In general case, $g(x) = \sum_{T \in \mathcal{I}} \hat{f}(T \cup S) \chi_T$.

5. For $a_1, a_2, a_3 \in \{-1, 1\}$, $NAE(a_1, a_2, a_3) = \frac{3}{4} - \frac{1}{4}a_1a_2 - \frac{1}{4}a_2a_3 - \frac{1}{4}a_1a_3$.

Hence, $Pr[NAE \text{ accepts } f] = \frac{3}{4} - \frac{1}{4}E[f(x)f(y)] - \frac{1}{4}E[f(y)f(z)] - \frac{1}{4}E[f(x)f(z)]$ where for each i,

(x_i, y_i, z_i) is chosen uniformly from $\{001, 010, 100, 011, 101, 110\}$. By symmetry,

$= \frac{3}{4} - \frac{3}{4}E[f(x)f(y)]$ where for each i (x_i, y_i) is chosen according to $\begin{cases} 00 & 1/6 \\ 01 & 1/3 \\ 10 & 1/3 \\ 11 & 1/6 \end{cases}$.

Equivalently, we choose x uniformly and choose $y = x + w$ where $w \sim \mu_{2/3}$. $\Rightarrow = \frac{3}{4} - \frac{3}{4}\langle f, T_{-\frac{1}{3}} f \rangle$.

$= \frac{3}{4} - \frac{3}{4}\sum_S (-\frac{1}{3})^{|S|} \hat{f}(S)^2$.

3a. Estimate $\hat{f}(S)$ to within $\pm\epsilon$ with conf. of $1-\delta$. $\hat{f}(S) = E[f(x)\chi_S(x)]$. By Chernoff,

number of samples is $O(\log 1/\delta / \epsilon^2)$.

6. $\sum_{S \in [n]} \hat{f}(S)^2 = E_x[f(x)^2]$, so again Chernoff.

7. $I(f) = \sum_S |S| \hat{f}(S)^2$, $4 \cdot Pr[f=1] \cdot Pr[f=-1] = 4 \cdot \frac{1}{2}(1 + \hat{f}(\phi)) \cdot \frac{1}{2}(1 - \hat{f}(\phi)) = 1 - \hat{f}(\phi)^2$.

So the inequality is equiv. to $\underbrace{\hat{f}(\phi)^2 + \sum |S| \cdot \hat{f}(S)^2}_{\geq \sum_S \hat{f}(S)^2 = 1} \geq 1$.

If we color the Hamming cube with two colors in a balanced way, then # of non-monochromatic

edges is $\geq 2^{n-1}$. In general $E(V', V-V') \geq \frac{4|V'| \cdot |V-V'|}{2 \cdot 2^n}$.

21.2.2008 Conj: $\forall \epsilon, \delta > 0$, $\exists k$ s.t. it is NP-hard given a label cover with unique constraints

to decide whether value $\geq 1-\epsilon$ or value $\leq \delta$.

Unique Games Hardness from any Tester

Thm: Suppose we have a tester $\underbrace{T}_{\text{with } q \text{ queries}}$ for $f: \{0,1\}^k \to [-1, 1]$ and $\epsilon, \delta, C, S \geq 0$

s.t. 1. All k dictatorships pass w.p. $\geq C$

2. If f is (ϵ, δ)-quasirandom then passes w.p. $\leq S$.

Then $\forall \eta > 0$, it is UG-hard given a CSP of the same "type" as our test,

decide whether value $\geq C - \eta$ or $\leq S + \eta$.

Example: We saw that Hast-Odd$_\delta$ accepts dictators w.p. $\geq 1-\delta$ and accepts

(ϵ^2, δ)-quasirandom w.p. $\leq \frac{1}{2} + \frac{1}{2}\epsilon$. Moreover, the tests were always of the form

"$x + y + z \stackrel{?}{=} 0/1 \pmod 2$". This gives that $\forall \eta > 0$, it is UG-hard to tell whether given

list of mod 2 equations with 3 variables per equation (MAX3LIN2) is $\geq 1 - \eta$ -

satis. or $< \frac{1}{2} + \eta$ satis.

(In fact, it is NP-hard)

Example: Consider the test that chooses $x \in \{0,1\}^k$ uniformly and $w \sim \mu_p$ and checks that $f(x) \neq f(x+w)$. The acceptance prob. is $\frac{1}{2} - \frac{1}{2}\langle f, T_{1-2p} f \rangle$.

Dictatorship is accepted w.p. p. It is known [MOO] that quasirandom functions are accepted w.p. $\leq \frac{1}{\pi}\operatorname{acos}(1-2p)$.

This implies hardness of $\dfrac{\frac{1}{\pi}\operatorname{acos}(1-2p)}{p}$ for MAX-CUT. For $p = 0.8466$ this gives UG-hardness of 0.8785. This exactly matches the Goemans-Williamson algorithm.

Proof (of Thm): Let $\lambda = \min(\eta/q, \eta\delta^2\varepsilon^3/18)$. By the UGC, $\exists k = k(\lambda)$ s.t. the following is NP-hard. Given a unique-LC on n vertices V and alphabet $[k]$, decide whether $\mathrm{val} \geq 1-\lambda$ or $\mathrm{val} \leq \lambda$. Moreover, it is known that we can assume that the graph is regular. We will reduce this problem to a CSP on $2^k \cdot n$ variables. We think of these variables as n Boolean functions $f_v : \{0,1\}^k \to \{-1,1\}$, one for each $v \in V$. We will complete the proof by proving two properties:

1. (Completeness) If $\exists$ a labeling $L: V \to [k]$ satisfying $\geq 1-\lambda$ of the constraints, then $\exists \{f_v\}_{v\in V}$ s.t. they satisfy $\geq c - q\lambda \geq c - \eta$ of the constraints of the CSP.
2. (Soundness) If $\forall$ labeling $L: V \to [k]$ at most $\lambda \leq \eta\varepsilon^2\delta^3/18$ fraction of constraints are satisfied, then $\forall \{f_v\}_{v\in V}$ the CSP is $< s+\eta$ satisfied.

Proof of 1: Assume $L: V \to [k]$ satisfies $\geq 1-\lambda$ of the constraints. Define f_v to be the $L(v)$-dictator function $(\chi_{\{L(v)\}})$. We now describe the reduction: Instead of describing the q-CSP, we specify a tester on the $\{f_v\}_{v\in V}$. For each $v \in V$, and one of its neighbors w, define $g_v^w : \{0,1\}^k \to \{1,-1\}$ by $g_v^w(x) = f_w(x_{\sigma_{w\to v}(1)}, x_{\sigma_{w\to v}(2)}, \dots, x_{\sigma_{w\to v}(k)})$. Our tester is the following:

1. Pick $v \in V$ uniformly.
2. Apply T to $\{g_v^w \mid w \text{ is a neighbor of } v\}$.

Notice, that this tester is of the same "type" as T.

The tester performs q queries, each to a random neighbor w of v. Since the graph is regular, each query satisfies that (v,w) is a uniformly distributed edge and hence is satisfied by L (i.e., $L(v) = \sigma_{w\to v}(L(w))$) w.p. $\geq 1-\lambda$. By union bound all q queries $w_1, w_2, \dots, w_q$ are such that (v, w_i) are satisfied by L w.p. $\geq 1 - q\lambda$.

In such a case, $g_v^{w_1} = g_v^{w_2} = \ldots = g_v^{w_q} = f_v = x_{L(v)}$, so T is applied to $x_{L(v)}$ and hence accepts w.p. $\geq C$. Overall, the acceptance prob. is $(1-q\lambda)C \geq C - q\lambda$, as required.

Proof of 2: We prove the contrapositive. Assume that we have $\{f_v\}_{v \in V}$ such that the tester passes w.p. $\geq S+\eta$. Our goal is to find a labeling $L: V \to [k]$ that satisfies $\geq \eta\delta^2\varepsilon^3/18$ of the constraints. For each $v \in V$, define a set of "candidate labels":
$\mathcal{L}(v) = \{i : \mathrm{Inf}_i^{(1-\delta)}(f_v) \geq \frac{\varepsilon}{2} \text{ or } \mathrm{Inf}_i^{(1-\delta)}(h_v) \geq \varepsilon\}$ where $h_v = \mathrm{avg}\{g_v^w : w \text{ neighbor of } v\}$.
By the bound we saw last time, $|\mathcal{L}(v)| \leq 3/\varepsilon\cdot\delta$. In the following we will see that for at least $\frac{\eta\cdot\varepsilon}{2}$ of the edges (v,w), we have that $\exists i$ s.t. $i \in \mathcal{L}(v)$ and $\sigma_{v\to w}(i) \in \mathcal{L}(w)$.
Define a labeling $L: V \to [k]$ by choosing for each $v \in V$ a label uniformly from $\mathcal{L}(v)$.
Since $\forall v, |\mathcal{L}(v)| \leq \frac{3}{\varepsilon\cdot\delta}$, the labeling L satisfies on expectation $\geq \frac{\eta\varepsilon}{2}\cdot\left(\frac{\varepsilon\delta}{3}\right)^2 =$
$= \frac{\eta\varepsilon^3\delta^2}{18}$ of the constraints, as required.

By an averaging argument, at least η of the v's are such that $\Pr[\text{test accepts} | v] > S$.
By our assumption on T, this implies that $\exists i$ s.t. $\varepsilon \leq \mathrm{Inf}_i^{(1-\delta)}(h_v)$ (so $i \in \mathcal{L}(v)$).

$$\varepsilon \leq \mathrm{Inf}_i^{(1-\delta)}(h_v) = \sum_{S \ni i}(1-\delta)^{|S|}\cdot \hat{h}_v(S)^2 = \sum_{S\ni i}(1-\delta)^{|S|}\cdot\left(\underset{w\sim v}{\mathbb{E}}[\hat{g}_v^w(S)]\right)^2 \underset{\text{Jensen, convexity}}{\leq} \sum_{S\ni i}(1-\delta)^{|S|}\cdot\underset{w\sim v}{\mathbb{E}}\left[\hat{g}_v^w(S)^2\right]$$

$$= \underset{w\sim v}{\mathbb{E}}\left[\sum_{S\ni i}(1-\delta)^{|S|}\cdot \hat{g}_v^w(S)^2\right] = \underset{w\sim v}{\mathbb{E}}\left[\sum_{T \ni \sigma_{v\to w}(i)}(1-\delta)^{|T|}\cdot \hat{f}_w(T)^2\right] = \underset{w\sim v}{\mathbb{E}}\left[\mathrm{Inf}_{\sigma_{v\to w}(i)}^{(1-\delta)}(f_w)\right].$$

By another averaging argument, we see that $\geq \varepsilon/2$ of v's neighbors w satisfy $\mathrm{Inf}_{\sigma_{v\to w}(i)}^{(1-\delta)}(f_w) \geq \frac{\varepsilon}{2}$, and hence $\sigma_{v\to w}(i) \in \mathcal{L}(w)$. In total, we found a $\frac{\eta\varepsilon}{2}$ fraction of the edges (v,w) such that $\exists i.\ i \in \mathcal{L}(v)$ and $\sigma_{v\to w}(i) \in \mathcal{L}(w)$, as required. $\blacksquare$

NP-hardness of MAX3LIN2

Assume we are given two Boolean functions, $f: \{0,1\}^k \to \{-1,1\}$, $g: \{0,1\}^\ell \to \{-1,1\}$, and a function $\pi: [k] \to [\ell]$. Consider the following test: Hastad$_{2,\delta}$

1. Choose $x \in \{0,1\}^\ell$, $y \in \{0,1\}^k$ uniformly, and $w \in \{0,1\}^k$ according to μ_δ.
2. Accept iff $g(x)\cdot f(\pi^{-1}(x)+y)\cdot f(y+w) = 1$, where $(\pi^{-1}(x))_i = x_{\pi(i)}$.

Lemma: $\Pr[\text{test accepts } f,g,\pi] = \frac{1}{2} + \frac{1}{2}\sum_{S\subseteq[k]}\hat{f}(S)^2\cdot\hat{g}(\pi_2(S))\cdot(1-2\delta)^{|S|}$, where $\forall j \in [\ell]$, $j \in \pi_2(S)$ iff $|\{i \in S : \pi(i) = j\}|$ is odd.

Proof: $\Pr[\text{test accepts } f,g,\pi] = \frac{1}{2} + \frac{1}{2}\underset{x,y,w}{\mathbb{E}}[g(x)\cdot f(\pi^{-1}(x)+y)\cdot f(y+w)] = \ldots$

$*\ \underset{x,y,w}{\mathbb{E}}[g(x)\cdot f(\pi^{-1}(x)+y)\cdot f(y+w)] = \underset{x,w}{\mathbb{E}}[g(x)\cdot f * f(\pi^{-1}(x)+w)] = \underset{x}{\mathbb{E}}[g(x)\cdot(T_{1-2\delta}(f*f))(\pi^{-1}(x))]$.

Define $h(x) = T_{1-2\delta}(f*f)(\pi^{-1}(x))$ for $x \in \{0,1\}^\ell$.

$$\ldots = E[g(x)h(x)] = \sum_{T\subseteq[\ell]} \hat{g}(T)\cdot\hat{h}(T) \underset{\text{Homework 1, Q2i}}{=} \sum_{T\subseteq[\ell]} \hat{g}(T)\cdot \sum_{\substack{S\subseteq[k]\\ \pi_2(S)=T}} \widehat{T_{1-2\delta}(f*f)}(S) =$$

$$= \sum_{S\subseteq[k]} \hat{g}(\pi_2(S))\cdot \widehat{T_{1-2\delta}(f*f)}(S) = \sum_{S\subseteq[k]} \hat{g}(\pi_2(S))\cdot \hat{f}(S)^2\cdot(1-2\delta)^{|S|}. \quad \blacksquare$$

Cor: If $f=\chi_{\{i\}}$ and $g=\chi_{\{j\}}$ with $\pi(i)=j$, then $\Pr[\text{accepts}]=1-\delta$.

Proof: Either from lemma or directly.

Def: For a function $f:\{0,1\}^k\to\{-1,1\}$ with $\hat{f}(\emptyset)=0$, define Q_f to be the distribution on $i\in[k]$ obtained by choosing $S\subseteq[k]$ with prob. $\hat{f}(S)^2$ and then choosing $i\in S$ uniformly. In other words, $\Pr[Q_f=i] = \sum_{S\ni i} \frac{\hat{f}(S)^2}{|S|}$.

Cor: Assume $\Pr[\text{test accepts } f,g,\pi] \geq \frac{1}{2}+\epsilon$ and moreover $\hat{f}(\emptyset)=\hat{g}(\emptyset)=0$.

Then, $\Pr_{\substack{i\sim Q_f\\ j\sim Q_g}}[\pi(i)=j] \geq \delta\cdot\epsilon^3$.

Proof: By the assumption, $2\epsilon \leq \sum_{S\subseteq[k]} \hat{f}(S)^2\cdot\hat{g}(\pi_2(S))\cdot(1-2\delta)^{|S|}$

Let $\mathcal{F}=\{S\subseteq[k] : \hat{g}(\pi_2(S))\cdot(1-2\delta)^{|S|} > \epsilon\}$. Then, $\sum_{S\notin\mathcal{F}} \hat{f}(S)^2\cdot\hat{g}(\pi_2(S))\cdot(1-2\delta)^{|S|} \leq \epsilon\cdot\sum\hat{f}(S)^2 = \epsilon$.

Therefore, $\epsilon \leq \sum_{S\in\mathcal{F}} \hat{f}(S)^2\cdot\underbrace{\hat{g}(\pi_2(S))\cdot(1-2\delta)^{|S|}}_{\leq 1} \leq \sum_{S\in\mathcal{F}}\hat{f}(S)^2$.

Now, $\Pr_{\substack{i\sim Q_f\\ j\sim Q_g}}[\pi(i)=j] \geq \sum_S \hat{f}(S)^2\cdot\hat{g}(\pi_2(S))^2\cdot\frac{1}{|S|}$, where the "$\frac{1}{|S|}$" appears because for each $j\in\pi_2(S)$ there is at least one $i\in S$ s.t. $\pi(i)=j$.

Using the ineq. $i\cdot(1-\delta)^i \leq \frac{1}{\delta}$, $\hat{g}(\pi_2(S))^2\cdot\frac{1}{|S|} \geq \delta\cdot\hat{g}(\pi_2(S))^2\cdot(1-\delta)^{|S|}$, which is at least $\delta\cdot\epsilon^2$ for all $S\in\mathcal{F}$. Hence, $\Pr[\ldots] \geq \delta\cdot\epsilon^2\sum_{S\in\mathcal{F}}\hat{f}(S)^2 \geq \delta\cdot\epsilon^3$. $\blacksquare$

Thm: $\forall \eta>0$ it is NP-hard to tell whether a given MAX3LIN2 instance has value $\geq 1-\eta$ or $\leq \frac{1}{2}+\eta$.