UNIQUE GAMES CONJECTURE

- Strengthening of $P \neq NP$ hypothesis
- posits hardness of a natural problem
- yields optimal hardness results for Max-Cut, Vertex Cover, ...
- truth not yet known
- Amazingly rich theory with interplay between
  1) hardness of approx.
  2) algo design via SDPs.
  3) Methods from Analysis of Boolean Functions, Geometry....

The Problem & the Conjecture

Max 2-LIN(p)

p → field.

2 → # vars in each equation

Input: equations of the form

$$\left\{ x_i + x_j = b \bmod p. \right.$$

2 vars in each equation

all mod p

Goal: Satisfy max frac of them

Def (Value): max frac of constraints satisfiable

Random Assignment: $\frac{1}{p}$

Theorem: If value = 1, can find

a sat assignment in poly time.

<u>Proof</u>: Solve linear equations via Gaussian elimination...

<u>CONJECTURE</u> [Khot'02]

$\forall \varepsilon > 0$, $\forall p$ large enough;

$(1-\varepsilon, \varepsilon)$ – MAX 2LIN(p) is NP-hard.

"It is NP-hard to 'find an ε sat assignment for a $1-\varepsilon$ sat instance".

<u>Best poly time algo</u>: $1-\varepsilon$, $1-O(\sqrt{\varepsilon \log k})$

"dependent on alphabet size"...

Can beat brute-force search:

Arora-Barak-Steurer '10

$2^{n^{\varepsilon}}$ time algo, if opt $= 1-\varepsilon$, round $= 1-\varepsilon^{1/6}$.

"indep of alphabet size"

Lots of interesting work on both algos & lower bounds...

Most recently.

Thm [2018] [2-to-2 Games Thm]

$(\frac{1}{2}, \varepsilon)$-UG is NP-Hard.

$1-\varepsilon \Rightarrow$ U.G.C.

Today & next 2 classes

→ a glimpse of this theory.

We will prove optimal NP hardness of Max-Cut assuming the U.G.C.

We will do it in a way that shows that there's a principled theory of such optimal hardness reductions.

And this theory directly builds hardness reductions from Integrality Gaps for SDP for a large class of problems

FAILURE OF SDP → UG Hardness...

Theorem [Khot Kindler Mossell O'donnell '04]

$\forall \varepsilon > 0$, $\alpha_{GW} + \varepsilon$ approx.

Max-Cut is NP-Hard assuming the U.G.C.

$\min_{\rho<0} \frac{\text{arc-cos}(\rho)}{\pi(\frac{1}{2}-\frac{1}{2}\rho)}$

Like the examples we saw in last lecture, this theorem we will prove this theorem via "gadget" reductions

"Gadget Reductions": $Q: P_1 \to P_2$.

P_2 instance = "local transformation" of instance for P_1.

modify constant size portions

E.g. Reduction for Vertex Cover that replaces each clause by 7 vertices

local.

each edge of the resulting IS instance depends only on 2 clauses in the input 3SAT instance.

Our gadgets need to get the exact $\frac{1}{2}-\frac{1}{2}\rho \to \frac{\text{arc-cos}(\rho)}{\pi}$ gap. So need some "geometry" to show up.

In fact, our gadgets will in a precise sense come from our Integrality Gap & Rounding Gap Instances for Max Cut.

Let me remind you of those...

Both were "embedded graphs" → each vertex came with a vector labeling it.

Integrality Gap (Feige Schechtman graph)

Vertices = (discretization) of unit sphere in d-dim

edge distribution: pick $\vec{u}, \vec{v}$ uniformly from S^{d-1} conditioned on $\langle \vec{u}, \vec{v} \rangle \leq \rho_*$.

SDP OBJ: $\mathbb{E}_{\{\vec{u},\vec{v}\} \sim \text{edge}} \left[\frac{1}{2} - \frac{1}{2} \langle \vec{u}, \vec{v} \rangle \right] = \frac{1}{2} - \frac{1}{2} \rho_*$.

Analysis: ① hemisphere cuts are optimal for this graph

② hemisphere cuts have value $\sim \frac{\text{arc-cos}(\rho_*)}{\pi}$

Same analysis as our rounding.

Rounding Gap

1) Vertices: Corners of hypercube $\{\pm \frac{1}{\sqrt{d}}\}^d$ scaled down to be unit vectors

2) edge dist^n: 1) pick $\vec{u}$ at random
2) pick $\vec{v}$ by flipping each coordinate of $\vec{u}$ indep w.p. $\frac{1}{2} - \frac{1}{2}\rho_*$.
3) output $\{\vec{u}, \vec{v}\}$.

$\vec{\mathbb{E}}\langle \vec{u}, \vec{v} \rangle = \rho_*$

SDP OBJ $= \frac{1}{2} - \frac{1}{2}\rho_*$.

True Max-Cut:, also $\frac{1}{2} - \frac{1}{2}\rho_*$

$$D_i = \{\vec{u} \mid \vec{u}_i = \frac{+1}{\sqrt{d}}\}.$$

$$\Pr_{\{\vec{u}, \vec{v}\} \sim \text{edge}} [D_i(\vec{u}) \neq D_i(\vec{v})] = \frac{1}{2} - \frac{\rho_*}{2}$$

Our gadget for Max-Cut $\to$ hypercube graph.

to analyze cuts in such graphs, useful to adopt "function view".

Every subset of vertices S $\longleftrightarrow$ $f : \{\pm 1\}^d \to \{\pm 1\}$

" Cuts

$f(x) = +1 \to x \in S$

$f(x) = -1, \to x \notin S$.

Value of cuts

$$= \Pr_{\{x,y\} \text{ edge dist}^n} [f(x) \neq f(y)]$$

Need machinery to reason about such quantities

BASIC FOURIER ANALYSIS OF BOOLEAN FUNCTIONS

$f : \{-1, 1\}^n \to \mathbb{R}$

↳ Boolean Function.

Fourier analysis ~ write f as linear combination of some nice functions, use it to prove properties of f

INNER PRODUCT

Let $f : \{-1, 1\}^n \to \mathbb{R}$
$g : \{-1, 1\}^n \to \mathbb{R}$

$$\langle f, g \rangle = \mathbb{E}_{x \sim \{-1,1\}^n} f(x) \cdot g(x) = 2^{-n} \cdot \sum_x f(x) \cdot g(x)$$

"treat f, g as vectors of length 2^n, take inner products, rescale by 2^{-n}". Note: $\mathbb{E}_x x_i = 0 \; \forall i$

NORM

$$\|f\|_2^2 = \langle f, f \rangle = \mathbb{E}_x f(x)^2$$

Def (Parity functions / Monomials)

For any $S \subseteq [n]$,

$$\chi_S = \prod_{i \in S} x_i \text{ is the}$$

"parity" function or monomial on S.

If odd # bits in S are -1, then -1

even # bits in S are -1, then 1

Observation [Orthonormality]

$\mathbb{E}_x \chi_S = 0$ if $S \neq \emptyset$

If $S \neq T$,

$\langle \chi_S, \chi_T \rangle = 0.$

If $S = T$, $\langle \chi_S, \chi_S \rangle = \|\chi_S\|_2^2$

$= 1$

Proof:

$$\chi_S \cdot \chi_T = \prod_{i \in S} x_i \cdot \prod_{i \in T} x_T$$

$$= \prod_{i \in S \cap T} x_i^2 \cdot \prod_{i \in S \Delta T} x_i$$

$$= \prod_{i \in S \Delta T} x_i$$

If $S \Delta T \neq \emptyset$, $\mathbb{E}_x \prod_{i \in S \Delta T} x_i = \prod_{i \in S \Delta T} \mathbb{E}_x x_i$

$= 0$

Observation:

The set of functions $\{\chi_S \mid S \subseteq [n]\}$ form a orthonormal basis w.r.t. the inner product above.

Proof: We proved that each χ_S has length: $\|\chi_S\|_2^2 = 1$

& $\forall S \neq T$ $\langle \chi_S, \chi_T \rangle = 0$

& there are 2^n such functions

& the dim of space $\leq 2^n$.

$\square$

Thus each f can be expanded as lin comb. of χ_S.

Definition (Fourier Transform)

$$f : \{-1,1\}^n \to \mathbb{R}$$

$$\forall x : f(x) = \underbrace{\sum_{S \subseteq [n]} \hat{f}(S) \cdot \chi_S(x)}_{\text{Fourier expansion}}$$

Fourier coefficient

"writing f in the basis of χ_S".

Observation

Let $f : \{-1,1\}^n \to \mathbb{R}$. Then

$$\forall S, \quad \hat{f}(S) = \mathbb{E}_x f(x) \cdot \chi_S$$

Correlation of f & χ_S.

Proof:

$$f(x) = \sum_S \hat{f}(S)\, \chi_S(x)$$

$$\mathop{\mathbb{E}}_x f(x)\cdot \chi_T(x) = \mathop{\mathbb{E}}_x \sum_S \hat{f}(S)\cdot \chi_S(x)\chi_T(x)$$

$$= \sum_S \hat{f}(S)\cdot \underbrace{\mathop{\mathbb{E}}_x \chi_S(x)\cdot\chi_T(x)}_{\substack{0 \text{ if } S\neq T \\ 1 \text{ if } S=T}}$$

$$= \hat{f}(T)$$

Observation: $f:\{-1,1\}^n \to \mathbb{R}$

Then, $\mathop{\mathbb{E}}_x f(x)^2 = \sum_S \hat{f}(S)^2$

Proof:

$$f(x) = \sum_S \hat{f}(S) \cdot \chi_S(x)$$

$$\mathbb{E}_x f(x)^2 = \frac{1}{2^n} \cdot \sum_x \sum_{S,T} \hat{f}(S) \cdot \hat{f}(T) \cdot \chi_S(x) \cdot \chi_T(x)$$

$$= \sum_{S,T} \hat{f}(S) \cdot \hat{f}(T) \cdot \mathbb{E}_x \chi_S \cdot \chi_T$$

$$= \sum_S \hat{f}(S)^2 .$$

Influence of Functions

$\mathrm{Inf}_i(f)$: influence of i-th variable on f.

<u>Def</u> (for Boolean valued functions)

$f : \{-1,1\}^n \to \{-1,1\}$.

$$\mathrm{Inf}_i(f) = \Pr_{x \sim \{\pm 1\}^n} [f(x) \neq f(x^{(i)})]$$

$x^{(i)}$: flip ith bit of x.

$$= \mathop{\mathbb{E}}_{x \sim \{\pm 1\}^n} \frac{1}{4}(f(x) - f(x^{(i)}))^2$$

"prob that at a random x, flipping ith bit changes the value of f"

Lemma

Let $f : \{-1,1\}^n \to \mathbb{R}$. Then,

$$\frac{1}{4} \mathbb{E}_x \left(f(x) - f(x^{(i)})\right)^2 = \sum_{S \ni i} \hat{f}(S)^2$$

Proof:

$$f(x) = \sum_S \hat{f}(S) \cdot X_S$$

$$f(x^{(i)}) = \sum_S \hat{f}(S)\, X_S^{(i)}$$

$$= \sum_{S \not\ni i} \hat{f}(S) \cdot X_S - \sum_{S \ni i} \hat{f}(S) \cdot X_S$$

Thus, $f(x) - f(x^{(i)}) = 2 \sum_{S \ni i} \hat{f}(S) \cdot X_S$.

or $$\underbrace{\frac{1}{2}\left(f(x) - f(x^{(i)})\right)}_{= g(x)} = \sum_{S \ni i} \hat{f}(S) \cdot X_S$$

<u>Parseval:</u> $\mathbb{E}_x\, g(x)^2 = \sum_{S \ni i} \hat{f}(S)^2$

$\qquad\qquad\qquad\qquad \|$

$\frac{1}{4} \mathbb{E}_x \left(f(x) - f(x^{(i)})\right)^2$

$\square$

<u>Examples:</u>

1) Dictator Function

$f(x) = x_j \rightarrow$ depends only on j^{th} bit.

$\mathrm{Inf}_j(f) = 1$! Very influential

2) $f(x) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)$ "mean function"

$\frac{1}{4} \mathbb{E}_x \left(f(x) - f(x^{(i)})\right)^2 = \frac{1}{n^2}$. "low influence function"

3) $f(x) = \text{MAJ}(x)$, n: odd

$$= \text{sign}(\Sigma_i x_i).$$

When does flipping a bit change the value of MAJ?

<u>Ans</u>: when $\Sigma_i x_i = 1$

or $\Sigma_i x_i = -1$.

At such an x, any bit flipped will change the value.

$$\text{Prob}[\Sigma_i x_i = 1 \text{ or } -1] = \frac{\binom{n}{\frac{n-1}{2}}}{2^n} + \frac{\binom{n}{\frac{n+1}{2}}}{2^n}$$

$$\sim O\left(\frac{1}{\sqrt{n}}\right)$$

"low influence"

All n bits have influence $\sim 1/\sqrt{n}$.

4) <u>PARITY FUNCTION</u>

$$f(x) = \prod_{i=1}^{n} x_i$$

flipping any bit at any x changes the output of PARITY FUNCTION.

every bit has influence 1.

$$\mathrm{Inf}_i(f) = 1 \quad \forall i$$

LOW DEGREE INFLUENCE

<u>Def</u>

$$\mathrm{Inf}_i^{C}(f) = \sum_{\substack{S \ni i \\ |S| \leq C}} \hat{f}(S)^2$$

"discounting the effect of higher degree parity functions."

Obs^n:

only C bits can be influential now.

$$\mathrm{Inf}_i^C(X_S) = \begin{cases} 0 & \text{if } |S| > C \\ 1 & \text{otherwise} \end{cases}$$

So, if $|S|$ is large, the low degree influence is small.

Lemma (Only a small # vars can be influential)

Let $f : \{-1,1\}^n \to [-1,1]$.

Then, for any $C > 0$

$$|\{ i \mid \mathrm{Inf}_i^C(f) \geq \varepsilon \}| \leq \frac{C}{\varepsilon}$$

Proof

$$\sum_{i=1}^{n} \mathrm{Inf}_i^{c}(f)$$

$$= \sum_{i=1}^{n} \sum_{S \ni i, |S| \le c} \hat{f}(S)^2$$

$$\le c \underbrace{\sum_{S} \hat{f}(S)^2}_{\le 1} \le c$$

since $\sum_S \hat{f}(S)^2 = \mathbb{E}_x f(x)^2 \le 1$

So, by Markov's inequality,

frac of i: $\mathrm{Inf}_i^{c}(f) \geq \varepsilon$

$$\le \frac{c}{\varepsilon}$$