Today
- Parallel Repetition Theorem (Comp)
- Hardness of MAX3LIN

CSS.330.1 : PCPs
Limits of Approximation Algorithms
Lecture 08 (2023-3-24)
Instructor: Prahladh Harsha

Recap from last time
(Wanted to prove following lemma).

Lemma: $\forall S \subseteq [k]$ for $i \in [k] \setminus S$

$\omega(G) + \varepsilon_i \triangleq \Pr[W_i / W_S]$ where $W_S = \bigwedge_{j \in S} W_j$

then $\mathbb{E}_{i \notin S}[\varepsilon_i] \leq O\left(\sqrt{\frac{1}{k-|S|}\left(k\log(|\Sigma_1|\cdot|\Sigma_2|) + \log\left(\frac{1}{P_S}\right)\right)}\right)$

where $P_S = \Pr[W_S]$.

$S \subseteq [k]$ - "won" co-ordinates.

Common Random String. - 3 parts

(i) Answers to co-ordinates in S
$(A_i B_i)|_{i \in S} = (A B)|_S$

(ii) Questions in co-ordinates S
$(X_i Y_i)|_{i \in S} \cdot (X Y)|_S$

C

(iii) Random question in co-ordinates in $\bar{S}$

$\forall i \notin \bar{S}$

$$V_i \sim_u \{0,1\}$$

$$T_i \leftarrow \begin{cases} X_i & \text{if } V_i = 0 \\ Y_i & \text{if } V_i = 1 \end{cases}$$

$$(V\,T)/_{\bar{S}}$$

R

Note: Last lecture R included C.

— Properties of RC

(1) $$XY/_{X_i = x \wedge Y_i = y \wedge W_S \wedge (R,C) = (r,c)}$$

$$= X/_{X_i = x \wedge W_S \wedge (R,C) = (r,c)} \times Y/_{Y_i = y \wedge W_S \wedge (R,C) = (r,c)}$$

(proved last time)

(2) $$RC/_{X_i = x \wedge Y_i = y \wedge W_S}$$

$$\approx_{\varepsilon_i} RC/_{X_i = x \wedge W_S}$$

$$\approx_{\varepsilon_i} RC/_{Y_i = y \wedge W_S}$$

$E[\varepsilon_i]$ - small $i \notin S$

- To prove

$$\{X_i Y_i RC / W_S\} = \{X_i Y_i / W_S\}\{RC / X_i Y_i W_S\}$$

$$\approx_{\varepsilon_i} \{X_i Y_i\}\{RC / X_i Y_i W_S\}$$

$$\approx_{\delta_i} \{X_i Y_i\}\{RC / X_i W_S\}$$

Similarly

$$\{X_i Y_i RC / W_S\} = \{X_i Y_i / W_S\}\{RC / X_i Y_i W_S\}$$
$$\approx_{\varepsilon_i} \{X_i Y_i\}\{RC / X_i Y_i W_S\}$$
$$\approx_{\delta_i} \{X_i Y_i\}\{RC / Y_i W_S\}$$

Today: Will prove the purple approximation

Recall from last lecture

Proposition: $\underbrace{U_1 \ldots U_n}_{\text{product}}$ & E - event $\Pr[E] \geq 2^{-d}$

then $\mathbb{E}_i\left[\,|U_i|_E - U_i|\,\right] \leq \sqrt{\frac{d}{n}}$

Notation: (1) $\{U_i / E\} \approx_{\varepsilon_i} \{U_i\}$ where $\mathbb{E}_i[\varepsilon_i] \leq \sqrt{\frac{d}{n}}$

(2) $\{U_i / E\} \overset{i}{\approx}_{\sqrt{\frac{d}{n}}} \{U_i\}$

Extensions of above proposition.

Proposition': $\underbrace{U_1 \ldots U_n, R}_{\text{r.v}}$ & $\underset{\text{event}}{E}$, such that

(*) $\forall x \in \mathrm{Supp}(R)$, $U_1 \ldots U_n / R = x$ is a product dist

(**) $\forall x \in \mathrm{Supp}(R)$, $\Pr[E / R = x] \geq 2^{-d}$

then $\{R/E\}\{U_i/RE\} \overset{i}{\approx}_{\sqrt{\frac{d}{n}}} \{R/E\}\{U_i/R\}$

Proposition ": $\underbrace{U_1 \ldots U_n, R, C}_{r.v}$ & $\underbrace{E}_{\text{event.}}$ such that

(*) $\forall r \in Supp(R)$ $U_1 \ldots U_n / R=r$ is a product dist

(*) $\forall r \in Supp(R)$, $Pr[E/R=r] \geq 2^{-d}$

(*) $Supp(C/R=r \wedge E) \leq 2^h$ $\forall r \in Supp(R)$

then $\{RC/E\}\{U_i/RCE\} \overset{i}{\approx}_{\varepsilon} \{RC/E\}\{U_i/R\}$.

where $\varepsilon = \sqrt{\frac{d+h}{n}}$.

Back to proof of $\{X_i Y_i\}\{RC/W_S X_i\}$

$\approx$

$\{X_i Y_i\}\{RC/W_S X_i Y_i\}$

$C = (AB)|_S$; $R = (XY)|_{\bar{S}} (VT)|_{\bar{S}}$

Recall $T_i \leftarrow \begin{cases} X_i & \text{if } V_i = 0 \\ Y_i & \text{if } V_i = 1 \end{cases}$

$\bar{T}_i \leftarrow \begin{cases} Y_i & \text{if } V_i = 0 \\ X_i & \text{if } V_i = 1 \end{cases}$

Apply Proposition " to $E = W_S$; $R = R$

$C = C$; $U_i = \bar{T}_i$

we get

$$\{RC/W_S\}\{T_i/R\} \overset{i\notin S}{\underset{\delta}{\approx}} \{RC/W_S\}\{T_i/RW_S C\}$$

$$\delta \leq \sqrt{\frac{1}{k-|S|}\left(\frac{1}{\Pr[W_S]} + |S|(\log|\Sigma_1| + \log|\Sigma_2|)\right)}$$

We get $\{RC/W_S\}\{\bar{T}_i/RW_S C\} \overset{i\notin S}{\underset{\delta}{\approx}} \{RC/W_S\}\{\bar{T}_i/R\}$

$$= \{RC/W_S\}\{\bar{T}_i/V_i T_i\}$$

$V_i = 0$ occurs w/ prob $1/2$ in which case $\bar{T}_i = Y_i$ & $T_i = X_i$.

Hence, $\{RC/W_S\}\{Y_i/RW_S C\} \overset{i\notin S}{\underset{2\delta}{\approx}} \{RC/W_S\}\{Y_i/X_i\}$.

$$R^{-i} = R \setminus (V_i T_i)$$

$$\{X_i Y_i\}\{CR^{-i}/X_i Y_i W_S\} \overset{i\notin S}{\underset{\delta}{\approx}} \{X_i Y_i/W_S\}\{CR^{-i}/X_i Y_i W_S\}$$

$$= \{X_i Y_i CR^{-i}/W_S\}$$

$$= \{Y_i CR^{-i}/W_S\}\{X_i/Y_i W_S CR^{-i}\}$$

$$\overset{i\notin S}{\underset{2\delta}{\approx}} \{Y_i CR^{-i}/W_S\}\{X_i/Y_i\}$$

$$= \{Y_i/W_S\}\{CR^{-i}/Y_i W_S\}\{X_i/Y_i\}$$

$$\overset{c\ \&\ S}{\approx_{\delta}} \{Y_i\}\{CR^{-i}/Y_i W_S\}\{X_i/Y_i\}$$

$$= \{X_i Y_i\}\{CR^{-i}/Y_i W_S\}$$

This completes the proof of inequality & hence the lemma $\boxtimes$

Where are we.

PCP Theorem: $\exists\ \varepsilon \in (0,1)$, Σ_1 & Σ_2 - constant-sized alphabets s.t

$gap_{1,1-\varepsilon}-LC(\Sigma_1, \Sigma_2)$ is NP-hard (under poly time reductions)

(even when restricted to projective instances)

Applying Parallel Repetition to above theorem

PCP-Theorem + Parallel Repetition:

$\forall\ \delta \in (0,1)$, there exist alphabets Σ_1 & Σ_2 ($|\Sigma_i| = poly(\frac{1}{\delta})$) s.t

$gap_{1,\delta}-LC(\Sigma_1, \Sigma_2)$ is NP-hard under $n^{O(\log \frac{1}{\delta})}$-time reductions.

(even when restricted to projective games).

Part II : Hardness of MAX3LIN2

MAX3LIN2:

Instance: Φ $\begin{cases} n \text{ variables } x_1 \ldots x_n \\ m \text{ linear eqns.} \\ \quad \left.x_{i_1} \oplus x_{i_2} \oplus x_{i_3} = b_i \right\} m \text{ eqns} \end{cases}$

Goal: Find a Boolean assignment that satisfies as many eqns as possible

—

$gap_{c,s}$-3LIN2 : YES: Instances Φ s.t at least c fraction of constraints can be satisfied.

$\left(\frac{1}{2} \leq s < c < 1\right)$

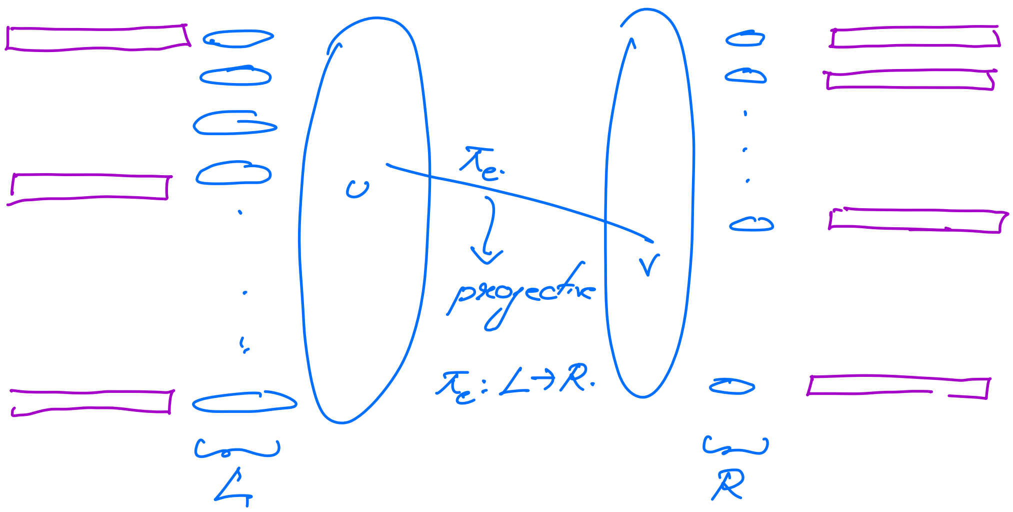
NO: Instances Φ s.t less than s fraction of constraints can be satisfied.

Theorem [Håstad] $\forall \varepsilon, \delta \in (0,1)$, $gap_{1-\varepsilon, \frac{1}{2}+\delta}$-3LIN2 is NP-hard

Reduction: We will show $\forall \varepsilon, \delta$. $\exists \mu$.

$$SAT \xrightarrow[\substack{\text{PCP Theorem} \\ + \\ \text{ll-repetition}}]{R_1} gap_{1,\mu}\text{-}LC(L,R) \xrightarrow[\text{Håstad.}]{R_2} gap_{1-\varepsilon, \frac{1}{2}+\delta}\text{-}3LIN2.$$

(L - left-hand side of
R - right hand side of



Proof: List of colors $\to$ Encoding of colors

This encoding must be testable

- using only tuple of 3 queries + predicates of form $x \oplus y \oplus z = b$.

- Codeword Test:
  Check if is a valid encoding

- Consistency Test
  Check if is an encoding of a coloring that satisfies the LC-projective predicates

Qn: What is a code that supports these properties?

Ans: Bellare-Goldreich-Sudan

Use the most wasteful code.
(aka long code)

Long Code

$LC : L \to \{0,1\}^{2^L}$

Think of 2^L as an index to the set of Boolean fns

$$\mathcal{F}_L \triangleq \{ f : L \to \{0,1\} \}$$

$$LC : L \to \{0,1\}^{\mathcal{F}_L}$$

$$a \mapsto (f(a))_{f \in \mathcal{F}_L}.$$

Qn: What is a candidate codeword test for LC?

$w \in \{0,1\}^{\mathcal{F}_L}$; $w : \mathcal{F}_L \to \{0,1\}$

w is a valid (long) code word if there exist an $a \in L$ s.t $w(f) = f(a)$, $\forall f \in \mathcal{F}_L$.

Or equivalently $w : \{0,1\}^L \to \{0,1\}$ is a (long) codeword if there $\exists a \in L$, $w(x) = x_a$ (dictator corresponding to a).

These are also called dictators

Obs: All dictators satisfy BLR Test
But so do all linear functions

Modify BLR-Test so that the prob of accepting linear fns w/ large support is small.

ε-perturbed BLR-Test^f: $(f : \{0,1\}^L \to \{0,1\})$

1. Pick $x, y \in_R \{0,1\}^L$
2. Pick $\eta \in \{0,1\}^L$ $\quad \eta_a \leftarrow \begin{cases} 0 & \text{w/p } 1-\varepsilon \\ 1 & \text{w/p } \varepsilon. \end{cases}$
3. Set $z \leftarrow x + y + \eta$
4. Accept if $f(x) + f(y) + f(z) = 0$.

Completeness: If f is a dictator (ie, $\exists a \in L$, $f(x) = x_a$)

then $\Pr_{x,y,\eta}[\varepsilon\text{-pert-BLR}^f \text{ accepts}] = 1-\varepsilon$.

Soundness analysis:

Convenient to work w/ $\{\pm 1\}$ instead of $\{0,1\}$

ie, $f : \{0,1\}^L \to \{\pm 1\}$

$$\Pr_{x,y,\eta}[\varepsilon\text{-perturbed-BLR}^f \text{ acc}] = \mathbb{E}_{x,y,\eta}\left[\frac{1 + f(x)f(y)f(z)}{2}\right]$$

Suppose $\Pr[\text{acc}] \geq \frac{1+\rho}{2}$.

$$\rho \leq \mathop{\mathbb{E}}_{x,y,\eta}\left[f(x)f(y)f(z)\right]$$

$$= \mathop{\mathbb{E}}_{x,y,\eta}\left[\sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \, \chi_\alpha(x)\chi_\beta(y)\chi_\gamma(x+y+\eta)\right]$$

$$= \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \mathop{\mathbb{E}}_{x}\left[\chi_{\alpha+\gamma}(x)\right] \cdot \mathop{\mathbb{E}}_{y}\left[\chi_{\beta+\gamma}(y)\right] \cdot \mathop{\mathbb{E}}_{\eta}\left[\chi_\gamma(\eta)\right]$$

$$= \sum_\alpha \hat{f}_\alpha^3 \cdot \mathop{\mathbb{E}}_{\eta}\left[\chi_\alpha(\eta)\right]$$

$$\mathop{\mathbb{E}}_{\eta}\left[\chi_\alpha(\eta)\right] = \mathop{\mathbb{E}}_{\eta}\left[(-1)^{\sum \alpha_i \eta_i}\right] = \mathop{\mathbb{E}}_{\eta}\left[\prod_i (-1)^{\alpha_i \eta_i}\right]$$

$$= \prod_i \mathop{\mathbb{E}}_{\eta_i}\left[(-1)^{\alpha_i \eta_i}\right] = \prod_{i:\alpha_i=1} \mathop{\mathbb{E}}_{\eta_i}\left[(-1)^{\eta_i}\right]$$

$$= \prod_{i:\alpha_i=1} (1-2\varepsilon) = (1-2\varepsilon)^{|\alpha|}$$

Plugging back

$$\rho \leq \sum_\alpha \hat{f}_\alpha^3 (1-2\varepsilon)^{|\alpha|} = \mathop{\mathbb{E}}_{\alpha \sim \hat{f}_\alpha^2}\left[\hat{f}_\alpha (1-2\varepsilon)^{|\alpha|}\right]$$

Hence, there exists an α s.t

$$\hat{f}_\alpha (1-2\varepsilon)^{|\alpha|} \geq \rho.$$

In particular $\hat{f}_\alpha \geq \rho$

$$(1-2\varepsilon)^{|\alpha|} \geq \rho \Rightarrow |\alpha| \leq O\left(\frac{1}{\varepsilon}\log\frac{1}{\rho}\right)$$

Soundness Claim: Suppose $\Pr[\varepsilon\text{-pert BLR}^f \text{ acc}]$ is at least $\frac{1+\rho}{2}$, then $\exists$ an $\alpha \in \{0,1\}^L$ s.t

(*) $\hat{f}_\alpha \geq \rho$

(*) $|\alpha| \leq O\left(\frac{1}{\varepsilon}\log\frac{1}{\rho}\right)$ (ie, support(α) is small)

Håstad's 3-bit PCP:

Redn from $\text{gap}_{1,\mu}\text{-}(L,C)$ to $\text{gap}_{1-\varepsilon,\frac{1}{2}+\delta}\text{-3LIN2}$.

Håstad's 3-bit PCP

Input: Label Cover Instance

$\Phi = (G = (U,V,E), L, R, \Pi = \{\pi_e : L \to R \mid e \in E\})$

Proof: $f_u : \{0,1\}^L \to \{\pm 1\}$, $\forall u \in U$

$f_v : \{0,1\}^R \to \{\pm 1\}$, $\forall v \in V$.

PCP

1. Pick $(u,v) \leftarrow_R E$ & π_e - be the corresponding projection.

2. Do the following for the triple (f_u, f_v, π_e) (for ease of notation f, g, π)

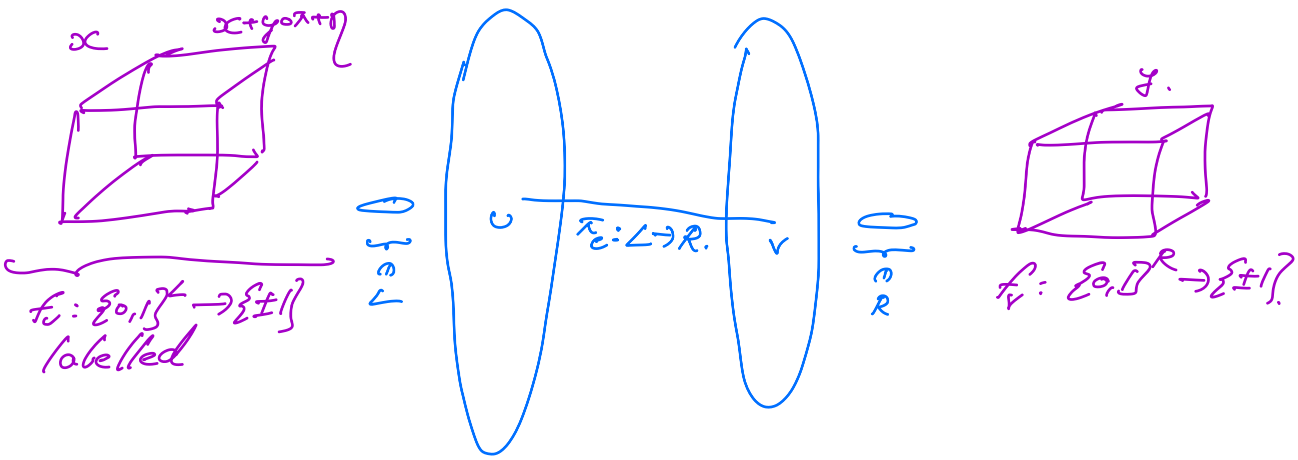
(a) Pick $x \in_R \{0,1\}^L$, $y \in_R \{0,1\}^R$

(b) Pick $\eta \in \{0,1\}^L$ s.t.

$$\eta_i \leftarrow \begin{cases} 0 & \text{w/ } 1-\varepsilon \\ 1 & \text{w } \varepsilon. \end{cases}$$

(c) $z \leftarrow x + y\circ\pi + \eta$

(d) Accept if $f(x)\cdot g(y)\cdot f(z) = 1$



Given $y \in \{0,1\}^R$ & $\pi: L \to R$

$y\circ\pi \in \{0,1\}^L$

$(y\circ\pi)_a = y_{\pi(a)}$.

Completeness: If Φ is an YES-instance of $\text{gap}_{1,\mu}$-LC & furthermore if $A: U \to L$ & $B: V \to R$ were the colorings that witnessed Φ is an YES-instance then

$$f_u = LC(A(u)) \quad , \forall u \in U$$
$$f_v = LC(B(v)) \quad , \forall v \in V.$$

$$\Pr_{(u,v)\, x,y,\eta}\left[\text{Håstad 3-bit PCP}^{\{f_u, f_v\}} \text{ acc}\right] = 1-\varepsilon.$$

$\overline{I}$ssue: Derailed by the all 1's function.

Folding: Valid longcodewords $f : \{0,1\}^L \to \{\pm 1\}$ satisfy

$$f(\bar{x}) = -f(x)$$

Assume: table is folded (ie, $f(\bar{x}) = -f(x)$)

Claim: $\forall \alpha$, $|\alpha|$ - even $\Rightarrow$ f is folded $\hat{f}_\alpha = 0$.

Pf: $\hat{f}_\alpha = \mathbb{E}[f(x) \chi_\alpha(x)] = -\mathbb{E}[f(\bar{x}) \chi_\alpha(x)]$

$$= -\mathbb{E}[f(\bar{x}) \chi_\alpha(\bar{x})]$$

$$= 0. \quad \boxtimes$$

Soundness Analysis.

Assume $\Pr_{(u,v),(x,y,\eta)}[\text{Håstad 3-bit PCP acc}] \geq \frac{1+\delta}{2}$.

Then,

$$\delta \leq \mathbb{E}_{(u,v)} \mathbb{E}_{x,y,\eta} \left[f_u(x) f_v(y) f_u(x + y \circ \pi + \eta) \right]$$

For at least a $\delta/2$-fraction of edges

$$\delta/2 \le \mathop{\mathbb{E}}_{x,y,\eta}\left[f_u(x)\, f_v(y)\, f_u(x+y\circ\pi+\eta)\right]$$

Fix such an edge (u,v)

$$\delta/2 \le \mathop{\mathbb{E}}_{x,y,\eta}\left[f(x)\,g(y)\,f(x+y\circ\pi+\eta)\right]$$

$$= \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha\, \hat{g}_\beta\, \hat{f}_\gamma \mathop{\mathbb{E}}_x\left[\chi_\alpha(x)\,\chi_\gamma(x)\right] \mathop{\mathbb{E}}_y\left[\chi_\beta(y)\cdot\chi_\gamma(y\circ\pi)\right] \mathop{\mathbb{E}}_\eta\left[\chi_\gamma(\eta)\right]$$

$$= \sum_{\alpha,\beta} \hat{f}_\alpha^2\, \hat{g}_\beta \mathop{\mathbb{E}}_y\left[\chi_\beta(y)\,\chi_\gamma(y\circ\pi)\right]\cdot(1-2\varepsilon)^{|\alpha|}$$

$$\chi_\gamma(y\circ\pi) = (-1)^{\sum_{i\in L}\gamma_i\cdot(y\circ\pi)_i} = (-1)^{\sum_{i\in L}\gamma_i\cdot y_{\pi(i)}} = (-1)^{\sum_{j\in R} y_j\cdot\sum_{i\in L:\pi(i)=j}\gamma_i} = (-1)^{\sum y_j\cdot\pi_2(\gamma)_j} = \chi_{\pi_2(\gamma)}(y)$$

Define $(\pi_2(\gamma))_j = \sum_{i:\pi(i)=j}\gamma_i$

For all the good edges

$$\delta/2 \le \sum_\alpha \hat{f}_\alpha^2\, \hat{g}_{\pi_2(\alpha)}\,(1-2\varepsilon)^{|\alpha|}.$$

Decoding a Labeling from the Proofs

Given $f_u : \{0,1\}^L \to \{\pm 1\}$, $\forall u \in U$
$f_v : \{0,1\}^R \to \{\pm 1\}$ $\forall v \in V$ } folded.

Define Randomized labelling $A : U \to L$
$B : V \to R$.

$A(u)$: 1. Pick $\alpha \leftarrow \{0,1\}^L$ w/ $\hat{f}_u^2(\alpha)$

2. Pick a random $a \leftarrow |\alpha|$.

$B(v)$: 1. Pick $\beta \leftarrow \{0,1\}^R$ w/ $\hat{f}_v^2(\beta)$

2. Pick a random $b \leftarrow |\beta|$

$$\Pr_{(u,v),A,B}\left[\pi_e(A(u)) = B(v)\right]$$

$$\geq \mathbb{E}_{u,v}\left[\sum_\alpha \hat{f}_u^2(\alpha) \sum_{\beta \subseteq \pi(\alpha)} \hat{f}_v^2(\beta) \frac{1}{|\alpha|}.\right]$$

For good-edges ($\delta/2$ - fraction)

$$\sum_\alpha \hat{f}_\alpha^2 \, \hat{g}_{\pi_2(\beta)} (1-2\varepsilon)^{|\alpha|} \geq \delta/2.$$

Need to relate $\displaystyle\sum_{\substack{\alpha \\ \beta \subseteq \pi(\alpha)}} \hat{f}_\alpha^2 \, \hat{g}_\beta^2 \frac{1}{|\alpha|}$

For a good edge

$$\sum_{\alpha} \sum_{\beta \subseteq \pi(\alpha)} \hat{f}_\alpha^2 \hat{g}_\beta^2 \frac{1}{|\alpha|} \geq \sum_{\alpha} \hat{f}_\alpha^2 \hat{g}_{\pi_2(\alpha)}^2 \frac{1}{|\alpha|}$$

$$= \left(\sum_{\alpha} \hat{f}_\alpha^2 \hat{g}_{\pi_2(\beta)}^2 \frac{1}{|\alpha|} \right) \left(\sum_{\alpha} \hat{f}_\alpha^2 \right)$$

$$\geq \left(\sum_{\alpha} \hat{f}_\alpha^2 \hat{g}_{\pi_2(\beta)} \frac{1}{\sqrt{|\alpha|}} \right)^2$$

$$\geq 4\varepsilon \left(\sum_{\alpha} \hat{f}_\alpha^2 \hat{g}_{\pi_2(\beta)} (1-2\varepsilon)^{|\alpha|} \right)^2$$

Since $\left(\frac{1}{\sqrt{2}} \geq \sqrt{4\varepsilon}(1-2\varepsilon)^2 \right)$ (via Taylor's thm).

$$\geq 4\varepsilon \left(\frac{\delta}{2} \right)^2 = \varepsilon\delta^2$$

$$\Pr_{u,v,A,B} \left[\pi(A(u)) = B(v) \right] \geq \frac{\delta^3 \varepsilon}{2}.$$

Choose $\mu = \delta^3\varepsilon/2$. ; Φ is not a NO-instance.

$\boxtimes$

Hence $\mathrm{gap}_{1-\varepsilon, \frac{1}{2}+\delta}$ -3LIN2 is NP-hard

---

MAX3SAT

$x \oplus y \oplus z = 1$

$x \vee y \vee z$

$\bar{x} \vee y \vee z$

$x \vee \bar{y} \vee \bar{z}$

$\bar{x} \vee y \vee \bar{z}$

$$1-\varepsilon \longrightarrow (1-\varepsilon)\cdot 1 + \varepsilon\cdot 0 = 1-\varepsilon$$

$$\frac{1}{2}+\delta \longrightarrow \left(\frac{1}{2}+\delta\right)\cdot 1 + \left(\frac{1}{2}-\delta\right)\cdot\frac{3}{4} = \frac{7}{8}+\frac{\delta}{4}$$

Corollary: $\mathrm{gap}_{1-\varepsilon,\frac{7}{8}+\delta}$-3SAT is NP-hard.