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) \to \text{field.}
\Rightarrow \text{# vars in each equation}

Input: equations of the form

\{ X_i + X_j = b \text{ 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.

Proof: Solve linear equations via Gaussian elimination --

**CONJECTURE [Khot’02]**

\[ \forall \varepsilon > 0, N \text{ large enough,} \]

\[ (1-\varepsilon, 3) - \text{MAX 2LIN}(p) \text{ is NP-hard.} \]

"It is NP-hard to "find an \( \varepsilon \) sat assignment for a 1-\( \varepsilon \) sat instance".

Best poly time algo: \( 1-\varepsilon, 1-O(\sqrt{\varepsilon \log k}) \)

Can beat brute-force search:
Arora-Bazak-Steuver’10

$2^{n^3}$ time algo, if $\text{opt} = 1-\varepsilon$, round $= \frac{1-\varepsilon}{16}$.

"indep of alphabet size"

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

Most recently:

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

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

$1-\varepsilon \implies \text{UGC}$.
Today & next 2 classes

→ a glimpse of this theory

We will prove hardness of Max-Cut assuming the UGC.

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 Mossel Odonnell ’04]

\[ f \approx 0 \implies d_{GW} + \varepsilon \text{ approx.} \]

\[
\text{Max-Cut is NP-Hard assuming the U.G.C.}
\]

\[
\min_{\varepsilon > 0} \frac{\pi(1/2 - \varepsilon/2)}{\pi(1/2)}
\]

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

"Gadget Reductions": \( R : P_1 \rightarrow P_2 \)

\( P_2 \) instance = "local transformation" of \( P_1 \) instance

- \( \text{modify constant size portions} \)
- \( \text{local} \)

E.g. Reduction for Vertex Cover that replaces each clause by 7 vertices
each edge of the resulting IS instance depends only on 2 clauses in the input 3SAT instance.

Our gadgets need to get the exact \( \frac{1}{2} - \frac{1}{2} \theta \rightarrow \frac{\text{arc-cos}(\theta)}{\pi} \) gap. So need some "geometry" to sharpen.

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.
Integral Gap (Feige Schechtman graph)

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

Edge distribution:  
- pick \( \mathbf{u}, \mathbf{v} \) uniformly from \( S^{d-1} \) conditioned on \( \langle \mathbf{u}, \mathbf{v} \rangle \leq \theta_\ast \).

\[
\text{SDP OBJ} : \mathbb{E} \left[ \frac{1}{2} - \frac{1}{2} \langle \mathbf{u}, \mathbf{v} \rangle \right]
\]

\[
\text{edge} = \frac{1}{2} - \frac{1}{2} \theta_\ast.
\]

Analysis:
1. Hemisphere cuts are optimal for this graph.
2. Hemisphere cuts have value \( \sim \frac{\text{arc-cos}(\theta_\ast)}{\pi} \).

Same analysis as our rounding.
Rounding Gap

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

2) Edges: 1) Pick \( \vec{u} \) at random.

2) Pick \( \vec{v} \) by flipping each coordinate of \( \vec{u} \) independently with probability \( \frac{1}{2} - \frac{1}{2} \delta \).

3) Output \( \{ \vec{u}, \vec{v} \} \).

- SDP OBJ = \( \frac{1}{2} - \frac{1}{2} \delta \).
- True Max-Cut: also \( \frac{1}{2} - \frac{1}{2} \delta \).

\( D_i = \{ \vec{u} \mid \vec{u}_i = \pm \frac{1}{\sqrt{d}} \} \).

\( \Pr[ D_i(\vec{u}^2) \neq D_i(\vec{v}^2) ] = \frac{1}{2} - \frac{1}{2} \delta \) for \( \vec{u}, \vec{v} \) on edge.
Our gadget for Max-Cut \( \to \) hypercube graph.

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

Every subset of vertices \( S \) corresponds to a function \( f : \{\pm 1\}^d \to \{\pm 1\} \).

- \( f(x) = +1 \) if \( x \in S \).
- \( f(x) = -1 \) if \( x \notin S \).

Value of cuts
\[
\Pr \left[ f(x) \neq f(y) \right] \quad \text{for } \{x, y\} \text{ edge distn}
\]

Need machinery to reason about such quantities.
**BASIC FOURIER ANALYSIS OF BOOLEAN FUNCTIONS**

\[ f : \mathbb{B}^n \rightarrow \mathbb{R} \]

Let \( f : \mathbb{B}^n \rightarrow \mathbb{R} \) be a Boolean function.

Fourier analysis ~ write \( f \) as a linear combination of some nice functions, use it to prove properties of \( f \).

---

**INNER PRODUCT**

Let \( f : \mathbb{B}^n \rightarrow \mathbb{R} \)

\( g : \mathbb{B}^n \rightarrow \mathbb{R} \)

\[ \langle f, g \rangle = \mathbb{E} f(x) \cdot g(x) \]

\[ x \in \{-1,1\}^n = \mathbb{E}_x f(x) \cdot g(x) \]
“treat \( f, g \) as vectors of length \( 2^n \), take inner products, rescale by \( 2^n \). Note: \( \sum_i x_i = 0 \iff f \)

\[
\text{NORM} \\
\|f\|_2^2 = \langle f, f \rangle = \sum_i f(x_i)^2
\]

\[\text{Def (Parity functions/Monomials)}\]

For any \( S \subseteq [n] \),

\[X_S = \prod_{i \in S} x_i \] 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]

\[ \mathbb{E}_x X_s = 0 \quad \text{if } S \neq \emptyset \]

If \( S \neq T \),
\[ \langle X_s, X_T \rangle = 0. \]

If \( S = T \),
\[ \langle X_s, X_s \rangle = \|X_s\|_2^2 = 1 \]

Proof:

\[ X_s \cdot X_T = \prod_{i \in S} x_i \cdot \prod_{i \in T} x_i \]

\[ = \prod_{i \in S} x_i^2 \cdot \prod_{i \in T} x_i \]

\[ = \prod_{i \in S \Delta T} x_i \]

If \( S \Delta T \neq \emptyset \),
\[ \mathbb{E}_x \prod_{i \in S \Delta T} x_i = \mathbb{E}_x \prod_{i \in S \Delta T} x_i \]

\[ = 0 \]
**Observation:**
The set of functions \( \{ X_s \mid s \leq [n] \} \) form an orthonormal basis w.r.t. the inner product above.

**Proof:** We proved that each \( X_s \) has length: \( \|X_s\|^2_2 = 1 \) & \( \forall s \neq t \langle X_s, X_t \rangle = 0 \) & there are \( 2^n \) such functions & the demn of space \( \leq 2^n \).

\[ \square \]

Thus each \( f \) can be expanded as demn comb. of \( X_s \).
Definition (Fourier Transform)

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

\[ \hat{f}(s) = \sum_{s \in \{n\}} \hat{f}(s) \cdot x_s(x) \]

"writing for the basis of \(x_s\)."

Observation

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

\( \forall s, \ \hat{f}(s) = \mathbb{E} f(x) \cdot x_s \)

\( \downarrow \)

Correlation of \( f \) & \( x_s \).
Proof:

\[ f(x) = \sum_{S} \hat{f}(S) \cdot x_{S}(x) \]

\[ \mathbb{E} f(x) \cdot x_{T}(x) = \mathbb{E} \sum_{S} \hat{f}(S) \cdot x_{S}(x) \cdot x_{T}(x) \]

\[ = \sum_{S} \hat{f}(S) \cdot \mathbb{E} x_{S}(x) \cdot x_{T}(x) \]

\[ = \hat{f}(T) \]

Observation: \( f : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R} \)

Then, \[ \mathbb{E} f(x)^{2} = \sum_{S} \hat{f}(S)^{2} \]
Proof: \( f(x) = \sum_{s} \hat{f}(s) \cdot x_s(x) \)

\[
\mathbb{E} f(x)^2 = \frac{1}{2^n} \sum_{s,t} \hat{f}(s) \hat{f}(t) \\
= \sum_{s,t} \hat{f}(s) \hat{f}(t) \\
\mathbb{E} x_s \cdot x_t \\
= \sum_{s} \hat{f}(s)^2.
\]
Influence of Functions

Inf_i(f) : influence of i-th variable on f.

Def (for Boolean valued functions)

\[ f : \{-1,1\}^n \rightarrow \{-1,1\} \]

\[ \text{Inf}_i(f) = \Pr_{x \in \{-1,1\}^n} \left[ f(x) \neq f(x^{(i)}) \right] \]

\[ = \mathbb{E}_{x \in \{-1,1\}^n} \frac{1}{4} (f(x) - f(x^{(i)}))^2 \]

"prob that at a random x, flipping i-th bit changes the value of f"
**Lemma**

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

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

**Proof**: $f(x) = \sum_{S} \hat{f}(S) \cdot x_S$

$$f(x^{(i)}) = \sum_{S} \hat{f}(S) \cdot x_S^{(i)}$$

$$= \sum_{S \ni i} \hat{f}(S) \cdot x_S - \sum_{S \ni i} \hat{f}(S) \cdot x_S$$

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

or $\frac{1}{2}(f(x) - f(x^{(i)})) = \sum_{S \ni i} \hat{f}(S) \cdot x_S^{(i)}$
Parseval: \[
\int g(x)^2 = \sum_{s\in\mathbb{I}} \hat{f}(s)^2
\]
\[
\frac{1}{4} \int \left( f(x) - f(x^{(i)}) \right)^2
\]

Examples:

1) Dictator Function

\[ f(x) = x^j \rightarrow \text{depends only on } j\text{th bit.} \]

\[ \text{Inf}_j(f) = 1! \quad \text{Very influential} \]

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

\[ \frac{1}{4} \int \left( f(x) - f(x^{(i)}) \right)^2 = \frac{1}{n^2} \quad \text{"low influence function"} \]
3) $f(x) = \text{MAJ}(x), \; n: \text{odd}$

$$= \text{Sign} (\sum_i x_i).$$

When does flipping a bit change the value of $\text{MAJ}$?

Ans: when $\sum_i x_i = 1$

or $\sum_i x_i = -1$.

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

$\text{Prob} [ \sum_i x_i = 1 \text{ or } -1 ] = \binom{n}{\lfloor n/2 \rfloor} + \binom{n}{n/2} \frac{1}{2^n}$

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

4) **Parity Function**
\[ f(x) = \prod_{i=1}^{n} x_i \]

flipping any bit at any \( x \) changes the output of \textsc{parity function}.

Every bit has influence 1.

\[ \text{Inf}_i(f) = 1 + i \]

**LOW DEGREE INFLUENCE**

Def

\[ \text{Inf}_i^c(f) = \sum_{S \in i} \hat{f}(S)^2 \]

"discounting the effect of higher degree parity functions."
Obs: only $C$ bits can be influential now.

\[ \text{Inf}_{i}(x_5) = 0 \text{ if } |S| > C \]

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

\[ \text{Lemma (Only a small \# vars can be influential)} \]

Let $f: \{-1,1\}^n \to [-1,1]$. Then, for any $C > 0$

\[ |\{i : |\text{Inf}_{i}^C(f) \geq 3\}| \leq \frac{C}{3} \]
Proof
\[
\sum_{i=1}^{n} \text{Inf}_{i}(f) 
\leq \frac{c}{3}
\]
\[
= \sum_{i=1}^{n} \sum_{s \in S_{i}, |s| = 1} \hat{f}(s)^{2} 
\leq c \sum_{s} \hat{f}(s)^{2} \leq c
\]
\[
\text{since } \sum_{s} \hat{f}(s)^{2} = \frac{\sum_{x} E(f(x)^{2})}{c} \leq 1
\]
So, by Markov's inequality,
\[
\text{frac of } i : \text{Inf}_{i}(f) \geq \varepsilon
\leq \frac{c}{3}
\]