Tsinghua University

Discrete Optimistic

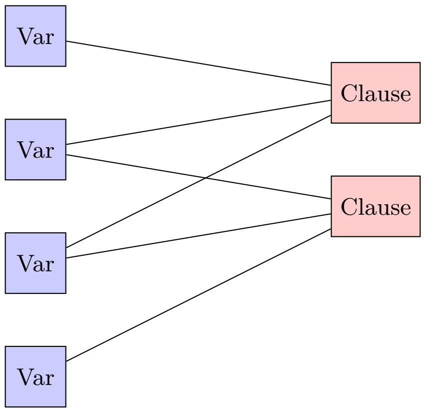
Fall 2024

Homework 5

Lin Zejin May 29, 2025

- **Collaborators:** I finish this homework by myself.

Problem 1. (a) Reduce from the instance of MAX-E3SAT-6.



Variables x_i have $\sigma(x_i) \in \{0,1\}$ and Clauses $c_i = x_{j_i}^1 \wedge x_{j_i}^2 \wedge x_{j_i}^3$ have $\sigma(c_i) \in [7]$ to represent the state of c_i.

Therefore, constraint is naturally induced.

In the instance of MAX-E3SAT, the radio of $|U|$ and $|V|$ is 2. So this is a regular Label-Cover Game for $K = 2, L = 7$ and $|V| = 2|U|$.

In the lecture we have proved that this is an instance of $\mathrm{MAX-LC}_{1,1-\epsilon}$ for some ϵ.

So $\mathrm{MAX-LC}_{1,1-\epsilon}$ is NP-Hard.

(b) We actually can construct another graph induced by (a).

We add $\bar{x}_i$ to the graph in (a) and add the induced constraints from c_i contains variable x_i to $\bar{x}_i$.

Here the Label-Cover Game is regular and symmetric.

Then for the $\mathrm{MAX-E3SAT-6}_{1,1-\epsilon}$ instance, the completeness is trivial.

Now we prove the soundness. That's because, if $\mathrm{OPT}_{\mathrm{MAX-E3SAT-6}} \le 1-\epsilon$, consider any $\sigma : U \to \{0,1\}, V \to [7]$. At least $(1-\epsilon)|V|$ clauses are not satisfied by $\sigma|_U$. For each clause, there exists at least one variable $x_i/\bar{x}_i$ such that do not satisfy the constraint.

So Verifier rejects with probability at least $(1-\epsilon)|V|/2|V| = (1-\epsilon)/2$. So the soundness property holds if we set $\epsilon' = \frac{1+\epsilon}{2}$.

So we prove that $\mathrm{GAP-LC}(K,L)_{1,1-\epsilon}$ is NP-Hard for some ϵ and K, L even if the graph is regular and symmetric.

By Raz' Paralled Repetition Theorem, we can reduce an instance of $\mathrm{GAP}-\mathrm{LC}(K,L)_{1,\delta}$ to the instance of $\mathrm{GAP}-\mathrm{LC}_{1,\exp(-\Omega(\frac{\delta^3 t}{\log t}))}$. Therefore, we finally prove that for any $\eta>0$, there exists K,L such that $\mathrm{GAP}-\mathrm{LP}(K,L)_{1,\eta}$ is NP-Hard.

Problem 2. (a) For a regular Label-Cover problem $G=(U,V,E)$ that every veritce in U matches k vertices in V, $|U|=|V|=n$, consider the k-uniform hypergraph $H=(V',E')$ where $V'=E$ and k-tuples are all $[(u,v_1),(u,v_2),\cdots,(u,v_k)]$ for $(u,v_i)\in E$. $[L']$ now represents the value of (u,v_i), *i.e.* $[L']=[L]\times[K]$. $[K]=[k+1]$.

The maps are defined as: For the labeling $\sigma:[V]\to[L]\times[K]$, $\sigma(u,v_i)=(l,k)$. If $\pi_{(u,v_i)}(k)=l$ is matching in Labek-Cover problem, then we let $\pi_e^i(\sigma(u,v_i))=k+1$. Otherwise, if (l,k) does not satisfy the constraint, then we let $\pi_e^i(\sigma(l,k))=i$.

So the constraint is weakly satisfied iff at least two edges in the k-tuples are satisfied in the constraint before. Also, the constraint is strongly satisfied iff all edges in the k-tuples are satisfied.

Completeness is trivial since if there is some label in the Label-Cover Game satisfy all constraint, then it can be naturally induced in the hypergraph.

Soundness is because: Assume $\mathrm{OPT}\geq\epsilon$ in k-ary-Consistent-Labeling problem. Then we choose all edges (u_i,v_j) that are satisfied in the Label-Cover Game, denoted as S. There are at least $2\epsilon n$ edges. Now we label each u_i,v_j one by one.

Since the graph G is regular, at most $2k-1$ edges in S have common vertice with an edge in S.

So each time we choose an arbitrary $e=(u,v)\in S$, label it with the label in k-ary-Consistent-Labeling and then we remove those edges in S who intersects with e.

In the end, for those vertices that have not been labeled yet, label it randomly.

Then at least $\frac{2\epsilon n}{2k}=\frac{\epsilon}{k}n$ edges are satisfied in Label-Cover-Game.

Therefore $\mathrm{OPT}\geq\frac{\epsilon}{k}$ for Label-Cover Game.

As a result, if $\mathrm{OPT}\leq\eta$ in Label-Cover Game, then $\mathrm{OPT}\leq k\eta$ in k-ary-Consistent-Labeling problem.

Since $\mathrm{MAX}-\mathrm{LC}_{1,\eta}$ is NP-Hard, to distingish instance with strong value 1 and weak value less than $k\eta$ k-ary-Consistent-Labeling problem is NP-Hard $\forall\eta>0$.

Here we end the proof.

(b)

In (a), we actually prove the l-regular case: *i.e.* each vertex u appears in l edges in the hypergraph.

For any hyperedge $e=(u_1,\cdots,u_k)$, let $T_e=\{1,2,\cdots,k\}^K$.

The unverse is $W=\cup_{e\in E}T_e\times\{e\}$, and set

$$S_{u,\alpha}=\bigcup_{u\in e,e\in E}\{x\in T_e:x_{\pi_i(\alpha)}=i\text{ where }e_i=u\}$$

Now we choose $|V|$ sets to cover this universe.

Completeness: If there is a instance with strong value 1, of course the max-coverage is 1.

Soundness: To prove if the instance has weak value less than δ, then the max-coverage is less than $\epsilon + 1 - (1 - \frac{1}{k})^k$. Suffices to prove that for each $\epsilon > 0$, $\exists \delta > 0$, if the max-coverage value is larger than $\epsilon + 1 - (1 - \frac{1}{k})^k$, then one can decode σ such that $\mathrm{Val} \geq \delta$.

Let $\mathrm{Sugg}(u) = \{\alpha : S_{u,\alpha} \text{ is chosen}\}$.

Claim 0.1.

$$\mathop{\mathbb{E}}_{e \in E} \sum_{u \in e} |\mathrm{Sugg}(u)| = \frac{1}{|E|} \sum_{e \in E} \sum_{u \in e} |\mathrm{Sugg}(u)| = \frac{l}{|E|} \sum_{u} |\mathrm{Sugg}(u)| = k$$

Here we use the l-regularity of the hypergraph, and $l|V| = k|E|$.

Denote

$$E_0 = \{e \in E : \exists u, v \in e, u \neq v, \pi_{i(u)}(\mathrm{Sugg}(u)) \cap \pi_{i(v)}(\mathrm{Sugg}(v)) \neq \emptyset, \text{ where } i(u), i(v) \text{ is the position of } u, v \text{ in } e\}$$

$$\tau = \mathop{\mathbb{E}}_{e \in E_0} \sum_{u \in e} |\mathrm{Sugg}(u)|, \gamma = \frac{|E_0|}{|E|}$$

For each u, if $\mathrm{Sugg}(u) \neq \emptyset$, then uniformly choose $\sigma(u) \sim \mathrm{Sugg}(u)$, else if $\mathrm{Sugg}(u) = \emptyset$, choose $\sigma(u)$ arbitary.

Claim 0.2. *At least $\frac{k-1}{k}$ of edges in $|E_0|$ has the property that*

$$\Pr[\textit{weakly satisfied}] \geq \frac{1}{k^2 \tau^2}$$

Proof. First, at least $\frac{k-1}{k}$ of edges in $|E_0|$ has the property that

$$\max_{u \in e} |\mathrm{Sugg}(u)| \leq k\tau$$

For $e \in E_0$, $\exists u, v$ such that

$u, v \in e, u \neq v, \pi_{i(u)}(\mathrm{Sugg}(u)) \cap \pi_{i(v)}(\mathrm{Sugg}(v)) \neq \emptyset,$ where $i(u), i(v)$ is the position of u, v in e. Then

$$\Pr[\pi_{i(u)}(\sigma(u)) = \pi_{i(v)}(\sigma(v))] \geq \frac{1}{|\mathrm{Sugg}(u)| \cdot |\mathrm{Sugg}(v)|} \geq \frac{1}{k^2 \tau^2}$$

the last inequality holds if $\max_{u \in e} |\mathrm{Sugg}(u)| \leq k\tau$.

Therefore, at least $\frac{k-1}{k}$ of edges in $|E_0|$ has the property that

$$\Pr[\text{weakly satisfied}] \geq \Pr[\pi_{i(u)}(\sigma(u)) = \pi_{i(v)}(\sigma(v))] \geq \frac{1}{k^2 \tau^2}$$

$\square$

Claim 0.3. *For $e \notin E_0$, coverage of $T_e \times \{e\}$ will less than $1 - (1 - \dfrac{1}{k})^{\sum_{u \in e} |\mathrm{Sugg}(u)|}$.*

Moreover, the total coverage of $\bigcup_{e \in E \setminus E_0} T_e \times \{e\}$ *is less than*

$$1 - \mathop{\mathbb{E}}_{e \in E \setminus E_0} (1 - \frac{1}{k})^{\sum_{u \in e} |\mathrm{Sugg}(u)|} \leq 1 - (1 - \frac{1}{k})^{\mathbb{E}_{e \in E \setminus E_0} \sum_{u \in e} |\mathrm{Sugg}(u)|} = 1 - (1 - \frac{1}{k})^{\frac{k - \gamma\tau}{1 - \gamma}}$$

Proof.

$$\text{non-coverage} = (1 - \frac{1}{k})^{\sum_{i=1}^{k} |\pi_i(\mathrm{Sugg}(e_i))|} \geq (1 - \frac{1}{k})^{\sum_{u \in e} |\mathrm{Sugg}(u)|}$$

□

Then by this claim, we have

$$1 - (1 - \frac{1}{k})^k + \epsilon < \gamma + (1 - \gamma)(1 - (1 - \frac{1}{k})^{\frac{k - \gamma\tau}{1 - \gamma}}) = 1 - (1 - \gamma)(1 - \frac{1}{k})^{\frac{k - \gamma\tau}{1 - \gamma}} < 1 - (1 - \gamma)(1 - \frac{1}{k})^{\frac{k}{1 - \gamma}}$$

Then $\gamma > \epsilon$.

And by

$$(1 - \gamma)(1 - \frac{1}{k})^{\frac{k - \gamma\tau}{1 - \gamma}} < (1 - \frac{1}{k})^k - \epsilon$$

we have

$$\tau < \frac{1}{\gamma} \left[k - (1 - \gamma) \log_{1 - \frac{1}{k}} \frac{(1 - \frac{1}{k})^k - \epsilon}{1 - \gamma} \right] < \frac{2k}{\gamma} < \frac{2k}{\epsilon}$$

Then by claim 0.2, the expected value of weakly satisfied edges is at least

$$\gamma \cdot \frac{k - 1}{k} \cdot \frac{1}{k^2 \tau^2} > \frac{\epsilon^3}{4k^5}$$

Take $\delta = \frac{\epsilon^3}{4k^5}$, then we can decode σ such that $\mathrm{Val}(\sigma) \geq \delta$.

Here we end the proof

Problem 3. Consider all values $d(r, v) \pmod{\frac{1}{2}}$. They divide $[0, \frac{1}{2})$ into $|V| + 1$ pieces of interval.(including the interval $[v, v]$ if exists) If we choose θ in each interval, edges that will be removed are the same, so the cost is the same.

As a result, we can try θ in each interval and find the minimum cost. This will be less than 2OPT.

Problem 4. (a) If a connected component has diameter at most k in the $(10, 0.1, 1, 1)$-expandar G, we prove that it has at most 10^k vertices.

By induction, $k = 1$ is trivial. Assume $k - 1$ holds for it. Assume subgraph G' has the maxmimum number of vertices. There isn't any vertex in G' that has distance less than $k - 1$ with each vertex in G' and also connects with other vertex u outside. Otherwise, u can be added to g', which causes contradiction with the maximum property. Then for k, any vertex in the graph with diameter $k - 1$ has degree 10 so at most 10^k vertices are connected to the graph. Since any vertex beyond G' has distance larger than k with some vertices in G' as we proved before, the expanded graph has at most 10^k vertices.

So each connected component has at most $10^{1/2\log_{10} n} = n^{1/2}$ vertices in this problem. As n large enough, $n^{1/2} < 0.1n$. For those connected components $S_1, \cdots, S_k$, removed edges are

$$|\partial S_1 \cup \partial S_2 \cup \cdots \partial S_k| = \frac{1}{2}\sum_{t=1}^{k} |\partial S_t| \geq \frac{1.01}{2}\sum_{t=1}^{k} |S_t| > 0.5n$$

So we must have deleted $\Omega(n)$ edges.

Now we set the pair (s_i, t_i) to be all (u, v) where $u, v \in G$ and distance between u and v is k.

Then for any possible connected component in multicut, vertices u, v in it have distance is less than k.

For a $(10, 0.1, 1.1)$-expandar graph, by (a) we removed at least $\frac{1}{2}n$ if $k = \frac{1}{2}\log_{10} n$.

However, in LP case, we can set $x_e = \frac{1}{k}$ for any edge e. Then the cost will be

$$\frac{1}{k} \cdot |E| = \frac{5}{k}|V| = \frac{5n}{k}$$

So the integral gap is $\Omega(\log n)$.

Problem 5. (a)

$$\begin{aligned}
\mathbb{E}(\text{cut value}) &= \sum_{(i,j)\in E} \omega_{ij} \cdot \frac{\arccos\langle v_i, v_j\rangle}{\pi} \\
&= \sum_{(i,j)\in E} \omega_{ij} - \sum_{(i,j)\in E} \frac{\frac{\pi}{2} + \arcsin\langle v_i, v_j\rangle}{\pi} \\
&\geq \sum_{(i,j)\in E} \omega_{ij} - \beta \cdot \sum_{(i,j)\in E} \omega_{ij}\sqrt{\frac{1+\langle v_i, v_j\rangle}{2}} \\
&\geq \sum_{(i,j)\in E} \omega_{ij} - \beta\left(\sum_{(i,j)\in E} \omega_{ij}\frac{1+\langle v_i, v_j\rangle}{2}\right)^{1/2}\left(\sum_{(i,j)\in E} \omega_{ij}\right)^{1/2} \\
&= 1 - \beta(1-\text{SDP})^{1/2} \\
&\geq 1 - \beta(1-\text{OPT})^{1/2}
\end{aligned}$$

where $\beta = \sup_{\alpha\in(-1,1)} \frac{\frac{\pi}{2}+\arcsin\alpha}{\sqrt{1+\alpha}} < +\infty$ by L'hospital rule.

Therefore, it is a $1-\epsilon$ vs. $1-\beta\sqrt{\epsilon}$ search algorithm.

(b)

Similar to max-cut. If we set $\mathbb{F}_2 = \{\pm 1\}$, then

$$\frac{1 - bx_ix_j}{2} = \begin{cases} 1 & x_i \oplus x_j = b \\ 0 & x_i \oplus x_j \neq b \end{cases}$$

where $1 \oplus 1 = -1 \oplus -1 = -1, 1 \oplus -1 = 1 \oplus 1 = 1$.

So the problem is to maximize the objective

$$\sum_{(i,j)\in E} \omega_{ij}\frac{1 - b_{ij}x_ix_j}{2}$$

Similarly, we set the SDP relaxation:

$$\min \sum_{(i,j)\in E} \omega_{ij} \frac{1 - b_{ij}\langle v_i, v_j\rangle}{2}$$

conditioned on $\|v_i\| = 1$.

After finding a minimum, we design a randomize algorithm as follows:

Uniformly sample $\vec{r} \sim S^{n-1}$.

Set $x_i = \operatorname{sgn}\langle \vec{r}, \vec{v}_i\rangle$.

Then

$$\begin{aligned}
\mathbb{E}(\text{cut value}) &= \sum_{(i,j)\in E} \omega_{ij} \cdot \frac{\arccos b_{ij}\langle v_i, v_j\rangle}{\pi} \\
&= \sum_{(i,j)\in E} \omega_{ij} - \sum_{(i,j)\in E} \frac{\frac{\pi}{2} + \arcsin b_{ij}\langle v_i, v_j\rangle}{\pi} \\
&\geq \sum_{(i,j)\in E} \omega_{ij} - \beta \cdot \sum_{(i,j)\in E} \omega_{ij} \sqrt{\frac{1 + b_{ij}\langle v_i, v_j\rangle}{2}} \\
&\geq \sum_{(i,j)\in E} \omega_{ij} - \beta \left(\sum_{(i,j)\in E} \omega_{ij} \frac{1 + b_{ij}\langle v_i, v_j\rangle}{2} \right)^{1/2} \left(\sum_{(i,j)\in E} \omega_{ij} \right)^{1/2} \\
&= 1 - \beta(1 - \text{SDP})^{1/2} \\
&\geq 1 - \beta(1 - \text{OPT})^{1/2}
\end{aligned}$$

where $\beta = \sup_{\alpha\in(-1,1)} \frac{\frac{\pi}{2} + \arcsin\alpha}{\sqrt{1+\alpha}} < +\infty$ by L'hospital rule.

Therefore, it is a $1 - \epsilon$ vs. $1 - \beta\sqrt{\epsilon}$ search algorithm.

Problem 6. First, for any uniformly weighted graph $G = (V, E, \omega)$, one can add another $|V|^2$ empty vertices such that the new graph is sparse. Apply algorithm A to this new graph, we can obtain a cut with value αOPT_G in $f(|V|^2, \frac{|E|}{|V|^2}) \in \text{poly}(|V|^2)$. Call this algorithm A'.

Given input graph $G = (V, E, w)$ and $\epsilon > 0$, set $\eta = \epsilon/\alpha$ and $w_0 = \frac{\min_{e\in E} \omega(e)}{cn^2}$ small enough for constant $c > 8$.

For each $e \in E$, let $k_e = \lceil \frac{w_e}{w_0} \rceil$.

Construct H by sampling each edge of G independently with probability $k_e p = k_e \min(1, \frac{c\log n}{\eta^2})$, and setting sampled edge weights to w_0/p.

Execute A on H to obtain cut $S \subseteq V$.

Lemma 0.4. *For any cut $S \subseteq V$:*

$$(1 - \eta)val_G(S) \leq val_H(S) \leq (1 + \eta)val_G(S)$$

with high probability $\geq 1 - 1/n$.

Proof. Since

$$\mathbb{E}[\text{val}_H(S)] = \sum_{e\in S} \omega_0/p \cdot k_e p \geq \text{val}_G(S)$$

and $\chi_e \le \frac{\omega_0}{p}$ By Chernoff bound, we have

$$\Pr[|\mathrm{val}_H(S) - \mathrm{val}_G(S)| \ge \eta \mathrm{val}_G(S)] \le 2e^{-\frac{\eta^2 \cdot \mathrm{val}_G(S) p}{3\omega_0}}$$

Then

$$\begin{aligned}
\Pr[\exists S \subset V, |\mathrm{val}_H(S) - \mathrm{val}_G(S)| \ge \eta \mathrm{val}_G(S)] &\le \sum_{S \subset V} 2e^{-\frac{\eta^2 \cdot \mathrm{val}_G(S) p}{3\omega_0}} \\
&\le \sum_{S \subset V} 2e^{-\frac{\eta^2 \cdot k_S p}{3}} \text{ where } k_S \text{ is the number of edges cut by } S \\
&= 2\sum_t |\{S \subset V : k_S = t\}| e^{-\frac{\eta^2 t p}{3}} \\
&\le 2\sum_t \binom{|E|}{t} e^{-\frac{\eta^2 t p}{3}} \\
&\le 2\sum_t \left(\frac{m}{t}\right)^t e^{-\frac{\eta^2 t p}{3}} \\
&\le 2\sum_t \left(\frac{n^2}{t}\right)^t e^{-\frac{ct \log n}{3}} \\
&= 2\sum_{t=1}^{m} \left(\frac{n^2}{t} \cdot n^{-\frac{c}{3}}\right)^t
\end{aligned}$$

Since $t \le n^2$, take $\frac{c}{3} > 4$, we have the lemma holds with high probability $1 - \frac{1}{n}$ when c large enough.

□

Apply algorithm A' on H we obtain a cut with value $\mathrm{val}_H(S) \ge \alpha \mathrm{OPT}_H \ge \alpha(1-\eta)\mathrm{OPT}_G = (\alpha - \epsilon/2)\mathrm{OPT}_G$, with high probability $1 - 1/n$, whose runtime is $f(|V|, \frac{|E|}{|V|}) = \mathrm{poly}(|V|)$

Then $\mathrm{val}_G(S) \ge \mathrm{val}_H(S) - \omega_0 \cdot n^2 \ge (\alpha - \epsilon)\mathrm{OPT}_G$.

So we construct a $\alpha - \epsilon$-approximating randomized algorithm for max-cut in all graph.

Problem 7. hyperplane cuts $\frac{\alpha}{\pi}$ edges in G_d with angle α.

Then totally, hyperplane cuts

$$\frac{\int_{\arccos \rho^*}^{\pi} \sin^{d-2}\alpha \frac{\alpha}{\pi}\,\mathrm{d}\alpha}{\int_{\arccos \rho^*}^{\pi} \sin^{d-2}\alpha\,\mathrm{d}\alpha} < \frac{\int_{\arccos \rho^*}^{\pi} (\pi-\alpha)^{(d-2)/2} \frac{\alpha}{\pi}\,\mathrm{d}\alpha}{\int_{\arccos \rho^*}^{\pi} (\pi-\alpha)^{(d-2)/2}\,\mathrm{d}\alpha} = \frac{\arccos \rho^*}{\pi} + O(\frac{1}{d})$$

The first inequality is because $\dfrac{\sin\alpha}{\sin\beta} > \dfrac{\sqrt{\pi-\alpha}}{\sqrt{\pi-\beta}}$ if $\alpha < \beta$. Thus the probability of α in the left is less than the probability of β in the right if $\alpha < \beta$.

Problem 8. $f : \{\pm 1\}^n \to \mathbb{R}$ is a linear combination of function $f : \{\pm 1\}^n \to \{\pm 1\}$, which can be written in the form of linear combination of Fourier base functions:

$$f(x) = \sum_{S \subset [n]} \hat{f}(S) \chi_S(x)$$

where $\chi_S(x) = \prod_{i \in S} x_i$ is a multilinear polynomial.
So it is expressible as a multilinear polynomial.

The uniqueness is because, if there is some multilinear polynoimal g such that $g(x) = f(x), \forall x \in \{\pm 1\}$. Then using Parserval's Theorem we obtain that

$$\sum_{S \subset [n]} (f \hat{-} g)(S)^2 = \mathop{\mathbb{E}}_{\vec{x} \sim \{\pm 1\}^n} (f(\vec{x}) - g(\vec{x}))^2 = 0$$

So $f - g = \sum_{S \subset [n]} (f \hat{-} g)(S) \chi_S = 0$.

Problem 9.

$$\langle f, g \rangle = \left\langle \sum_{S \subset [n]} \hat{f}(S) \chi_S, \sum_{S \subset [n]} \hat{f}(S) \chi_S \right\rangle = \sum_{S_1, S_2 \subset [n]} \hat{f}(S_1) \hat{g}(S_2) \langle \chi_{S_1}, \chi_{S_2} \rangle = \sum_{S \subset [n]} \hat{f}(S) \hat{g}(S)$$

However, if we let $f = \chi_{\{x\}}, g = \chi_{\{y\}}, h = \chi_{\{x,y\}}$, then

$$\mathop{\mathbb{E}}_{\vec{t}} \chi_{\{x\}}(t) \chi_{\{y\}}(t) \chi_{\{x,y\}}(t) = \mathop{\mathbb{E}}_{\vec{t}} t_x^2 t_y^2$$

But

$$\hat{f}(S) \hat{g}(S) \hat{h}(S) \equiv 0, \forall S \subset [n]$$

due to they are Fourier basis.