t\}_{t \geq 0}$

$$N_t = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} S_k$$

Where $p = (1-e^{-t})/2$

- We show by direct point-wise comparison

$$\|M_{Sen}(\boldsymbol{S})\| < O(\|M_{Sen}(\boldsymbol{N})\|)$$

- Use known result: $\|M_{Sen}(\boldsymbol{N})\| \leq 2\sqrt{2}$

Step 1+ Step 2

$$\|M\downarrow \boldsymbol{S} f\| \leq O(\|M\downarrow Sen(\boldsymbol{S}) f\| + \|\boldsymbol{f}\|) \leq$$
$$O(\|M\downarrow Sen(\boldsymbol{N}) f\| + \|\boldsymbol{f}\|) \leq O(\|\boldsymbol{f}\|)$$

Step 1 of Proof

$$\|M_{\mathbf{S}}\| < O(\|M_{Sen(\mathbf{S})}\| + \|R_{\mathbf{S}}\|)$$

- Bounding the norm of $R_{\mathbf{S}}$:

(a) Stein's application of Cauchy-Schwartz,

(b) Spectral bounds on the family S.

- S_k resembles $N_{k/n}$ (since N_t approx the average of S_k for $k = nt \pm \sqrt{nt(1-t)}$.
- While direct comparison is difficult, we argue that spectra of those two operators are similar.

Step 1 of Proof

- S_k resembles $N_{k/n}$ (since N_t approx the average of S_k for $k=nt\pm\sqrt{nt(1-t)}$).
- While direct comparison is difficult, we argue that spectra of those two operators are similar.
- N_t has evals $(1-2t)^x$ for character χ_y, $|y|=x$.
- S_k has evals $kraw_k(x)$ for character χ_y, $|y|=x$.
- Show that $kraw_k(x)$ has similar behavior to $(1-2k/n)^x$.

- **Lemma**: For k,x ≤ n/2, $kraw_k(x) \leq \exp(-\Omega(k\cdot x/n))$.

Step 2 of Proof

- Step 2:

"Noise operators" $\boldsymbol{N}=\{N_t\}_{t\geq 0}$

$$N_t = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} S_k$$

Where $p=(1-e^{-t})/2$

- We show by direct point-wise comparison

$$\|M_{Sen}(\boldsymbol{S})\| < O(\|M_{Sen}(\boldsymbol{N})\|)$$

- Use known result: $\|M_{Sen}(\boldsymbol{N})\| \leq 2\sqrt{2}$

Step 2 of Proof

$\|M{\downarrow}Sen(\boldsymbol{S})\|$ bound. Base this upon

Ergodic maximal inequalities:

T = doubly stochastic matrix

Form the semigroup

$$\mathsf{T} = \{T^r\} \quad (r > 0)$$

"Senate operators" built from T:

$$Sen(\mathsf{T})_r = (1/r)\, \Sigma_{1 \le k \le r}\; T^k$$

Kakutani, Yosida, Hopf, Dunford, Schwartz: under (hypotheses we satisfy),

(*) $$\|M{\downarrow}Sen(\boldsymbol{T})\| < \infty.$$

Spherical-Mean Maximal Inequality: method

Specifically

$$\mathrm{Sen}(\mathbf{N})_T = (1/T) \int_0^T N_t \, dt$$

Have

(*) $$\| M_{Sen(N)} \| < \infty.$$

Intuition: while N_t is very different from S_r, $\mathrm{Sen}(\mathbf{N})_T$ is not so different from $\mathrm{Sen}(\mathbf{S})_r$

Final piece of puzzle:

$$\| M_{Sen(\mathbf{S})} \| < \| M_{Sen(N)} \|$$

by showing stochastic domination of the set $\mathrm{Sen}(\mathbf{S})$ by the set $\mathrm{Sen}(\mathbf{N})$.

Future

Applications?

UG on Hypercube?

Other graphs where maximal inequality holds?

THANK YOU!