CS 294

CSPs

V: set of variables

D: constant size domain → w.r.t. size of the input

Local Constraints (constant # of variables / constraint)

Questions in CSP

Satisfiability: For an input I, ∃ a satisfying assignment?

↳ Dichotomy Theorem [91~18]

Approximability: find an assignment satisfying max # constraints

↳ PCP Theorem [91~92] → Unique Games Conjecture

↳ Satisfiable Instances, Promise CSPs

Counting: find the # of satisfying solutions

Random CSP: Average case analysis

↳ Structural Questions: 1) $Pr[I \text{ is satisfiable}] = 1$ or 0?

↳ Solution Space: 1) # of solutions 2) Geometry of solutions

↳ Algorithmic Questions: How to solve a random input algorithmically?

Quantum CSPs: Local Hamiltonians

Meta-Questions: 2SAT $\in$ P, but 3SAT $\in$ NPComplete. Why?

↳ This is usually hard, but CSP has a good expainable theory!

$\Rightarrow$ Given a description of a CSP Γ, is $\Gamma \in P$ or $\Gamma \in NPC$?

$\Rightarrow$ What is the best approximation ratio for Γ?

Satisfiability

Def) CSP: A CSP Γ is specified by:

1) Constant sized domain D

2) Finite set of relations $\{R_1, \dots, R_k\}$ over the domain

ex) $D = \{0,1\}$. $R_1 = \{(0,0),(0,1),(1,0)\}$. $R_2 = \{(0,0),(1,1),(1,0)\}$.

$R_3 = \{(0,0),(0,1),(1,1)\}$. $R_4 = \{(0,1),(1,0),(1,1)\}$

An instance of CSP Γ is: Variables $V = \{v_1, \dots, v_n\}$. Constraints $C_1 \dots C_m$, where each $C_i \equiv$ some R_i applied to a subset of V.

ex) Γ has a relation R which is binary & symmetric. $R \subseteq D^2$

↳ R is a graph H: Instance of $\Gamma \Leftrightarrow$ Graph G on n vertices

CSP(Γ_H): $\exists$ a map $f: G \to H$ s.t. $\forall (u,v) \in G$, $(f(u), f(v)) \in H$.

$H = \triangle$ → this corresponds to 3-Coloring! What about

$H =$  (in P) or $H =$  (in NPC)?

[Hell-Nesteril]: $\Gamma_H \in P$ if H is bipartite, $\in NPC$ otherwise.

Schaefer's Theorem: Among all Boolean CSPs, following are in P:
1) 2-SAT 2) Horn-SAT 3) Dual Horn SAT 4) Lin Eq. 5) Trivial CSP
And the rest are in NPC. ↳ all variables are negated

[Feder-Vardy]: Every CSP is in P or NP-Complete. (Conjecture)

Polymorphisms

ex) Linear Equations: $[A][x] = [b]$ over $\mathbb{Z}_2$.

Suppose we have 3 solutions to the linear systems, $Ax = Ay = Az = b$.

Consider $A(x+y-z) = b$. $(x+y-z)$ is also a solution.

Polymorphisms combine solutions of a CSP to produce another solution.

Concretely, $h: \mathbb{Z}_2^3 \to \mathbb{Z}_2$, $h(a,b,c) = a-b+c \pmod 2$ is a polymorphism for LINEQ since $\forall Ax=b$, $\forall x,y,z$ solutions of the following produces a new solution $w = h(x,y,z)$.

$$X = [x^{(1)}, x^{(2)}, \cdots, x^{(n)}]$$
$$y = [y^{(1)}, y^{(2)}, \cdots, y^{(n)}]$$
$$Z = [z^{(1)}, z^{(2)}, \cdots, z^{(n)}]$$

$$W = X+y-Z = [x^{(1)}+y^{(1)}-z^{(1)}, \cdots, x^{(n)}+y^{(n)}+z^{(n)}]$$

Some polymorphisms: 2-SAT → Maj(a,b,c), HORN-SAT → AND(a,b),
Dual HORN-SAT → OR(a,b), LINEQ → a+b-c, Trivial CSP → any h.

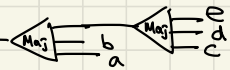
ex) $x \xrightarrow{x \Rightarrow y} y \xrightarrow{\bar{y} \Rightarrow z} z \xrightarrow{\bar{z} \Rightarrow \bar{w}} w$, $y \xrightarrow{\bar{y} \Rightarrow \bar{w}} w$ (For 2-SAT)

x	0	1	1	0
y	1	1	1	0
z	0	0	1	0
Maj(x,y,z)	0	1	1	0

Every CSP has a trivial polymorphism: $h(a,b,c) := b$ (dictator function)

"Algebraic" Dichotomy: CSP $\Gamma \in P$ iff it has a nontrivial polymorphism. (otherwise $\Gamma \in$ NPC)

↳ We can recursively combine h for more polymorphisms! (universal algebra*)

ex) $Maj(a,b,c) \to Maj(a,b,Maj(c,d,e))$ 

[Hardness]: no nontrivial polymorphism if $\Gamma \in$ NPC. (03~04)

[Algorithm]: nontrivial polymorphism of Γ implies $\Gamma \in P$. (18~19)

Theorem) CSP Γ_1 reduces to CSP Γ_2 iff $Poly(\Gamma_2) \subseteq Poly(\Gamma_1)$.

Polymorphisms as Solution to CSP (WLOG, 3-SAT): Consider $Poly^{(3)}$(3-SAT), a polymorphism on 3 bits of 3-SAT candidates.

0	0	0	0	1	1	1	→ $h_1(x) = x_1$
0	0	1	1	0	0	1	→ $h_2(x) = x_2$
0	1	0	1	0	1	0	→ $h_3(x) = x_3$

$h(0,0,0)$ $h(0,0,1)$ - - - - - - $h(1,1,1)$ where $h(a,b,c): \{0,1\}^3 \to \{0,1\}$

X_{000} X_{001} - - - - - - X_{111}

$DICT^{(3)}_{3SAT} := (\text{Variables} := \{X_{000}, \ldots, X_{111}\}$, → 8 variables

Constraints := {every constraint satisfied by the dictator function}).

↳ any assignment on $DICT^{(3)}_{3SAT}$ is a function on 3 bits.

Claim: any satisfying assignment for $DICT^{(3)}_{3SAT}$ is a polymorphism for 3SAT.

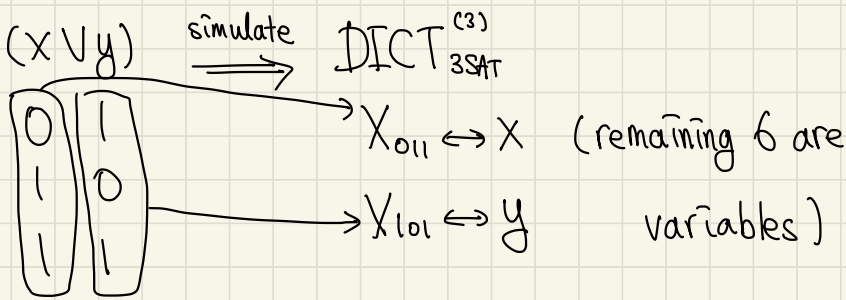
Generally, $\forall \Gamma, \forall k$, a satisfying assignment of $DICT^{(k)}_{\Gamma}$ is a k-bit poly. of Γ.

Def) Dictatorship Gadget: $DICT^{(k)}_{\Gamma}$ for all Γ and arity k.

$DICT^{(3)}_{3SAT}$:= instance of 3SAT on 8 variables.

↳ 3SAT has no nontrivial poly. ⇔ the only satisfying assignment to $DICT^{(3)}_{3SAT}$ are $h_1(X_1,X_2,X_3)=X_1, \ldots, h_3(X_1,X_2,X_3)=X_3$ (dictatorships)

If we were to reduce 2SAT → 3SAT:



Generally, if Γ has k satisfying conditions, reduce it to $DICT^{(k)}_{\Gamma}$.

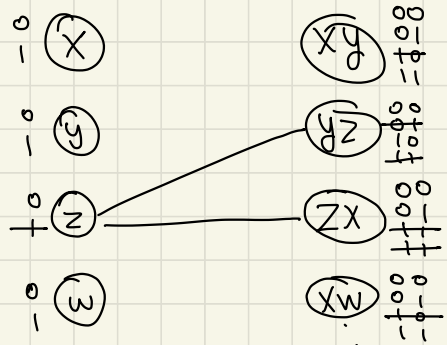
Then, $DICT^{(k)}_{\Gamma}$ proves hardness iff we cannot "add more rows" to it.

Algorithms for CSP: Suppose $\text{Poly}(\Gamma)$ is nontrivial. Then $\Gamma \in P$.

↳ Two big ideas: 1) Local Propagation 2) Gaussian Elimination over $\mathbb{F}_p$.

Local Propagation: Consider a 2SAT instance, $\bar{z} \wedge (\bar{x} \vee w) \wedge (\bar{x} \wedge y)$.

Look at local states and eliminate infeasible options. After it stops,



there are sets of plausible local states.

A width k-local prop. $\to \forall$ k-tuples $S \subseteq V$, $\exists$ a set of partial assignments $A_S \subseteq \{0,1\}^{|S|}$.

↳ partial assignments A_S are consistent $\Rightarrow T \subseteq S$, A_T is consistent with A_S if every assignment $\sigma_T \in A_T$ can be extended to an assignment in A_S.

Def) A CSP Γ is width k if a width k-local prop. decides Γ.

* LINEQ is not bounded-width! (Intuitively, local conditions do not imply anything about the system as a whole)

* Local prop. can be simulated by LP, i.e. LPs that execute local prop.

↳ Variables: $P_{S,\alpha}$ for $S \subseteq V$, $\alpha \in \{0,1\}^{|S|}$ where $P_{S\alpha} := \mathbb{1}\{S \text{ is in state } \alpha\}$

Constraints: (S is assigned some α) $\sum_{\alpha} P_{S,\alpha} = 1$.

(extension consistency of S) $P_{S\cup\{i\},\alpha\circ 0} + P_{S\cup\{i\},\alpha\circ 1} = P_{S,\alpha}$

(violation of S) $P_{S,\alpha} = 0$ if α violates a constraint of CSP for S.

* Gaussian Elimination is not captured by LP/Convex Optimization!

↳ However, a general algorithm for CSP must involve both LP & GE!

Randomized Algorithm for LINEQ on $\mathbb{F}_2$: start with $O(n^2)$ random inputs. Delete any candidates inconsistent with the first constraint's parity. Then, produce $O(n^2/2)$ inputs via polymorphism. Repeat over constraints.

Probabilistic Checkable Proofs

Def) Proof System: A verifier (Claim, Proof) $\to \{0,1\}^{\text{accept/reject}} \in P$

Completeness: True claims have proof (if Claim=True, $\exists$ proof s.t $V[c,p]=1$)

Soundness: False claims have no proof (if claim=False, $\forall$ proof $V[c,p]=0$)

↳ If we make the verifer probabilistic, we constrain on $Pr[V[c,p]=0/1] \geq 2/3$.

PCP: verifier is probabilistic and it reads only $\theta(1)$ bits of proof!

Def) PCP(q,L): q := # of queries into proof, L := length of proof

Naive PCP for 3SAT: Claim: ϕ is satisfiable, Proof: $x \in \{0,1\}^n$.

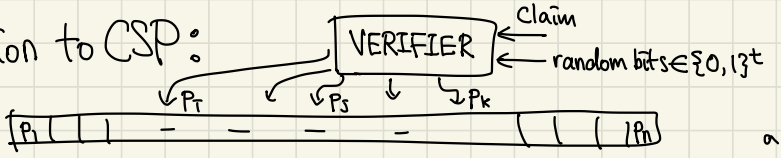
↳ Verifier: Picks a random clause C from ϕ, checks if C is satisfied.

⇒ Completeness is evident, 1.

⇒ Soundness: ϕ is not satisfiable, then $\forall x, \exists c$ s.t. x violates C.

↳ $\Pr[\text{verifier rejects } x] \geq \frac{1}{\#\text{clauses}} = O\left(\frac{1}{\text{poly}(n)}\right)$. //

Connection to CSP:



For a single execution of the PCP, the verifier expects a constant # of bits $\{P_1 \ldots P_k\}$ to satisfy a certain constraint on those bits.

The set of all possible executions gives a set of constraints on local bits. ⇒ CSP instance I_V.

$$\max_{\text{proof } p}\{\Pr[V[c,p]]\} \Leftrightarrow \max_{p \sim \text{assignment of proof bits}}\{\text{val}(I_V, p)\}$$

(val: fraction of constraints of I_V satisfied by p)

Theorem) PCP Theorem [9!]: $NP = PCP(\Theta(1), \text{poly}(n))$

Relevance: Proofs ("Computation") are brittle objects, i.e. one false claim in a proof can break the entire proof. Then, it's surprising that we can check a proof's validity with only constant # of samples!

Exponential Sized PCP for NP

Claim: $NP \subseteq PCP(30, 2^{n^2})$.

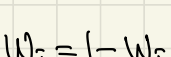
The NPC problem: QUAD-EQ.

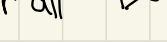
Variables: $X_1, \dots, X_n \in \{0,1\} = \mathbb{F}_2$. ex) $X_1X_2 - X_3 = 0$

Constraints: quadratic equations each on 3 variables.

Goal: Is there a satisfying assignment (decision problem)

Proof of QUAD-EQ$\in$NP: CIRCUIT-SAT $\leq_p$ QUAD-EQ.

Consider a CIRCUIT-SAT with NOT and AND gates. Variables will be all wires in the circuit, $w_1, \dots, w_n$. For all  , $w_iw_j = w_k$.

For all  , $w_i = 1 - w_j$. The output $w_n = 1$. Reduction is complete. //

Hadamard Codes (Truth table of linear function)

↳ Encode$(l_1, \dots, l_n) \in \mathbb{F}_2^n \Rightarrow$ Table of linear function $L(x) = \langle l, x \rangle = \sum_i l_i x_i$ $(\mathbb{F}_2^n \to \mathbb{F}_2)$

$\Rightarrow$ In effect, we encode n bits into 2^n bits $(0, l_1, l_2, l_3, l_1+l_2, \dots)$

Self-Correction: Suppose $\tilde{L}: \mathbb{F}_2^n \to \mathbb{F}_2$ that is ε-close to a linear function (Hadamard code) L. $\Pr[x \in \mathbb{F}_2^n, L(x) = \tilde{L}(x)] \geq 1 - \varepsilon$.

$\forall x \in \mathbb{F}_2^n$, we want $L(x)$ from $\tilde{L}(x)$. Consider $\tilde{L}(x+y) - \tilde{L}(y)$.

* A true linear function L would yield $L(x+y)-L(y)=L(x)$.

Claim: $\Pr[\tilde{L}(x+y)-\tilde{L}(y)=L(x)]\geq 1-2\varepsilon$ for a random $y\sim \mathbb{F}_2^n$.

Proof: Fix x. When y is random, $x+y$ is uniformly random over $\mathbb{F}_2^n$.

↳ $\Pr[\underbrace{\tilde{L}(x+y)=L(x+y)}_{e_1}]\geq 1-\varepsilon$, $\Pr[\underbrace{\tilde{L}(y)=L(y)}_{e_2}]\geq 1-\varepsilon$.

⇒ By union bound, $\Pr[e_1\cap e_2]\geq 1-2\varepsilon$. //

Theorem) Linearity Testing*(later): Suppose $\tilde{L}:\mathbb{F}_2^n\to\mathbb{F}_2$ satisfies $\Pr_{x,y}[\tilde{L}(x+y)=\tilde{L}(x)+\tilde{L}(y)]\geq 1-\varepsilon$, then $\exists$ a linear function L s.t. $\Pr_x[\tilde{L}(x)=L(x)]\geq 1-\varepsilon$.

Intuition: To test linearity, sample $x,y\sim\mathbb{F}_2^n$. Test whether $\tilde{L}(x)+\tilde{L}(y)=\tilde{L}(x+y)$ (Linearity test is sound)

A Testing Problem: $\vec{l}=(l_1,\dots,l_n)\in\mathbb{F}_2^n$. Test if $\vec{l}=\vec{0}$?

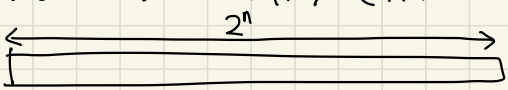
Suppose we have access to $L(x)=\langle l,x\rangle$. Then, check $L(x)=0$ for a random $x\in\mathbb{F}_2^n$.

Claim: Suppose $\vec{l}\neq\vec{0}$. $\Pr_x[L(x)=\langle l,x\rangle=0]\geq 1/2$. (Exercise)

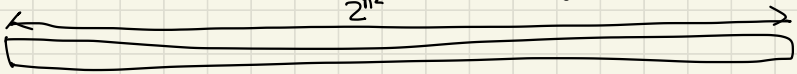
Construction of PCP for QUAD-EQ: Variables $l_1\dots l_n\in\mathbb{F}_2$.

Quadratic equations $q_i(l)=0$ for $i=1\dots m$. Suppose l satisfies.

PCP Proof: $L(x) := \langle l, x \rangle$ (Hadamard encoding of l)

↳ $L(x)$: 

$H(x): \mathbb{F}_2^{n^2} \to \mathbb{F}_2$, $H(x) :=$ Hadamard encoding of $l^{\otimes 2} = \{l_i l_j \mid i,j \in [n]\}$ ($\to |l^{\otimes 2}| = n^2$)

↳ $H(x)$: 

↳ $H((C_{ij})) = \sum_{i,j} C_{ij} l_i l_j = \langle C, l^{\otimes 2} \rangle$

Verifier: Suppose the prover gives $\tilde{L}(x)$ and $\tilde{H}(x)$. First, test $\tilde{L}$ is ε-close to a linear function using 3 bits in linear time. Next test $\tilde{H}$ is ε-close to a linear function $\hat{H}(c) = \sum_{ij} M_{ij} C_{ij}$ (3 bits). If these are not ε-close, we will reject with probability ε. Now, we can pretend we have access to L and $\hat{H}$ via self-correction. We test if $M_{ij} = l_i l_j \; \forall i,j$ (linear to quadratic consistency). If so, we are convinced we have access to true L and H. The goal is to test whether $q_i(l) = 0$ for all $q_1, \dots q_m$. Pick $r_1 \dots r_m \in \mathbb{F}_2^m$. Let $R(l) := \sum_i r_i q_i(l)$. Test if $R(l) = \sum r_{ij} l_i l_j + \sum s_i l_i + b = 0$. This is easy since we can rewrite $R(l) = H(r_{ij}) + L(s) + b$. We recover $H(r_{ij})$ and $L(s)$ each with 2 bits of $\hat{H}$ and $\hat{L}$. Then, R gives a possible violation of l based on $r_1 \dots r_m$. //

Linearity Testing

Input: $f: \mathbb{F}_2^n \to \mathbb{F}_2$, want to test if f is close to linear.

→ Pick $x, y \in \mathbb{F}_2^n$ u.a.r. Check $f(x+y) = f(x) + f(y)$. [BLR]
Blum-Luby-Rubinfield

Thm) $\forall$ function f, $\exists$ a linear function g s.t. $\text{dist}(f,g) = \Pr_n[f(n) \neq g(n)]$
$\leq \frac{1}{2} \underbrace{\Pr[\text{BLR rejects } f]}_{\hookrightarrow \delta}$. ($\frac{1}{2}$ can be removed if done with Fourier)

$\mathbb{F}_2$ Facts:

1) $+1 = -1 \pmod 2 \to x+y = x-y$, $x+y+y = x$.

2) fix $x \in \mathbb{F}_2^n$. $y \sim \mathbb{F}_2^n$ u.a.r. $\Rightarrow x+y \sim$ u.a.r. $\mathbb{F}_2^n$.

3) fix $v \neq 0$, $\Pr_x[\langle v, x \rangle \neq 0] = \frac{1}{2}$.

Reminder) $f(x) = f(x+z) - f(z)$ if f is linear.

Define $g(x) := \underset{z \in \mathbb{F}_2^n}{\text{Maj}}(f(x+z) - f(z))$.

Claim A: g is a linear function if $\delta < \frac{1}{20}$. ($\delta \to \Pr[\text{BLR rejects } f]$)

Claim B: $\text{dist}(f,g) = \Pr_x[f(x) \neq g(x)] \leq 2\delta$.

Let $P_x := \Pr_y[g(x) = f(x+y) - f(y)]$ $\forall x$.

"Surprising" Claim: $\forall x,\ P_x \geq 1-2\delta$.

Proof: Consider $\Pr_{y,z}[\overbrace{f(x+y)-f(y) = f(x+z)-f(z)}^{A}]$ for a fixed x. assume $g(x)=0$.

↳ $f(x+y)-f(y)$ is a bit with bias P_x. $\Rightarrow \Pr[A] = P_xP_x + (1-P_x)(1-P_x)$
$= P_x^2 + (1-P_x)^2$.

Now, consider event $A: f(x+y)-f(y) = f(x+z)-f(z)$

$\Leftrightarrow f(x+y)+f(x+z) = f(y)+f(z)$ (we can ignore $\pm$)

$\rightarrow \Pr[A] = \Pr_{y,z}[f(x+y)+f(x+z) = f(y)+f(z)]$.

↳ $f(y)+f(z) = f(y+z)$ with probability $(1-\delta)$.

↳ $f(x+y)+f(x+z) = f(x+y+x+z) = f(y+z)$ with prob. $(1-\delta)$.

$\Rightarrow \Pr[A] \geq 1-2\delta$. $\Rightarrow P_x^2 + (1-P_x)^2 = \Pr[A] \geq 1-2\delta$.

$\Rightarrow \delta \geq P_x(1-P_x) \geq \frac{1}{2}(1-P_x) \Rightarrow P_x \geq 1-2\delta$. //
($P_x \geq 1/2$)

Proof of Claim A: Need to prove $g(x)+g(y) = g(x+y)\ \forall x,y$.

$g(x) \ominus f(x+z) - f(z)$ with prob. $1-2\delta$. (due to surprising claim)

$$
\begin{array}{rl}
 & g(x) \ominus f(x+z) - f(z) \\
 & \quad\quad + \quad\quad\quad + \\
+ & g(y) \ominus f(y+w) - f(w) \quad // \\
\hline
 & g(x+y) \ominus f(x+y+z+w) - f(z+w) \quad //
\end{array}
$$

(the two sums marked ⑪; w.p. $1-\delta$, w.p. $1-\delta$)

$\rightarrow$ w.p. $1-8\delta$ over z & w, all $\ominus$ happens $\Rightarrow \exists z,w$ s.t. this happens. → one instance is enough. (witness)

↳ $1-8\delta > 0 \rightarrow \delta < \frac{1}{8}$

$\Rightarrow g(x+y) = g(x)+g(y)$.

Proof of Claim B: $\Pr_{x,y}[f(x)+f(y)\neq f(x+y)]=\delta$ (def. of BLR)

$\Leftrightarrow \Pr_{x,y}[f(x)\neq f(x+y)-f(y)]=\delta$.

Let $Bad := \{x \mid f(x)\neq g(x)\}$. $\Pr_y[f(x)\neq f(x+y)-f(y) \mid x\in Bad] \geq \frac{1}{2}$. ($Maj_y(f(x+y)-f(y))$)

$\Pr[f(x)+f(y)\neq f(x+y)] \geq \Pr[x\in Bad]\cdot$ [ ] $\geq \frac{1}{2}\cdot \Pr[x\in Bad]$

$\Rightarrow \Pr[x\in Bad] \leq 2\cdot\Pr[f(x)+f(y)\neq f(x+y)] = 2\delta$. //

Exp. PCP with Linearity Testing: When testing $\tilde{L}$ and $\tilde{H}$, use BLR to test if they are ε-close to a linear function. Each rejects with prob. $\Omega(\varepsilon)$, and if not, $\exists$ linear functions L and H s.t. $dist(\tilde{L},L)\leq\varepsilon$ and $dist(\tilde{H},H)\leq\varepsilon$. Then, pick $x,y\in\mathbb{F}_2^n$ u.a.r.

Test $(\sum l_i x_i)(\sum l_i y_i) = L(x)\cdot L(y) = H(xy^T) = \sum H_{ij}x_i y_i = \sum l_i l_j x_i y_j$.

More $\mathbb{F}_2$ Facts:

$\Pr_{x,y\in\mathbb{F}_2^n}[x^T M y = \sum M_{ij}x_i y_j \neq 0] \geq 1/4 \quad \forall M\neq 0$.

Proof: $M\neq 0 \Rightarrow \exists$ a row $M_i \neq 0 \Rightarrow (My)|_i = \langle M_i, y\rangle$.

$\forall M_i\neq 0$, $\Pr_y[\langle M_i,y\rangle\neq 0]=1/2 \Rightarrow$ w.p. $\geq 1/2$, $My\neq 0$.

$\rightarrow \Pr_x[\langle x, My\rangle \neq 0] \geq \Pr[My\neq 0]\cdot\Pr[\langle x,My\rangle\neq 0 \mid My\neq 0] \geq \frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$. //

$\Rightarrow H(xy^T)-L(x)\cdot L(y) = x^T H y - x^T ll^T y = x^T(H-ll^T)y$. If $H-ll^T\neq 0$, by the above fact, $\Pr[x^T(H-ll^T)y\neq 0]\geq 1/4$. $\Rightarrow$ can detect $H\neq ll^T$ w.p. $\geq 1/4$.

Now, take a random linear combination of constraints $\{q_i\}$. This can just be written as some $\sum A_{ij} l_i l_j + \sum b_i l_i + C = H(A_{ij}) + L(b) + C = 0$.

If any of $q_1(l), \ldots, q_k(l)$ are not $\emptyset$, w.p. $\frac{1}{2}$, $\Pr_l[\sum_i r_i q_i(l) \neq 0] \geq \frac{1}{2}$. //

* Soundness Boosting for Lin. Test: say we sample $x, y \in \mathbb{F}_2^n$ t times. If f is ε-far from linear, Prob[success] $\leq (1-\varepsilon)^t$ with $3t$ bits of reading. For t bits, best possible soundness is $\frac{1}{2^t}$.

* PCP that reads t bits accept wrong proof w.p. $\frac{2t}{2^t}$.

Suppose S: $NP \subseteq PCP(t, \text{poly}(n), \varepsilon) \Leftrightarrow \exists$ a CSP with constraints on t bits s.t. it is NP-Hard to approximate better than ε-factor.

For $\varepsilon = \frac{2t}{2^t}$, if $S(\varepsilon)$ is true, $\forall$ CSPs with constraints on t bits, $\exists$ a $\frac{2t}{2^t}$-factor approximation algorithm.

Expander Graphs

Most connected graph: K_n, $\binom{n}{2}$ edges. To break K_n, we need to delete a constant fraction of edges.

Def) Edge Expansion: For $G(V,E)$, $\forall S \subseteq V$ s.t. $|S| \leq n/2$, $\phi(S) = \frac{E(S,\bar{S})}{d|S|}$.

(ratio of edges crossing $S \to \bar{S}$; # of edges incident in S)

↳ Also, $\phi(S) = \Pr_{u \in N(v)}[u \notin S]$ where $N(v)$ is neighbors of v.

Def) Conductance: $\phi(G) = \min_{|S| \le n/2} \{\phi(S)\}$.

Def) Expander (1): a d-regular (degree $= d$) graph $G(V,E)$ is a (α, d)-expander if $\phi(G) \ge \alpha = \Omega(1)$. (combinatorial)

Theorem) [LPS] $\exists$ constant degree expander graphs with constant expansion.

↳ Pick a prime p. Vertices are $\{0, 1, \dots, (p-1), \infty\}$. Edges are:

$x \to x+1 \pmod p$, $x \to x-1 \pmod p$, $x \to 1/x \pmod p$.

This is a degree-3 graph and is an expander. //

Theorem) A random d-regular graph is an expander w.h.p.

* Notation: Adjacency Matrix $A_{ij} := \mathbb{1}\{(i,j) \in E\}$.

Fact) Every real symmetric matrix A has:

1) n orthonormal eigenvectors $\vec{v}_1, \dots, \vec{v}_n$ with eigenvalue $\lambda_1, \dots, \lambda_n$.

↳ $A = \sum_i \lambda_i v_i v_i^T$. for $\lambda_1 > \lambda_2 > \dots > \lambda_n$, $\lambda_1 = \max_{x \in \mathbb{R}^n} \left\{\frac{x^T A x}{\|x\|^2}\right\}$, $\lambda_2 = \max_{x \perp v_1} \{\}$.

Def) Expander (2): A d-regular graph is a (d, λ) expander if $\max\{\lambda_2, |\lambda_n|\} \le \lambda$. (spectral)

$\forall d$-regular graph, $\lambda_1 = d$, $\vec{v}_1 = \vec{1}$, and all other λs live in range $[-d, d]$.

Theorem) In a d-regular graph, $\underline{\phi(G)^2 \le 1 - \lambda/d \le 2\phi(G)}$ where

$\lambda := \max\{\lambda_2, |\lambda_n|\}$ ↳ Cheeger's Inequality

$A = d \cdot \vec{1}\cdot\vec{1}^T + \overbrace{\sum_{i>1} \lambda_i \vec{v}_i\vec{v}_i^T}^{M}$ where $\vec{1} := \left[\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right]$.

For a complete graph, $A_c = \vec{1}\cdot\vec{1}$. So, if M is "negligible", A feels like a complete graph! → $\|M\| = \max\{\lambda_2, |\lambda_n|\} = \lambda$, $\|Mx\| < \lambda\|x\|$.

Claim) $\phi(G) \ge \frac{1}{2}\left(1 - \frac{\max\{\lambda_2, |\lambda_n|\}}{d}\right)$.

Proof: Fix a set S. Let $\vec{1}_S = [1\{v_1 \in S\}, \dots, 1\{v_n \in S\}]$.

$\vec{1}_S^T A \vec{1}_S = \sum_{i,j} A_{ij} \cdot 1_S(i) \cdot 1_S(j) = \sum_{i,j \in S} A_{ij} = 2E(S,S)$.

Then, $E(S,\bar{S}) = d|S| - E(S,S) = d|S| - \vec{1}_S^T A \vec{1}_S$.

→ $\phi(S) = 1 - \frac{\vec{1}_S^T A \vec{1}_S}{d|S|}$. We define a new graph L as such:

Def) Laplacian Matrix: $L := d(Id) - A$.

$\vec{x}^T A \vec{x} = 2\sum_{(i,j)\in E} x_i x_j$. $\vec{x}^T L \vec{x} = \sum_{(i,j)\in E} (x_i - x_j)^2 = d\sum_i x_i^2 - 2\sum_{(i,j)\in E} x_i x_j$.

$\lambda_1(L) = d - \lambda_1(A)$, $\lambda_2(L) = d - \lambda_2(A)$, ...

$\phi(S) = \frac{E(S,\bar{S})}{|S|}$.

Claim: $\underline{E(S,\bar{S}) = \vec{1}_S^T L \vec{1}_S}$. $\sum_{(i,j)\in E} \underbrace{(1_S(i) - 1_S(j))^2}_{1\{(i,j) \text{ crosses } S \to \bar{S}\}} = \#$ of crossing edges.

$\to \vec{1}_S = \underbrace{\frac{|S|}{\sqrt{n}}}_{\alpha} \cdot \vec{1} + \vec{v}$ where $\vec{v} \perp \vec{1}$. ($\frac{|S|}{\sqrt{n}}$ is $\langle \vec{1}, \vec{1}_S \rangle$)

$\to \vec{u}^T L \vec{u} = \sum_i \lambda_i(L) \cdot \langle \vec{v}_i, \vec{u} \rangle^2 \to \vec{1}_S^T L \vec{1}_S = \lambda_1(L) \cdot \langle \vec{1}_S, \vec{1} \rangle^2 + \sum_{i>1} \overbrace{\lambda_i(L)}^{\geq (d-\lambda)} \langle \vec{1}_S, \vec{v}_i \rangle^2$

$\geq (d-\lambda) \cdot \sum_{i>1} \langle \vec{1}_S, \vec{v}_i \rangle^2$. $\left(\|\vec{1}_S\| = \sqrt{S},\ \|\vec{v}\|^2 = |S| - \frac{|S|^2}{n} = |S|(1 - \frac{|S|}{n}). \right)$

$\geq (d-\lambda) \cdot |S| \cdot (1 - \underbrace{\frac{|S|}{n}}_{\geq 1/2}) \to \phi(S) = \frac{E(S,\bar{S})}{d|S|} = (1 - \frac{\lambda}{d})(1 - \frac{|S|}{n}) \geq \frac{1}{2}(1 - \frac{\lambda}{d})$. //

Def) Expander (3): Graphs on which random walks "mix" in $\Theta(\log n)$ steps. $\pi_u^t :=$ distribution of endpoint v after t steps. i.e. in $t = \Theta(\log n)$ steps, $\| \pi_u^t - \text{Uniform} \| \leq \frac{1}{n^2}$.

Fact) $\pi_u^{t+1} = \frac{1}{d} \cdot A \cdot \pi_u^t$. In particular, $\pi_u^t = \left(\frac{A}{d}\right)^t \pi_u^0$.

Fact) Expander Graphs behave like Complete Graphs.

↳ Expander Mixing Lemma: G is a (n, d, λ) expansion. $S, T \subseteq V$. ($S \cap T = \emptyset$?)

$\frac{E(S,T)}{|E|}$ is how much of the edges cross $S \to T$. For a complete graph K_n, this is $\frac{|S| \cdot |T|}{\binom{n}{2}}$. $\left| \frac{E(S,T)}{|E|} - \frac{|S| \cdot |T|}{\binom{n}{2}} \right| \leq \lambda \sqrt{\frac{|S| \cdot |T|}{n^2}}$

Application) Reduce the randomness needed for <u>error reduction</u>.

Setup: A is a random algo. $\Pr_x[A(x,r) \text{ is correct}] \geq 0.6$.

Run $A(x, r_1), \ldots, A(x, r_k)$ and return the majority vote.

Total randomness used is $k \cdot |r|$ bits. Success prob $= O(1 - e^{-\Omega(k)})$.

Idea: Let A use R bits of randomness. Fix an expander graph on vertices $\{0,1\}^R$. Let $r_1 :=$ new random point on $\{0,1\}^R$. $r_2 \dots r_k$ are generated by a random walk in the expander. Use $r_1 \dots r_k$ as the random bits for trials of A. We only need $R + k\log(d)$ randomness and $r_1 \dots r_k$ still satisfies randomness properties of A.

Lemma 1) Let $G(V,E)$ be a (n, d, λ)-expander ($|V| = n$, deg $= d$, $\lambda^{(A)} = \max\{\lambda_2, |\lambda_n|\}$). Let $F \subseteq E$ be a subset of the edges ("bad edges"). $\Pr[t\text{-th step of walk} \in F \mid 0\text{-th step} \in F] \le \frac{|F|}{|E|} + \left(\frac{\lambda}{d}\right)^{t-1}$.

Proof: Let $v :=$ starting point. Let $Y_v := |F \cap N(v)|/d$. (prop. of bad edges)

Let $X_v := \Pr[v \text{ is the first vertex} \mid 0\text{-th step} \in F] = \frac{|F \cap N(v)|}{|F|} \cdot \frac{1}{2} = \frac{d\,Y_v}{2|F|}$.

Then, $\bar{A}^{t-1} \cdot \vec{X} =$ prob. dist. of w (vertex after $t-1$ steps) where $\bar{A} = \left(\frac{A}{d}\right)$.

$P = \sum_w \left(\bar{A}^{t-1} \cdot \vec{X}\right)_w \cdot Y_w$ ($[\text{prob. that } (t-1)\text{-th vertex is } w] \cdot [t\text{-th step} \in F \mid w]$).

$\vec{X} = \vec{X}^{\parallel} + \vec{X}^{\perp}$ where $\vec{X}^{\parallel}$ is the component along $\vec{1}$ and $\vec{X}^{\perp}$ is orth. to $\vec{1}$.

$\vec{X}^{\parallel} = \langle \vec{X}, \vec{1}(\frac{1}{\sqrt{n}}) \rangle \cdot \left(\vec{1}\left(\frac{1}{\sqrt{n}}\right)\right) = \left(\sum \vec{X}^{\parallel}\right) \cdot \frac{1}{n} \cdot \vec{1} = \frac{1}{n}\vec{1}$. ($\bar{A}\vec{1} = \vec{1}$)

$\bar{A}^{t-1}\vec{X} = \bar{A}^{t-1}\vec{X}^{\parallel} + \bar{A}^{t-1}\vec{X}^{\perp} = \vec{X}^{\parallel} + \left(\frac{A}{d}\right)^{t-1}\vec{X}^{\perp}$.

$P = \langle \bar{A}^{t-1}\vec{X}, \vec{Y} \rangle = \frac{2|F|}{d}\langle \bar{A}^{t-1}\vec{X}, \vec{X} \rangle = \frac{2|F|}{d}\langle \vec{X}^{\parallel} + \left(\frac{A}{d}\right)^{t-1}\vec{X}^{\perp}, \vec{X}^{\parallel} + \vec{X}^{\perp} \rangle$

$= \frac{2|F|}{d}\left[\langle \vec{x}^{\parallel}, \vec{x}^{\parallel}\rangle + \langle \left(\frac{A}{d}\right)^{t-1}\vec{x}^{\perp}, \vec{x}^{\perp}\rangle\right]$ ($\vec{1}/\sqrt{n}$) (since $\vec{x}^{\parallel} \perp \vec{x}^{\perp}$, $\left(\frac{A}{d}\right)^{t-1}\vec{x}^{\parallel} = \vec{x}^{\parallel}$)

$\leq \frac{2|F|}{d}\left(\frac{1}{n} + \|\left(\frac{A}{d}\right)^{t-1}\vec{x}^{\perp}\| \cdot \|\vec{x}^{\perp}\|\right)$ (Cauchy-Schwartz)

$\leq \frac{2|F|}{d}\left(\frac{1}{n} + \left\{\left(\frac{\lambda}{d}\right)^{t-1} \cdot \|\vec{x}^{\perp}\|\right\} \cdot \|\vec{x}^{\perp}\|\right)$, and use the bound $\|\vec{x}^{\perp}\|^2 \leq \|\vec{x}\|^2$.

$* \|\vec{x}\|^2 = \sum_v x_v^2 \leq \overbrace{\sum_v x_v}^{1} \cdot (\max_w x_w) = \max_w(x_w) = \max_w\left(\frac{|F \cap N(w)|}{2|F|}\right) \leq \frac{d}{2|F|}$.

$\rightarrow \leq \frac{2|F|}{d}\left(\frac{1}{n} + \left(\frac{\lambda}{d}\right)^{t-1} \cdot \frac{d}{2|F|}\right) = \frac{|F|}{|E|} + \left(\frac{\lambda}{d}\right)^{t-1}$. //

Upshot: no matter how F is selected, we can "escape" F w.h.p.

Dinur's PCP Proof

Recall: PCP Theorem $\Rightarrow NP \subseteq PCP(q, \ell)$ (queries, length of proof)

Def) 2CSP: Given by $G(V, E)$ (V: variables, E: constraints), a constraint graph, a constant size Σ ($|\Sigma| = k$).

An assignment is $\sigma: V \to \Sigma$. Constraints are $\{C_e \subseteq \Sigma \times \Sigma \mid e \in E\}$ (C_e: set of allowed config of edge e).

σ satisfies edge $e = (u, v)$ if $(\sigma(u), \sigma(v)) \in C_e$. The value of σ,

$Val(\sigma)$ = fraction of constraints satisfied $= \frac{1}{|E|} \cdot$ (# of edges satisfied by σ).

$OPT(G) = \max_\sigma\{Val(\sigma)\}$. $gap(G) = 1 - OPT(G)$.

MAX-2CSP: Given $G(V, E)$, compute $OPT(G)$.

Proof of PCP Theorem: 3SAT(some NP-C problem) $\xrightarrow{(\phi) \to (G)}$ 2CSP preserving completeness (ϕ is satisfiable $\Rightarrow OPT(G) = 1$) and soundness (ϕ is

unsatisfiable $\Rightarrow OPT(G_1) < 0.9$, gap > 0.1).

Idea: Pick 3-Coloring for ϕ, which is already a 2-CSP.

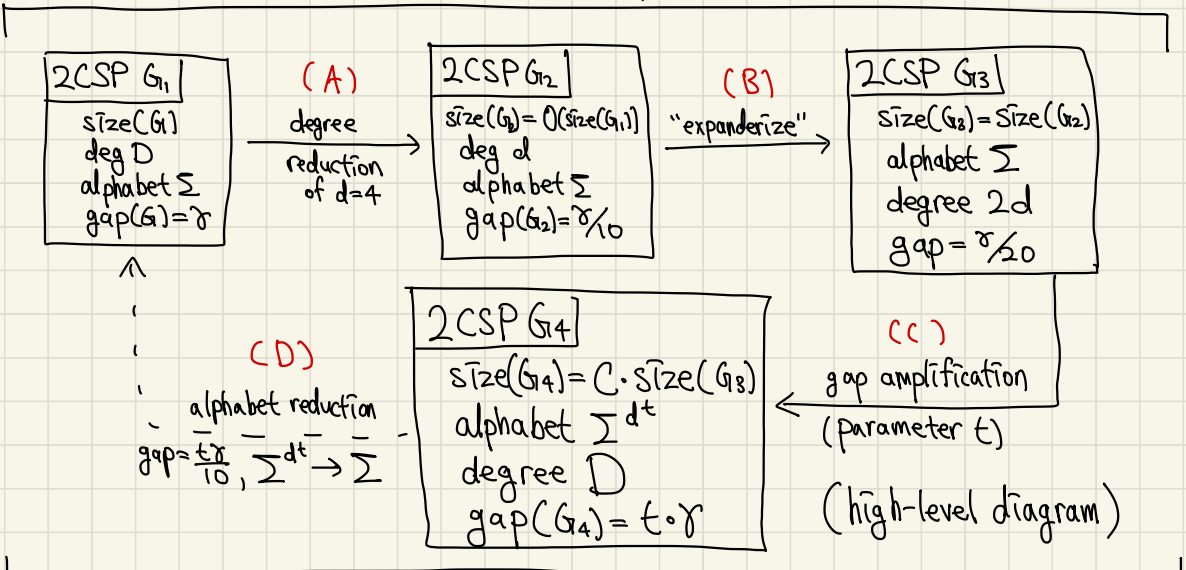
3COL $\in$ NP-C. $\Rightarrow$ Completeness: $gap(G_1) = 0$. Soundness: $\forall$ coloring of graph G_1, $\exists$ a violated edge $\Rightarrow gap(G_1) \geq \frac{1}{|E|} \simeq \frac{1}{n^2}$. → how to make this a constant?

Idea: Perform a series of reductions to amplify the gap to, say, 0.1.

The Reduction: [2CSP size$(G_1)=n$. gap$(G_1)=\gamma$. Σ. G] $\longrightarrow$ [2CSP size$(G')=C(1)\cdot n$. gap$(G')=2\gamma$. Σ. G']

↳ repeat this reduction $\Theta(\log n)$ times, starting from 3-COL(G_1). Then,

size$(G^*) = C^{\log n} \cdot n = poly(n)$, while gap is amplified.



(High-level diagram)

(A)

Degree Reduction: $deg(G_1) = D$. Produce a reduction to G' s.t. $deg(G') = 4$.

Suppose $\exists v$ with > 4 constraints. Make D copies of it, $v_1, v_2, \dots, v_D$. Assign

a single edge to each of them. Then, connect $v_1, \dots, v_D$ with a degree 3 expander of equality constraints. → # of vertices in G' = # of edges in G, and $\deg(G') = 3+1$ (3 for expander, 1 for single edge).

Soundness: Suppose $\text{gap}(G) \geq \gamma \Rightarrow \text{gap}(G') \geq \gamma/40$. Call the set $\{v_1, \dots, v_D\}$ cloud(v). Fix any assignment $\sigma' : V(G') \to \Sigma$ for G'.

Define $\sigma : V(G) \to \Sigma$ where $\sigma(v) = \underset{\text{↳(most popular)}}{\text{Plurality}}\{\sigma(v^{(i)}) \mid v^{(i)} \in \text{cloud}(v)\}$.

Define $\text{Bad}_v := \{v^{(i)} \mid \sigma'(v^{(i)}) \neq \sigma(v)\}$, the vertices inconsistent with equality.

Proof of Soundness: $\forall$ original edge $(u^{(i)} \to v^{(i)}) \in$ constraints in G' from G, if σ violates $(u^{(i)} \to v^{(i)})$, then either 1) σ' also violates that constraint, or 2) $\sigma(u) \neq \sigma'(u^{(i)})$ or 3) $\sigma(v) \neq \sigma(v^{(i)})$. Then, # of edges violated by $\sigma \leq$ # of original edges violated by $\sigma' + \sum_v |\text{Bad}_v|$ (union).

Also, # of edges violated by $\sigma \geq \gamma \cdot |E|$.

$\Rightarrow \underbrace{\#\text{ of original edges violated by } \sigma'}_{\alpha} + \sum_v |\text{Bad}_v| \geq \gamma \cdot |E|$.

Case 1) $\alpha \geq \gamma/2 \cdot |E| \geq \gamma/8 \cdot |E'|$ since $|E'| = 5/2\,|E|$ (some constant mult)

Case 2) $\sum_v |\text{Bad}_v| \geq \gamma/2\,|E|$, total # of minority vertices $\geq \gamma/2 \cdot |E|$. By expansion of degree-3 graphs, in cloud(v), at least $|\text{Bad}_v|/5$ equality constraints are violated (some constant fraction). ⇒ # of violated constraints $\geq \sum_v |\text{Bad}_v|/5 \geq \gamma/10\,|E| \geq \gamma/40\,|E'|$. //

(B) Expanderizing: $\deg(G)=4$, but not an expander. $\text{gap}(G)=\gamma/40$.

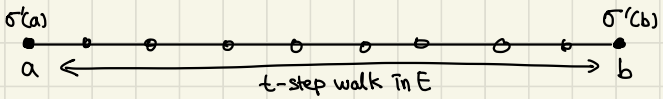
$G'=G\cup H$ where $H:=$ some degree-4 expander. $V(G')=V(G)$, but $E(G')=E(G)\cup E(H)$. Constraints on H are trivial $C_e=\sum\times\sum$ $\forall e\in H$. (↳ all possible pairs)

$\text{gap}(G')=\text{gap}(G)/2$.

(C) Gap Amplification: G is a (n,d,λ)-spectral expander $\to G'$.

Idea: $V'=V$. For assignment in G, $\sigma: V\to\Sigma$, in G', $\sigma'(v):=$ "opinion of v on all vertices up to distance t", i.e. $\forall w$ s.t. $\text{dist}(v,w)\le t$, $(\sigma'(v))_w\in\Sigma$, which is "v's opinion on value on value to be assigned to w". $|\sigma'(v)|<1+d+d^2+\cdots+d^t$. Thus, $\sigma': V\to\Sigma^{1+d+\cdots+d^t}$. Constraints are:

$E':=\{$all "t-step" ($\xleftarrow{\text{random}}$) walks in $G\}$, i.e. $(a,b)\in E'\Leftrightarrow(a,b)$ are endpoints of a t-step walk in G.



$\sigma'(a)$ holds opinion on all vertices on the path, as well as $\sigma'(b)$. Check if they are consistent along all vertices on the path AND that they satisfy the constraints on each edges on the path.

Completeness: If $\exists\sigma$ satisfying G, then let σ' be constructed by it. //

Soundness Goal: If $\text{gap}(G)=\gamma$, then we want $\text{gap}(G')\ge t\gamma$.

↳ σ' is any assignment, and it will violate $t\gamma|E'|$ or more edges.

Proof Strategy: Use σ' to define an assignment $\sigma: V \to \Sigma$. Define $\sigma(u) := \text{Plurality}\{(\sigma'(v))_u \mid v \in N_t(u)\}$ ($*N_t(u)$ means t-distance neighbors of u)

Since $\text{gap}(G) \geq \gamma$, σ violates γ-fraction of edges. Let $F := \{\text{edges violated by } \sigma\} \subseteq E$. Consider a t-step walk in G (an edge in G'). If the walk contains an edge in F, say (u,v), $(\sigma(u), \sigma(v))$ violates it. Since σ is a majority vote, $(\sigma'(a))_u \simeq \sigma(u)$, and $(\sigma'(b))_v \simeq \sigma(v)$, and the "average" opinion of a and b also violates (u,v), thus $(\sigma'(a), \sigma'(b))$ violates the t-step walk. $\Rightarrow$ When a t-step walk contains an edge in F, the t-step walk is violated. (w.c.p., intuitively)

In an expander graph, every step of a walk is a "random edge".

↳ $\Pr[\text{hit an edge in } F \text{ in a t-step walk}] \simeq 1-(1-\gamma)^t \simeq t\gamma$. ($\gamma \to$ gap of G')

Formalisms: First, we concretely define "t-step walks".

Def) After Stopping Random Walk (ASRW):

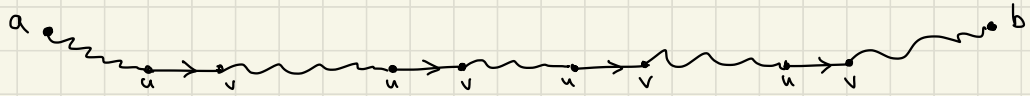
1) Pick $v \in V$ u.a.r. 2) Take a random step to a neighbor.

3) Stop w.p. $1/t$. 4) Else, go to step 2. $\to E[\text{length of walk}] = 1+t$.

Def) Before Stopping Random Walk (BSRW): Interchange steps 2 and 3 from ASRW. $E[\text{length of walk}] = t$.

↳ In both cases, $\mathbb{E}[\text{length of walk}] \simeq t$, mimicking a t-step walk. $\Theta_{u\to v,=k}$

Lemma) Fix a $k \in \mathbb{N}$, $G(V,E)$. Consider $[\text{ASRW} \mid \underline{\text{exactly}}\ k \text{ steps on edge } (u,v)]$.



1) Distribution of b in $\Theta_{u\to v,=k} \equiv$ BSRW starting at v w.p. $1/t$.

2) Distribution of a in $\Theta_{u\to v,=k} \equiv$ ASRW starting at u w.p. $1/t$.

3) a and b are independent.

Consider $\Theta_{u\to v,\geq k} := [\text{ASRW} \mid \text{at least } k \text{ steps on } (u,v)]$. It satisfies the above 1,2,3. $\Theta_{u\to v,=k} = \Theta_{u\to v,\geq k} - \Theta_{u\to v,\geq(k-1)}$ with abuse of notation, and the "subtraction" of two distributions still preserves properties. //

Now, we formally define $E' := \{$ all ASRWs (of $o(t \text{ th} \Sigma)$ steps) in $G\}$, i.e. $(a,b) \in E' \iff$ (a,b) are endpoints of an ASRW. Then, G' is a weighted multigraph. ($\to w := \Pr[\text{walk happens}]$)

We can normalize all weights into integers, then replace them with multiedges.

For $\forall$ edge $u\to v$ in walk s.t. $\text{dist}(a,u) \leq t$ AND $\text{dist}(b,v) \leq t$, check whether $((\sigma'(a))_u, (\sigma'(b))_v)$ satisfy the constraint on (u,v).

Formal Proof: Start with $\sigma' : V \to \sum^{1+\cdots+d^t}$ in G'. Define $\sigma(v) :=$

$\text{Plurality}\{(\sigma'(b))_v \mid b \text{ is an endpoint of BSRW starting at } v \wedge \text{dist}(b,v) \leq t\}$

σ must violate the set of edges F where $\underline{|F| \geq \gamma|E|}$ ($\because \text{gap}(G) \geq \gamma$).

Def) Faulty Step: In an ASRW $a \rightsquigarrow b$, an edge $u \to v$ is a faulty step if: 1) $(\sigma'(a))_u = \sigma(u)$ 2) $(\sigma'(b))_v = \sigma(v)$ 3) $(\sigma(u), \sigma(v))$ violate constraint (u,v).

Obs) If an ASRW $a \rightsquigarrow b$ has a faulty step, constraint (a,b) is violated by σ'.

$\Pr[\text{ASRW } a \rightsquigarrow b \text{ is violated by } \sigma'] \geq \Pr[\exists \text{ faulty step } u \to v \in a \rightsquigarrow b]$. ($\to \frac{|F'|}{|E'|}$)

Define $S :=$ # of steps of a walk, $N_F :=$ # of steps that are faulty edges

$N :=$ # of faulty steps in ASRW. $N_* := N \cdot 1\{S \leq B\}$. ($B = \Theta(t \ln \Sigma)$)

$\Rightarrow \Pr[\exists \text{ a faulty step } u \to v \in a \rightsquigarrow b] = \Pr[N_* > 0]$. (observe $N_* \leq N \leq N_F$)

Goal: $\Pr[N_* > 0] \geq \frac{E[N_*]^2}{E[N_*^2]} \geq t\gamma$. (second moment) ($E[N_*]^2 \to$ lower bound, $E[N_*^2] \to$ upper bound)

$E[S] = \underline{t+1}$. $E[N_F \mid S = \ell] = \sum_{i=1}^{\ell} E[\mathbb{1}\{i\text{-th step} \in F\}] = \sum_{t=1}^{\ell} \sum_{(u,v) \in F} E[\mathbb{1}\{t\text{-th step} = (u,v)\}]$

$= \sum_{t=1}^{\ell} \sum_{(u,v) \in F} \Pr[t\text{-th step} = (u,v)] = \sum_{t=1}^{\ell} \sum_{(u,v) \in F} \frac{1}{|E|} = \underline{\ell \frac{|F|}{|E|}}$. $E[N] = \sum_{t=1}^{\infty} \sum_{(u,v) \in F} E[\mathbb{1}\{t\text{-th step is a faulty step on } (u \to v)\}]$

$= \sum_{(u,v) \in F} \sum_{t=1}^{\infty} \Pr[t\text{-th step is } (u \to v) \text{ and } (u \to v) \text{ is faulty}] = \sum_{(u,v) \in F} E[(\#\text{ of } (u \to v) \text{ steps}) \times \mathbb{1}\{(u \to v) \text{ is a faulty step}\}]$

$= \sum_{(u,v) \in F} \sum_{k=1}^{\infty} k \cdot \Pr[\#\text{ of } (u \to v) \text{ steps} = k] \cdot \underline{\Pr[(u \to v) \text{ is a faulty step} \mid \#\text{ of } (u \to v) \text{ steps} = k]}$.

$\to \Pr[(\sigma'(a))_u = \sigma(u), (\sigma'(b))_v = \sigma(v) \mid \#(u \to v) \text{ steps} = k]$

$= \Pr[(\sigma'(a))_u = \sigma(u) \mid \#(u \to v) \text{ steps} = k] \cdot \Pr[(\sigma'(b))_v = \sigma(v) \mid \#(u \to v) \text{ steps} = k]$

$= \Pr_{\text{BSRW from } u}[(\sigma'(a))_u = \sigma(u)] \Pr_{\text{BSRW from } v}[(\sigma'(b))_v = \sigma(v)] \geq \left(\frac{1}{|\Sigma|}\right)\left(\frac{1}{|\Sigma|}\right)$. ($= \Theta(t\gamma \frac{1}{|\Sigma|^2})$)

$\to \geq \frac{1}{|\Sigma|^2} \sum_{(u,v) \in F} \sum_{k=1}^{\infty} k \cdot \Pr[\#\text{ of } (u \to v) \text{ steps} = k] = \frac{1}{|\Sigma|^2} \cdot E[N_F] \geq \underline{(t+1) \cdot \frac{|F|}{|E|} \cdot \frac{1}{|\Sigma|^2}}$. //

$E[N^*] = E[N \cdot 1\{S \leq B\}] \geq \underline{\frac{t\gamma}{8|\Sigma|^2}}$ (without proof).

$E[N^{*2}] \le E[N^2] \le E[N_F^2] = E[(\sum_i X_i)^2]$ where $X_i := \mathbb{1}\{i\text{-th step} \in F\}$.

$\to \sum_{i=1}^{\infty} E[X_i] + 2\sum_{i=1}^{\infty}\sum_{j>i} E[X_i X_j] = E[N_F] + 2\sum_{i=1}^{\infty} Pr[i\text{-th} \in F] \sum_{j>i} Pr[j\text{-th} \in F \mid i\text{-th} \in F]$.

Reminder: G is a (n,d,λ)-expander, $F \subseteq E$. Then,

$Pr[(t+\ell)\text{-th step} \in F \mid t\text{-th step} \in F] \le \frac{|F|}{|E|} + \left(\frac{\lambda}{d}\right)^{\ell}$.

$\to \le E[N_F] + 2\sum_{i=1}^{\infty} \underbrace{Pr[i\text{-th} \in F]}_{\frac{|F|}{|E|}} \cdot \underbrace{Pr[\#\text{ of extra steps} \ge j-i]}_{(1-1/t)^{j-i}} \cdot \underbrace{Pr[j\text{-th} \in F \mid i\text{-th} \in F]}_{\le \frac{|F|}{|E|} + (\frac{\lambda}{d})^{j-i}}$

$\le t \cdot \frac{|F|}{|E|} = \Omega(t\gamma)$. $\Rightarrow Pr[N_* > 0] \ge \frac{\Omega(t^2\gamma^2)}{\Theta(t\gamma)} \ge t\gamma$. //

(D)

Alphabet Reduction: $G(V,E)$, $\Sigma = [2^k]$. constraint $C_{uv} \subseteq \Sigma \times \Sigma$.

$u,v \in \{\ell\text{-bit strings}\}$. we want a constant # of bits. Consider the PCP verifier that reads $(u,v) \in E$ and verifies $(\sigma(u), \sigma(v))$ satisfies constraints C_{uv}.

↳ This reads 2ℓ bits now, but can we reduce # of bits read?

$\Leftrightarrow$ How to check if $(\sigma(u), \sigma(v))$ satisfies C_{uv} by reading $\le 2\ell$ bits?

↳ Solve by recursion, or, an exponential-sized PCP! (proof composition)

Recall: For a circuit C with K bit input, $\exists$ a PCP of size $2^{O(K^2)}$ s.t. by reading 20 bits of the proof, we can check that $\exists z$ s.t. $C(z) = 1$.

↳ Our circuit is $C(\sigma(u), \sigma(v)) = \mathbb{1}\{\sigma(u), \sigma(v) \text{ satisfies } C_{uv}\}$.

$\Rightarrow$ prover gives exp-sized proof that "Circuit C_{uv} is satisfiable". But is $(\sigma(u), \sigma(v))$ satisfiable, i.e. $C_{uv}(\sigma(u), \sigma(v)) = 1$??

Digression: Assignment Testers (PCP Proximity Test)

Def) Reduction Algorithm: A q-query tester $AT(\gamma>0, \Sigma_0)$ is a RA P,

P: {Boolean Circuit C} $\to$ {Constraint System Γ of 1) Boolean variables X,

2) Auxillary Variables Y}. Γ satisfies properties:

1) If $C(X)=1$, then $\exists Y$ s.t. (X,Y) satisfy all constraints of Γ.

2) If X is δ-far from any satisfying assignments, then $gap(\Gamma) \geq \gamma \cdot \delta$.

$\Leftrightarrow \forall Y$, (X,Y) violates $\gamma \cdot \delta$ constraints.

↳ Checks if X is close to a real satisfying solution to C.

⇒ Now, verifier can check if C_{uv} is satisfied by something close to $(\sigma(u), \sigma(v))$.

Let E, an error correcting code, : $\{0,1\}^\ell \to \{0,1\}^{100\ell}$ s.t. $\text{dist}_{\text{Hamming}}(E(\alpha), E(\beta)) \geq 0.3$. Then, use $E(\sigma(u))$ (↳ $\sigma'(u)$) and $E(\sigma(v))$ (↳ $\sigma'(v)$) (of each length 100ℓ) and some modified constraint C'_{uv} that checks 1) $\sigma'(u)$ & $\sigma'(v)$ are valid codewords and 2) $E^{-1}(\sigma'(u))$ & $E^{-1}(\sigma'(v))$ satisfies C_{uv}. A proximity tester for C' will then give whether $(\sigma'(u), \sigma'(v))$ is close to satisfying C'. But in the codomain of E, being close is equivalent to being equal since every point is sparse. Thus, proximity testing suffices for testing C_{uv}.

Fix $\sigma' : V' \to \{0,1\}^{100\ell}$. If $gap(G_1) \geq \varepsilon \Rightarrow AT$ rejects w.p. $\varepsilon/100$.

Define $\sigma(u) :=$ nearest codeword decoding of $\sigma'(u) \in \{0,1\}^\ell$. We know that $\sigma(u)$ violates gap(G) constraints. Suppose edge (u,v) is violated by σ. Then, nearest codewords to $\sigma'(u)$ & $\sigma'(v)$ violate C_{uv}.

$\Rightarrow (\sigma'(u), \sigma'(v))$ is far from satisfying C_{uv} (by virtue of being close to a wrong one).

Reduction from Q-CSP ($\nearrow$ Q-query PCP) $\rightarrow$ 2-CSP ($\nearrow$ 2-query PCP): Let Q-CSP = 3-SAT. $C_1, \ldots, C_m$ are clauses, $X_1, \ldots, X_n$ are variables. For (clause − variable) edge, constraint is consistency between $C_i = (a \vee b \vee c)$ and $V_j = 0$ or 1. If gap(3SAT-instance) $= \gamma$, gap(graph) $\geq \frac{\gamma}{3}$. $\Rightarrow$ AT can be reduced to 2-queries.

Recall: For Exp-sized PCP for Circuit-SAT, to check $\exists x$ s.t. $C(x) = 1$, we constructed a reduction to a system of quadratic equations. The prover wrote down a Hadamard encoding of Z and ZZ^T, where they claim that Z satisfies all of the quad eqs. $\rightarrow$ constructs AT as desired.

SDP-Based Algorithms

MAXCUT: $G(V,E) \rightarrow (S, \bar{S})$ s.t. $|E(S,\bar{S})|/|E|$ is maximized.

↳ can be thought of as a CSP that assigns $\{0,1\}$ membership to S.

$\Rightarrow$ NP-Hard to solve exactly, how well can we approximate it?

PCP says that as long as $P \neq NP$, $\exists \varepsilon \ll 1$ s.t. $\nexists$ a $(1-\varepsilon)$-approx. factor.

↳ ∴ we proved a reduction of $(\underset{(3\text{-COL})}{NP\text{-}C}) \xrightarrow{(G)} (\text{Gap } 2\text{-CSP})$ s.t.:

if 3-COL is satisfiable, $gap(H) = 0$, and if not, $gap(H) > \varepsilon$, say 10^{-6}.

If we have a more fine-grained approx. factor α, say 10^{-12}, then we have a decision oracle for an NP-C problem, so we cannot have it. //

We can also reduce $(\text{Gap } 2\text{-CSP}) \xrightarrow{(H)} (MAXCUT)$, where $val(H) = 1$

$\Rightarrow val(H') = \beta$, $val(H) < 1 - 10^{-6} \Rightarrow val(H') < (1-10^{-6}) \cdot \beta (1-\gamma)$.

Def) Positive Semi-Definiteness: symmetric $M \succeq 0$ iff,

1) $\lambda_i(M) \geq 0 \ \forall i \in [n]$, 2) $M = VV^T$ where $v_1, \ldots, v_n \in \mathbb{R}^n$ s.t. $\langle v_i, v_j \rangle = M_{ij}$.

3) $\forall v$, $v^T M v \geq 0$.

Fact: PSD in $\mathbb{R}^{n \times n}$ forms a convex cone, and $\exists$ a seperation oracle s.t. if $M' \notin PSD(\mathbb{R}^{n \times n})$, we can find a seperating hyperplane between PSD & M'.

↳ M' has a negative $\lambda_i(M')$ w/ v, so $v^T M' v < 0 \Rightarrow \sum M'_{ij} v_i v_j < 0$.

Then, since $\forall M \in PSD$, $v^T M v \geq 0$, $\sum M_{ij} v_i v_j \geq 0$, so $f(M) = \sum M_{ij} v_i v_j$ is a seperating oracle. //

Naïve Rand. Algo.: $E[\#\text{ cuts}] = \frac{1}{2}|E|$. (not useful, a crude lower bound)

Consider a degree $D=100$ random graph and a bipartite graph of degree $D=(1-\varepsilon)D$. LP cannot distinguish between these, so it is no better than $\frac{1}{2}|E|$.

[Goeman-Williamson]

SDP-based MAXCUT: $X_1,\dots,X_n$ are variables, $X_i=\begin{cases}+1 & \text{if } i\in S\\ -1 & \text{if } i\notin S\end{cases}$.

Obj: $\max_{x\in\{0,1\}^n} \text{Val}_G(x)=\frac{1}{|E|}\sum_{(i,j)\in E}\frac{(X_i-X_j)^2}{4}$ s.t. $(X_i-X_j)^2=4$ iff i,j are seperated.

Relaxation: $X_i\in\{\pm 1\}\Rightarrow V_i\in\mathbb{R}^d$. Let $OPT_d=\max_{V_i\in\mathbb{R}^d,\|V_i\|^2=1}\sum_{(i,j)}\|V_i-V_j\|^2$.

At $d=n$, OPT_n is actually a convex program (solvable in P)!

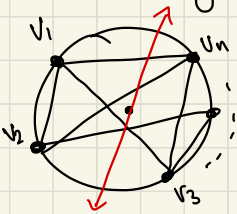
↳ $OPT_n=\max_{V_i\in\mathbb{R}^n,\|V_i\|^2=1}\frac{1}{|E|}\sum_{(i,j)}\frac{\|V_i-V_j\|^2}{4}=\max_{V_i\in\mathbb{R}^n,\,V_i\cdot V_i=1}\frac{1}{|E|}\sum_{(i,j)}\frac{(V_i-V_j)\cdot(V_i-V_j)}{4}$

$=\max\frac{1}{|E|}\sum_{(i,j)}\frac{(V_i\cdot V_i+V_j\cdot V_j-2V_i\cdot V_j)}{4}\Rightarrow\max\frac{1}{|E|}\sum_{(i,j)}V_i\cdot V_i+V_j\cdot V_j-2V_i\cdot V_j$ sub. to $V_i\cdot V_i=1$.

Now consider an LP defined $V_i\cdot V_i=M_{i,i}$, $V_j\cdot V_j=M_{j,j}$, $V_i\cdot V_j=M_{i,j}$. Then, $M=(M_{ij})$ is a PSD, and we can rewrite to $\max_M\frac{1}{|E|}\sum(M_{ii}+M_{jj}-2M_{ij})$ subject to $M_{ii}=1$ and $M\succeq 0$. Optimize over M to get $V=(v_1,\dots,v_n)$. //

* SDP $\simeq$ LP over inner products of vectors $v_1,\dots,v_N\in\mathbb{R}^N\simeq$ Opt. over PSD cone.

So $OPT_n(G)=SDP(G)\geq MAXCUT(G)=OPT_1(G)$. In a $\mathbb{R}^n$ unit ball, the graph will now be embedded on the surface. Take a random hyperplane cut through the origin, $H_r=\{x\in\mathbb{R}^n\mid r\cdot x=0\}$.



Output $S=\{i\mid v_i\cdot r\geq 0\}$, $\bar{S}=\{i\mid v_i\cdot r<0\}$. 0.878 factor. (proof omitted)

Consider again $\max_{x\in\{\pm1\}^n} \mathrm{Val}_G(x)$. How do we "convexify" it? The most convenient way is to optimize over probability distributions.

↳ $\max_{\mu} \mathop{\mathbb{E}}_{x\sim\mu}[\mathrm{Val}_G(x)] = \sum_{x\in\{\pm1\}^n} \mu(x)\cdot \mathrm{Val}_G(x)$ where $\mu(x) = \Pr[x \text{ is selected}]$.

Now, the set of all μ is convex. However, $|\mu| = 2^n$. Can we do better?

Def) Moments of μ: μ is a prob. dist. over $\in\{0,1\}^n$. degree 1 moments are $\mathop{\mathbb{E}}_{x\sim\mu}[x_1], \dots, \mathop{\mathbb{E}}_{x\sim\mu}[x_n]$, degree 2 is $\mathop{\mathbb{E}}_{x\sim\mu}[x_i x_j], \dots$ and so on.

$$\mathop{\mathbb{E}}_{x\sim\mu}[\mathrm{Val}_G(x)] = \mathop{\mathbb{E}}_{x\sim\mu}\left[\frac{1}{|E|}\sum_{(i,j)}\frac{(x_i-x_j)^2}{4}\right] = \frac{1}{|E|}\sum_{(i,j)}\left(\mathop{\mathbb{E}}_{\mu}[x_i^2] + \mathop{\mathbb{E}}_{\mu}[x_j^2] - 2\mathop{\mathbb{E}}_{\mu}[x_i x_j]\right)/4$$

→ this is a linear function over all deg. 2 moments of μ!

⇒ (lossy step) $\max_{\{\text{deg 2 moments}\}} \frac{1}{|E|}\sum_{(i,j)} \frac{\mathbb{E}[x_i^2]+\mathbb{E}[x_j^2]-2\mathbb{E}[x_i x_j]}{4}$, let $X_{ij} := \mathbb{E}[x_i x_j]$.

→ $\frac{1}{|E|}\max_{\{X_{ij}\}} \sum_{(i,j)} X_{ii} + X_{jj} - 2X_{ij}$ sub. to $X_{ii}=1$, and $X \succeq 0$. This is because moment matrices are PSD, and X is a deg. 2 moment matrix.

Then, this is a SDP problem. In fact, we can impose a deg. d moment matrix for a finer resolution of the constraint.

Current knowledge: for MAXCUT, deg 2 (SOS-SDP) → 0.878 factor is the best known, and deg 4 → unknown?? However, under Unique Games Conjecture (an unproven hardness assumption), deg 2 SOS is indeed optimal in P.

i.e. It is NP-Hard to beat it (when the instance is not satisfiable).

The Whole Point: this scheme yields a more general way to SDP-fy any reasonable NP-Complete problem, say 3-SAT. And again, UGC says that deg. 2 SOS will be optimal (basic SDP relaxation).

Hastad's 3 Query PCP

PCP(3 bits, poly(n)) → Completeness $(1-\varepsilon)$, Soundness $(\frac{1}{2}+\varepsilon)$ (The goal)

Fourier Analysis of the Boolean Cube: $\{0,1\}^n \equiv \mathbb{Z}_2^n$.

$x, y \in \{0,1\}^n \Rightarrow x \oplus y$, which is just $x+y \pmod 2$ forms a group.

Def) Character: $\chi : \{0,1\}^n \to \mathbb{R}$ s.t. $\chi(a \oplus b) = \chi(a) \cdot \chi(b)$.

ex) $\chi_i : \{0,1\}^n \to \mathbb{R}$. $\chi_i(x) = (-1)^{x_i}$. $\chi_i(x \oplus y) = (-1)^{x_i \oplus y_i} = (-1)^{x_1} \cdot (-1)^{x_2}$

$\chi_i(x) = (-1)^{x_i}$ are characters.

The list of all characters: $\forall S \subseteq [n]$, $\chi_S(x) := (-1)^{\oplus_{i \in S} x_i} = \prod_{i \in S} \chi_i(S)$.

Inner Product on functions on $\{0,1\}^n$: $\forall f, g : \{0,1\}^n \to \mathbb{R}$, $\langle f, g \rangle := \frac{1}{2^n} \sum_x f(x) g(x) = \mathop{E}_x [f(x) g(x)]$. This behaves well; $\langle \chi_i, 1 \rangle = \mathop{E}_x [(-1)^{x_i} \cdot 1] = 0 \Rightarrow \langle \chi_S, 1 \rangle = 0 \ \forall S \neq \emptyset$.

$\chi_S(x) \cdot \chi_T(x) = (-1)^{\sum_{i \in S} x_i} (-1)^{\sum_{i \in T} x_i} = (-1)^{\sum_{i \in S \Delta T} x_i} = \chi_{S \Delta T}(x) \Rightarrow \langle \chi_S, \chi_T \rangle = 0$ unless $S = T$.

$\langle \chi_S, \chi_S \rangle = \mathop{E}_x [\chi_S(x)^2] = 1$. $\therefore \{\chi_S\}$ is an orthonormal basis of set of all $f : \{0,1\}^n \to \mathbb{R}$.

$\Rightarrow$ we can write ANY function in the basis of χ_S!

Suppose $f:\{0,1\}^n \to \mathbb{R}$. Then, $f = \sum_S \langle f, \chi_S \rangle \cdot \chi_S = \sum_S \hat{f}_S \chi_S$ where $\hat{f}_S := \mathop{\mathbb{E}}_x[f(x)\cdot\chi_S(x)]$. Thus, $\{f(x) \mid x \in \{0,1\}^n\} \iff \{\hat{f}_S \mid S \subseteq [n]\}$.

Theorem) Parseval's Identity: Given $f:\{0,1\}^n \to \mathbb{R}$, $\underline{\|f\|^2 = \frac{1}{2^n}\sum_x f(x)^2 = \sum \hat{f}_S^2}$.

Proof: $\|f\|^2 = \langle f, f\rangle = \langle \sum_S \hat{f}_S \chi_S, \sum_T \hat{f}_T \chi_T\rangle = \sum_S \hat{f}_S^2 + \cancel{\sum_{S\neq T} \hat{f}_S \hat{f}_T \underbrace{\langle \chi_S, \chi_T\rangle}_{0}}^{0}$. //

Linearity Test: Given $f:\{0,1\}^n \to \{0,1\}$, test if f is close to linear, i.e. $f(x) \simeq \sum_i b_i x_i$ for most x.

The Test: Pick $x, y \in \{0,1\}^n$. Test if $f(x)+f(y) = f(x+y)$.

Analysis: $\Pr_{x,y}[f(x)+f(y)-f(x+y)=0]$? Define $F(x) := (-1)^{f(x)} = \begin{cases} -1 & \text{if } f(x)=0 \\ 1 & \text{if } f(x)=1. \end{cases}$

$\to \Pr_{x,y}[(-1)^{f(x)}\cdot(-1)^{f(y)}\cdot(-1)^{-f(x+y)} = 1] = \underbrace{\mathop{\mathbb{E}}_{x,y}\left[\frac{1+F(x)F(y)F(x+y)}{2}\right]}$.

$\because$ Observe that $\mathbb{1}\{a=0\} = \frac{1+(-1)^a}{2}$. $\mathbb{1}\{f(x)+f(y)-f(x+y)=0\} \equiv \frac{1+F(x)F(y)F(x+y)}{2}$.

$\to \frac{1}{2} + \frac{1}{2}\mathop{\mathbb{E}}_{x,y}[F(x)F(y)F(x+y)] = \frac{1}{2} + \frac{1}{2}\mathop{\mathbb{E}}_{x,y}\left[\left(\sum_S \hat{F}_S \chi_S(x)\right)\left(\sum_T \hat{F}_T \chi_T(y)\right)\left(\sum_U \hat{F}_U \chi_U(x+y)\right)\right]$

$= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U} \hat{F}_S \hat{F}_T \hat{F}_U \underbrace{\mathop{\mathbb{E}}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x+y)]}_{\mathop{\mathbb{E}}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x)\chi_U(y)]}$

$= \mathop{\mathbb{E}}_x[\chi_S(x)\chi_U(x)]\mathop{\mathbb{E}}_y[\chi_T(y)\chi_U(y)] = \langle \chi_S, \chi_U\rangle\langle \chi_T, \chi_U\rangle = \mathbb{1}\{S=T=U\}$.

$\to \frac{1}{2} + \frac{1}{2}\sum_{S,T,U} \hat{F}_S \hat{F}_T \hat{F}_U \mathbb{1}\{S=T=U\} = \frac{1}{2} + \frac{1}{2}\sum_S \hat{F}_S^3$.

$\Rightarrow$ If $\frac{1}{2} + \frac{1}{2}\sum_S \hat{F}_S^3 \geq \frac{1}{2} + \delta$, then $\sum_S \hat{F}_S^3 \geq 2\delta$. We know that $\sum_S \hat{F}_S^2 = 1$.

$\underline{2\delta \leq \sum_S \hat{F}_S^3 \leq \sum_S \hat{F}_S^2 \cdot \hat{F}_S \leq \sum_S \hat{F}_S^2 \cdot \left(\max_T\{\hat{F}_T\}\right) \leq \max_T\{\hat{F}_T\}}$.

$F_\tau(x) = \mathbb{E}_x[f(x)\chi_\tau(x)] = \mathbb{E}_x[(-1)^{f(x)} \cdot (-1)^{\sum_{i\in\tau} x_i}] = 2\Pr_x[f(x) = \sum_{i\in\tau} x_i] - 1.$

⇒ maximum alignment is bounded by below! //

Recall: MAX 3-SAT is hard to approximate better than $1 - 10^{-6}$.

In general, NP → 3SAT ϕ ⇒ if NP is unsat., $gap(\phi) \geq 10^{-6}$.

There is also a reduction of 3SAT $\phi \to$ 2-CSP $G \Rightarrow gap(\phi) \propto gap(G)$

↳ How would we increase the gap?

Idea: Parallel Repetition. Let $H^{(k)} := (U \cup V, E)$ where $U :=$ k-tuples of clauses from ϕ, $V :=$ k-tuples of variables from ϕ. $(C_{i_1}, \dots, C_{i_k}) \leftrightarrow (x_{i_1} \dots x_{i_k})$ $\overset{\in U}{\quad}$ $\overset{\in V}{\quad}$

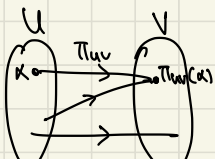
iff $(x_{i_1} \in C_{i_1}) \wedge (x_{i_2} \in C_{i_2}) \cdots \wedge (x_{i_k} \in C_{i_k})$. Let an assignment A:

$U \to \{0,1\}^{3k}$

$V \to \{0,1\}^{k}$.

Theorem) $val(\phi) = 1 \Leftrightarrow val(H^{(k)}) = 1$. $val(\phi) \leq 1-\varepsilon \Leftrightarrow val(H^{(k)}) = e^{-\Omega(k)}$. ← stated w/o proof here.

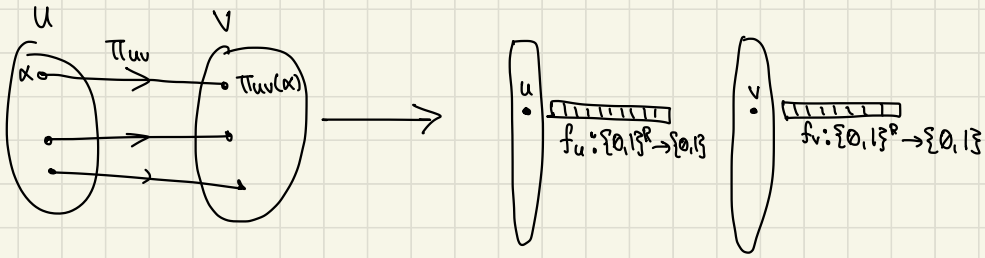
Label Cover (2-CSP): $H := (\overset{U \cap V = \emptyset}{U \cup V}, E, \overset{\to \text{alphabet}}{[R]}, \{\pi_{uv} : [R] \to [R] \mid (u,v) \in E\})$.



An assignment $A : U \cup V \to [R]$ satisfies an edge (u,v) if $\pi_{uv}(A(u)) = A(v)$.

↳ ∃ reduction NP → Label Cover s.t. $val(NP) < 1 \Rightarrow val(H) < \frac{1}{|R|^{0.1}}$.

Proof Idea: Label Cover → 3-LIN (lin. eq. mod 2 w/ 3 variables). i.e. $x+y+z = \{0,1\}$

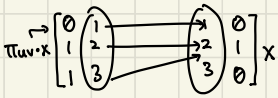


If $A(u) = j \in \{1, \dots, R\}$, we want $f_u(x) = x_j$ (dictator function). Also, $A(v) = \ell \in \{1, \dots, R\}$, we want $f_v(y) = y_\ell$. This the most redundant code.

Consider an inefficient test: Test linearity on f_u & f_v. Test consistency to the map by π_{uv}, i.e. if $f_u(x) = x_j$, then $f_v(y) = y_{\pi_{uv}(j)}$, or, if $f_u(x) \equiv j$-th dictator $\Rightarrow f_v(y) \equiv \pi_{uv}(j)$-th dictator.

For a random x, check $f_u(\pi_{uv} \circ x) = f_v(x)$ where $(\pi_{uv} \circ x)_j := x_{\pi_{uv}(j)}$. (a "pullback")

↳ But this only checks for linearity.



Noise Test: Pick $x \in \{0,1\}^n$. Let $\tilde{x} :=$ ε-noisy x, $\tilde{x}_i = \begin{cases} x_i \text{ w.p. } 1-\varepsilon \\ 0/1 \text{ w.p. } \varepsilon \end{cases}$.

Test if $f_u(x) = f_u(\tilde{x})$. Observe that for a dictator, $f_u(x) = x_i$, $\Pr[f_u(x) = f_u(\tilde{x})] = 1 - \varepsilon$. For linear functions not a dictator, say, a parity function $f_a(x) = \sum_{\tau=1}^{\ell} x_i \pmod 2$, $\Pr[f_a(x) = f_a(\tilde{x})] = \frac{1}{2} + \frac{(1-\varepsilon)^\ell}{2}$.

$\Rightarrow$ Dictators are robust, others are not.

* Hastad's 3 bit PCP wraps all of these tests into one 3-bit query!

1) Pick $(u,v) \in E$ u.a.r. $f_u, f_v : \{0,1\}^R \to \{0,1\}$.

2) Pick $x, y \in \{0,1\}^R$ u.a.r.

3) Test if $\underbrace{f_v(x) = f_u(\pi_{uv} \circ x + y)}_{\text{consistency}} \underbrace{- f_u(\tilde{y})}_{\text{linearity}}$. $\hookrightarrow y+\mu$ where $\mu_i := \begin{cases} 0 & \text{w.p. } 1-\varepsilon \\ 1 & \text{w.p. } \varepsilon \end{cases}$ (small noise) $\hookrightarrow$ noise

Caveat: $\vec{0}(x)$ is a solution, so we insist that $f_u(x) = 1 - f_u(-x)$, same for v.

$\Rightarrow$ The equation is a linear equation of 3 variables! $\to$ 3LIN

Recall that the goal is to get completeness $(1-\varepsilon)$ and soundness $(\frac{1}{2}+o(1))$.

$\hookrightarrow$ Completeness naturally follows from the true dictator, which is robust.

Soundness Analysis: We shall prove that $\mathrm{Val}(3\mathrm{LIN}) \geq \frac{1}{2}+\delta \Rightarrow \mathrm{Val}(\mathrm{Label\ Cover}) \geq \delta'$.

Proof: $F_v(x) = F_u(\pi \cdot x \oplus y \oplus \mu) - F_u(y) \Leftrightarrow F_v(x) \oplus F_u(\pi \cdot x + y + \mu) \oplus F_u(y) = 0$.

Define $f_u(x) = (-1)^{F_u(x)}$. Then, $\Leftrightarrow \dfrac{1 + f_v(x) \cdot f_u(\pi \cdot x + y + u) \cdot f_u(y)}{2}$.

$$\Pr_{u,v,x,y,\mu}[F_v(x) = F_u(\pi \cdot x + y + \mu) - F_u(y)] = \frac{1}{2} + \frac{1}{2} \mathop{\mathbb{E}}_{u,v,x,y,\mu}[f_u(\pi \cdot x + y + \mu) \cdot f_u(y) \cdot f_v(x)] \geq \frac{1}{2} + \delta.$$

Let $A := f_u, B := f_v$. Then, $\mathop{\mathbb{E}}_{u,v,x,y,\mu}[A(\pi \cdot x + y + \mu) A(y) \cdot B(x)] \geq 2\delta$. Expanding,

$$\mathop{\mathbb{E}}_{x,y,\mu}\left[\left(\sum_s \hat{A}_s \chi_s(\pi \cdot x + y + \mu)\right)\left(\sum_\tau \hat{A}_\tau \chi_\tau(y)\right)\left(\sum_u \hat{B}_u \chi_u(x)\right)\right]$$

$$= \sum_{s,\tau,u} \hat{A}_s \hat{A}_\tau \hat{B}_u \mathop{\mathbb{E}}_{x,y,\mu}[\chi_s(\pi \cdot x)\chi_s(y)\chi_s(\mu)\chi_\tau(y)\chi_u(x)]$$

$$= \sum_{s,\tau,u} \hat{A}_s \hat{A}_\tau \hat{B}_u \underbrace{\mathop{\mathbb{E}}_x[\chi_s(\pi \cdot x)\chi_u(x)]}_{\hookrightarrow\ ?^*} \underbrace{\mathop{\mathbb{E}}_y[\chi_\tau(y)\chi_s(y)]}_{\hookrightarrow\ \mathbb{1}\{S=T\}} \underbrace{\mathop{\mathbb{E}}_\mu[\chi_s(\mu)]}_{\hookrightarrow\ \prod_{i \in S} \mathbb{E}[\chi_i(\mu)] = (1-2\varepsilon)^{|S|}}$$

Lemma) $\chi_S(\pi \circ x) = \chi_{\pi_2(S)}(x)$ where $\pi_2(S) := \{j \mid j \text{ appears odd \# of times in } \pi(S)\}$

Proof: $S \subseteq \{1,\dots,R\}$. $\pi(S)$ is a multiset $\subseteq \{1,\dots,R\}$. $\chi_S(\pi \cdot x) = \prod_{i \in S} \chi_i(\pi \cdot x)$

$= \prod_{i \in S} \chi_{\pi(i)}(x) = \prod_{i \text{ pulled odd \#}} \chi_{\pi(i)}(x) \cdot \prod_{i \text{ pulled even}} \chi_{\pi(i)}(x) \quad (\text{even} \to 1) \quad = \chi_{\pi_2(S)}(x)$. //

$\to^* \mathop{E}_x[\chi_S(\pi \cdot x)\chi_u(x)] = \mathop{E}_x[\chi_{\pi_2(S)}(x)\chi_u(x)] = \mathbb{1}\{\pi_2(S) = u\}$.

$\Rightarrow$ Simplifies to $\sum_S \hat{A}_S^2 \hat{B}_{\pi_2(S)} (1-2\varepsilon)^{|S|} > 2\delta$. Now, recall that by Parseval's Identity, $\sum_S \hat{A}_S^2 = 1$. Sample some $S \sim$ w.p. $\hat{A}_S^2 \Rightarrow S \subseteq [R]$.

Decoding Procedure: Pick $\ell \in \mathbb{N}$.

1) For each vertex u, sample $S_u \sim (\hat{f}_{u,S_u})^2 \Rightarrow S_u \subseteq [R]$, $S_v \sim (\hat{f}_{v,S_v})^2$, where $|S_u|, |S_v| < \ell$ (only prefer small subsets).

2) Assign $A(u) \leftarrow$ random label from S_u

$\Rightarrow \mathop{E}_{u,v}\left[\sum_S \hat{A}_S^2 \hat{B}_{\pi_2(S)} (1-2\varepsilon)^{|S|}\right] \le \mathop{E}_{u,v}\left[\sum_S \hat{A}_S^2\right]^{1/2} \mathop{E}_{u,v}\left[\sum_S \hat{A}_S^2 \hat{B}^2_{\pi_2(S)} (1-2\varepsilon)^{2|S|}\right]^{1/2}$ ($\sum_S \hat{A}_S^2 \to 1$, $\mathop{E}_{u,v}[\sum_S \hat{A}_S^2 \hat{B}^2_{\pi_2(S)} (1-2\varepsilon)^{2|S|}] \ge \delta^2$)

$\to \mathop{E}_{u,v}\left[\sum_{S: |S| < \ell} \hat{A}_S^2 \hat{B}^2_{\pi_2(S)}\right] \ge \delta^2/10 \quad \left(\ell \leftarrow \frac{(0 \log(1/\delta)}{\varepsilon}\right)$

$\to$ For a random edge, $\Pr[\pi_2(S_u) = S_v] \ge \delta^2/10$.

$\to \Pr[\pi_{uv}(A(u)) = A(v)] \ge \delta^2/10 \cdot 1/\ell^2 = \delta'(\delta, \varepsilon)$. //

Unique Games

Recall that $3LIN_p$ is a system of lineqs mod $\mathbb{Z}_p$ with 3 variables each.

↳ It is NP-Hard to distinguish $\text{Val}(3LIN_p) \ge (1-\varepsilon) \leftrightarrow \text{Val}(3LIN_p) < 1/p + \varepsilon$

What about 2LIN$_p$? This seems easier than 3LIN$_p$.

Conj) Unique Games Conjecture: $\forall \varepsilon, \exists p$ s.t. it is NP-Hard to distinguish $Val(2LIN_p) \geq 1-\varepsilon \leftrightarrow Val(2LIN_p) < \varepsilon$. [03' Khot]

Def) Unique Games: a generalization of 2LIN$_p$, which is Label Cover where all $\pi_{uv}: [R] \to [R]$ are bijections.

Remark: $\exists$ algorithm A s.t. on input UG instance of value $(1-\varepsilon)$, A outputs an assignment with value $\left(\frac{1}{p}^{\varepsilon/2-\varepsilon}\right)$. [CMM 06']

Remark2: UG on expanders is easy [09']

Remark3: Subexponential time algorithm for UG: runtime 2^{n^ε}. [11']

Remark4: deg 8 SoS SDP relaxation solves "some important UG instances" [12']

[19'] 2-to-1 problem is NP-Hard!

Hardness of MAXCUT

"Assuming Unique Games Conjecture, it is NP-Hard to beat 0.878 factor."

Recall that $MAXCUT(G) = \max_{x \in \{\pm 1\}^n} \left\{ \frac{1}{|E|} \sum_{(i,j) \in E} (x_i - x_j)^2 \right\}$. The SDP assigns $v_i \in \mathbb{R}^n$ $\forall i$ where $\|v_i\|^2 = 1$ s.t. every vertex gets embedded on a unit sphere that maximizes distances between connected vertices. Then a random

cut through the origin determines the rounded cut. → 0.878 factor.

↳ Can we specify a worst case instance where 0.878 is tight?

Consider the "Sphere Graph" (S_α): vertices are all points on S^n, unit sphere on $\mathbb{R}^n$. Edge set is $\{(x,y) \in S^{n-1} \mid \|x-y\| = \alpha\}$. What is the SDPVal$(S_\alpha)$?

All points x are naturally mapped to $\vec{x} \in \mathbb{R}^n \Rightarrow$ SDPVal(G_α) = average squared edge length $= \frac{\alpha^2}{4}$. OPTCut(S_α) = Half-Space Cut (by Borell). Then,

$$\text{OPT val}(S_\alpha) = \Pr_{x,y}[\text{halfspace } H \text{ cuts } x,y] = \Pr_H[H \text{ cuts } x,y \mid \|x-y\| = \alpha] = \emptyset(\alpha).$$

$$\Rightarrow \frac{\text{OPTval}(S_\alpha)}{\text{SDPval}(S_\alpha)} = 0.878.$$

Dictatorship Testing Gadget (H): an instance of MAXCUT where its vertices are hypercubes, $\{\pm 1\}^R$ ($R \to$ const.). A cut is a function $F: \{\pm 1\}^R \to \{\pm 1\}$. We want:

1) Completeness: If $F(x) = x_i$, then $\text{val}_H(F) = \mathop{\mathbb{E}}_{e \in H}[\mathbb{1}\{F(x) \neq F(y)\}] = \frac{\alpha^2}{4}$.

2) Soundness: If F is "far from" all dictators, $\text{val}_H(F) \leq \text{OPT}(S_\alpha)$. ($\downarrow$ SDP(S_α))

↳ What do we mean by "far from"? → Influences: Given $F: \{\pm 1\}^R \to \{\pm 1\}$, the influence of the i-th coordinate is $\Pr_{x \sim \{\pm 1\}^R}[F(x) = F(x \oplus e_i)]$ ($x \oplus e_i$: flipping ith bit).

ex) Suppose $F(x) = x_1$. Then $\text{Inf}_i(F) = \mathbb{1}\{i=1\}$.

$\to \text{Inf}_i(F) := \mathop{\mathbb{E}}_x\left[\frac{(F(x)-F(x \oplus e_i))^2}{4}\right] = \sum_{S \ni i} \hat{F}_S^2$ by Fourier Expansion.

* Inf is well defined for non-boolean functions, too: $F(x) = \mathop{\mathbb{E}}_{x_{-i}}[\mathop{\text{Var}}_{x_i}(F(x))]$.

⇒ Being "far from" a dictator is then having $\forall i$, $Inf_i(F) < \tau = o(1)$.

ex) $Maj(X)$ is far from dictator since $Inf_i(Maj) = O(1) \cdot \frac{1}{\sqrt{R}}$.

Edge set of H is $\{(x,y) \in V \mid Ham(x,y) = \frac{\alpha^2}{4} \cdot R\}$. We can sample edges as follows: pick $x \in \{\pm 1\}^R$, and construct y s.t. $y_i = \begin{cases} x_i & \text{w.p. } 1-\rho \\ -x_i & \text{w.p. } \rho \end{cases}$ where $\rho = 1 - \frac{\alpha^2}{4}$.

Then, we can construct a weighted graph s.t. $w(x,y) = Pr[(x,y) \text{ is sampled}]$.

Lemma) If F is a dictator, then $Val_H(F) = SDP(S_{\alpha})$. [Completeness] $(\alpha^2/4)$

Proof: WLOG, let $F(x) = x_1$. $Val_H(F) = \Pr_{(x,y) \in E_H}[F(x) \neq F(y)] = \Pr_{(x,y)}[x_1 \neq y_1] = \frac{\alpha^2}{4} = SDP(S_{\alpha})$ by construction. //

Soundness: Suppose $\max_i Inf_i(F) < \tau = o(1)$. We want $\Pr_{(x,y)}[F(x) \neq F(y)] \leq OPT(S_{\alpha})$ → (from "Sphere Graph"). Observe that we can view y as a noisy copy of x. Then $\Pr_{(x,y)}[F(x) \neq F(y)]$ is interpreted as noise sensitivity.

Def) Noise Stability: $Pr[F(x) = F(y) \mid y = Noise_{\rho}(x)]$. $E[F] = \emptyset$ (balanced).

→ Dictators are the most noise stable functions of ρ!

Theorem) $Maj(X)$ is the most stable out of functions far from dictators, where $NS_{\rho}(F) \leq NS_{\rho}(Maj) + \varepsilon(\tau)$.

* Recall Gaussian, $N(0,1) \in \mathbb{R}$. $N(0,1)^R \leftarrow (g_1 \dots g_R)$ each of $\mu_i = 0, \sigma_i^2 = 1$.

$\vec{g} \sim N(0,1)^R$ gives a distribution on $\mathbb{R}^R$. Then $\|\vec{g}\| \simeq \sqrt{R}$, i.e. it is spherically symmetric and almost behaves like a unit vector. Now imagine $(g_i, h_i) \leftarrow N((0,0), \begin{bmatrix} 1 & \frac{2p-1}{2} \\ \frac{2p-1}{2} & 1 \end{bmatrix})$ s.t. $E[g_i h_i] = \frac{2p-1}{2}$. With such correlation, we can construct $\vec{g}$ & $\vec{h}$ s.t. they have a certain distance apart.

* CLT says that for $F(x) = \frac{1}{\sqrt{R}} \sum_{i=1}^{R} x_i$, fix any "nice" distribution D over $\mathbb{R}$ (↳ fourth moments are bounded by second) where $\mathop{E}_{x \sim D}[x] = 0$, $\mathop{Var}_{x \sim D}(x) = 1$. Then $F(x) \mid x \sim D^R \xrightarrow[R \to \infty]{} F(g) \mid g \sim N(0,1)^R$.

Theorem) Invariance Principle: F is a polynomial of $(x_1, \dots, x_n)$ of degree D and $\max_i Inf_i(F) < \tau$. Then $F(x) \mid x \sim D^R \longrightarrow F(g) \mid g \sim N(0,1)^R + O_\tau(1)$.

Proof Scheme for Soundness: We have H (hypercube) and S_α (sphere graph). We have $F: \{\pm 1\}^R \to \{\pm 1\}$ s.t. $\max_i Inf_i(F) < \tau$ (far from dictators). By Fourier analysis, F is a polynomial. Then $\Pr_{x \sim y}[F(x) \neq F(y)] = \Pr_{g \sim h}[F(g) \neq F(h)]$ ($x \sim y$: ↳ booleans, $\mathbb{E}[\frac{1-F(x)F(y)}{2}]$; $g \sim h$: ↳ $\sim N(0,1)$, $\mathbb{E}[\frac{1-F(g)F(h)}{2}]$). By Invariance Principle, cut ratio in F is preserved to S_α, i.e. $Val_H(F) \simeq \underbrace{Val_{S_\alpha}(F)}_{\leq OPT(S_\alpha)}$.

Invariance Principle (Formally): Let $x_1 \dots x_n \overset{\forall i}{\sim} \{\pm 1\}$, $g_1 \dots g_n \sim N(0,1)$. Let $P(z_1, \dots, z_n)$ be a polynomial of degree $\leq D$ and $\forall i, Inf_i(P) = \sum_{S \ni i} P_S^2 < \tau$. Fix any function $\psi: \mathbb{R} \to \mathbb{R}$ where $\|\psi^{(3)}\| < B$. Then $E[\psi(P(x_1 \dots x_n))] \simeq E[\psi(P(g_1 \dots g_n))] + error(D, \tau)$ where $error(D, \tau) \to 0$ as $\tau \to 0$.

Theorem) Under UGC, it is NP-Hard to approximate MAXCUT better than α.

Proof Idea: Reduce UGC → MAXCUT via dictatorship test. (proof omitted)

Construction of Dictatorship Test: Fix $G(V,E)$ ($|V|=n$), an instance of MAXCUT. Fix any k solutions $x_1,\dots,x_k \in \{0,1\}^n$ where $\forall x_i$, $\text{Val}_G(x_i) \geq C$. Let the dictatorship testing graph be H_G. Column wise merge $x_1,\dots,x_k$ to get n k-tuples $X^{(1)},\dots,X^{(n)}$. Take $V(H_G) := \{0,1\}^k$, and $\forall i,j \in G \Rightarrow (X^{(i)}, X^{(j)}) \in H_G$. Suppose $f(z) = z_i$ for some $z \in \{0,1\}^k$. Then $\text{Val}_{H_G}(f) = \text{Val}_G(x_i)$ since we only observe one specific row of each column. By assumption, $\text{Val}_G(x_i) \geq C$. (Completeness) Now suppose f is far from dictator.

What is the "best" dictatorship test (i.e. what is the smallest α (α_{MAXCUT})?

→ $\nexists$ dictatorship test with $\frac{\text{soundness}}{\text{completeness}} < \alpha$.

$\Rightarrow \nexists f: \{0,1\}^k \to \{0,1\}$ s.t. it is far from dictator, and $\forall$ graph G & $\forall x_1,\dots,x_k \in \{0,1\}^n$ s.t. $\text{Val}_G(x_i) \geq C$, $\text{val}_G(f(x_1,\dots,x_k)) \geq \alpha \cdot C$. (approximate polymorphism)

Def] α-Approximate Polymorphism: (distribution of) functions $F: \{0,1\}^k \to \{0,1\}$ s.t. $\forall G$, $\forall x_1,\dots,x_k$ where $\text{Val}_G(x_i) \geq C$, $E[\text{Val}_G(F(x_1,\dots,x_k))] \geq \alpha \cdot C$. ($F$: pointwise eval)

↳ Dictators are 1-approximate polymorphisms but not interesting.

→ α_{MAXCUT} = max F far from dictators but possible α-approx. poly.

Question: α-Approx. Poly. for MAXCUT $\Rightarrow$ α-Approx. Algorithm for MAXCUT?

→ for some $k \in \mathbb{N}$, say 10^6, $\exists$ function $F: \{\pm 1\}^k \to \{\pm 1\}$ far from dictators.

Say we have $G(V,E)$, $|V| = n$. If we have $X_1, \dots, X_k \in \{\pm 1\}^n$, and $val_G(X_i) \geq c$.

Applying $F(X_1, \dots, X_k)$ gives $val_G(F(X_1, \dots, X_k)) \geq \alpha \cdot c$. But we don't have x_is.

→ Solve SDP for MAXCUT(G) to get $v^{(1)}, \dots, v^{(n)} \in \mathbb{R}^n$. Pick $z \in N(0,1)^n$ and apply $v^{(1)} \cdot z, v^{(2)} \cdot z, \dots, v^{(n)} \cdot z =: g$. Then $E[(v^{(i)} \cdot z) \cdot (v^{(j)} \cdot z)] = v^{(i)} \cdot v^{(j)}$.

Now, use $g_1, \dots, g_k$ in the place of $x_1, \dots, x_k$. Apply $F(g_1, \dots, g_k)$.

→ Since F is far from dictator, it is low influence, the invariance principle says that $E_g[val_G(F(g_1, \dots, g_k))] \geq \alpha \cdot c$. //

⇒ Generally, an α-approx. poly. gives an α-approx. algorithm.

Theorem) Under UGC, $\nexists$ α-approx. poly. $\Leftrightarrow$ $\exists$ α-dictatorship tests

⇒ MAXCUT is hard to approximate to α.

Upshot: Polymorphisms "determine" the complexity of a CSP. This extends to decision and approximation problems.

Counting # of solutions? $\in \#P$. Which counting problems are in FP (poly time solvable), and which are in #P-Complete?

↳ A counting CSP $\in$ FP iff it admits a Maltzer Polymorphisms, where $f(a,b,b) = f(b,b,a) = a$ (such as for 3LIN, $f(x,y,z) = x-y+z$).

Else, it is in #P-Hard.

Two Lines of Work (related to polymorphisms):

1) Approximation under Perfect Completeness: given a fully satisfiable instance, approximate the value. [KMB]

↳ If 3LIN is $(1-\varepsilon)$ satisfiable, the problem is hard (to beat random), but if it is satisfiable, we can just use Gaussian Elimination!

↳ Betweenness CSP: For $P_1, \ldots, P_n$, find a permutation π subject to constraints of the form $\pi(P_i) < \pi(P_j) < \pi(P_k)$ (or flipped). Completeness gives $\alpha = 1/2$ instead of $1/3$!

2) Promise CSPs: ex) Given a 3-Colorable graph G, color it using as few colors as possible → ∃SDP-based algorithm to get a $n^{0.19\ldots}$ coloring. Also, it is NP-Hard to get a 6-coloring.

One emerging topic: Quantum CSPs (analog of PCP in quantum?)