Hardness Of Approximation

Lecture 6: MAXCUT approximation algorithm and UGC-hardness

Instructor: Irit Dinur and Amey Bhangale *Scribe: Tom Ferster*

1 Reminder from Last Week

Last week we finished the hardness proof of gap($\frac{1}{2}+\epsilon, 1-\epsilon$)-3LIN. Then, we defined the influence of a function in order to construct a k-bit dictatorship test.

Definition 1.1 (influence of bit i on f)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$. The influence of bit i on f is defined as:*

$$Inf_i(f) \coloneqq Pr[f(x) \neq f(x^{(i)})] = \sum_{i \in S} \widehat{f}(S)^2$$

where $x^{(i)}$ is the vector x where the i'th bit is flipped.

For any dictator $dict_i$ we have $inf_i(dict_i) = 1$ and $inf_{j \neq i}(dict_i) = 0$ but also other parity functions have high influences, so we defined the following:

Definition 1.2 (δ-influence of bit i on f)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$. The δ-influence of bit i on f is defined as:*

$$Inf_i^{1-\delta}(f) \coloneqq \sum_{i \in S} (1-\delta)^{|S|-1} \widehat{f}(S)^2$$

Then we showed that f is "far from dictator" if $\forall i \in [n] : Inf_i^{1-\delta}(f) \leq \epsilon$. We also talked about the unique games conjecture.

In this scribe I will use a slightly different definition of influence:

Definition 1.3 (k-degree influence of bit i on f)**.** *Let $f : \{0,1\}^n \to \mathbb{R}$. The k-degree influence of bit i on f is defined as:*

$$Inf_i^k(f) \coloneqq \sum_{i \in S,\, |S| \leq k} \widehat{f}(S)^2$$

2 MAXCUT Approximation Algorithm

In this section we will show the Goemans-Williamson algorithm [GW95] which approximates the MAXCUT problem with ratio 0.878567. For simplicity we will show the algorithm for unweighted graphs, but it applies to weighted graphs as well.

2.1 The MAXCUT problem

Definition 2.1. *Let $G = (V, E)$ be an unweighted graph. The problem is to find a set $S \subset V$ such that the number of edges between S and $\bar{S}$ is maximized. So the value that should be maximized is $\frac{|E(S,\bar{S})|}{|E|}$.*

Until 1995 the best approximation ratio known to the MAXCUT problem was $\frac{1}{2}$. The following are $\frac{1}{2} - $ *approximation* algorithms:

- Start with an arbitrary set S. As long as there is a vertex in S or in $\bar{S}$ such that moving it to the other set increases the cut, move the vertex to the other set. When the algorithm can't make any move it means that any $v \in V$ has more neighbors on the other set, so the size of the cut is at least $\frac{1}{2}|E|$, which means it must be a $\frac{1}{2} - \textit{approximation}$ algorithm.
- Take a random partition $(S, \bar{S})$. For any $e \in E$ the probability both its endpoints are on different sides (which means e is in the cut) is $\frac{1}{2}$, so in expectation the size of the cut is $\frac{1}{2}|E|$.

2.2 the Goemans-Williamson algorithm

The algorithm starts with solving a relaxation of the the following integer linear program (ILP), for which the objective function maximizes the cut in G.

maximize

$$\sum_{(i,j)\in E} \frac{1 - x_i x_j}{2}$$

subject to

$$\forall i \in V : x_i \in \{-1, 1\}$$

When we take $S = \{i \mid x_i = 1\}$, the size of the cut is equal to the objective function value because if $x_i = x_j$ then $\frac{1-x_i x_j}{2} = 0$, and if $x_i \neq x_j$ then $\frac{1-x_i x_j}{2} = 1$.

This integer linear program is NP-hard to solve, so we introduce the following relaxation. We substitute each variable x_i with a unit vector $v_i \in \mathbb{R}^n$ (when $n = |V|$). We get the following semi-definite program (SDP), which can be solved optimally using the ellipsoid algorithm (up to a negligible additive error) :

maximize

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

subject to

$$\forall i \in V : \|v_i\| = 1.$$

When all vectors belong to a set of two antipodal vectors $\{v, -v\}$, and we define $S = \{i \mid v_i = v\}$ (see figure 1) the objective function will represent the size of the cut defined by S, so we have

$$OPT_{SDP} \geq OPT_{ILP} = MAXCUT(G)$$

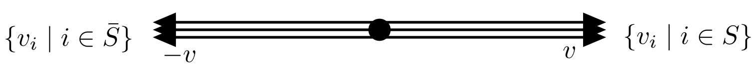

Figure 1: All vectors belong to $\{v, -v\}$.

Therefore, the solution to the SDP is an upper bound on the maximal cut size. We now have to round the solution in order to get a valid cut.

Let $v_i^* \in \mathbb{R}^n$ be an optimal solution to the SDP. Let r be a random unit vector in $\mathbb{R}^n$ sampled from a spherically symmetric distribution. The coordinates of r are i.i.d from $N(0, 1)$ and then the vector is normalized. Let H be the hyperplane tangent to r. We use it to divide the vertices to S and $\bar{S}$.

$$v_i^* \cdot r \geq 0 \Rightarrow i \in S$$

$$v_i^* \cdot r < 0 \Rightarrow i \in \bar{S}$$

The algorithm ends by outputting the partition $(S, \bar{S})$ as shown in figure 2.

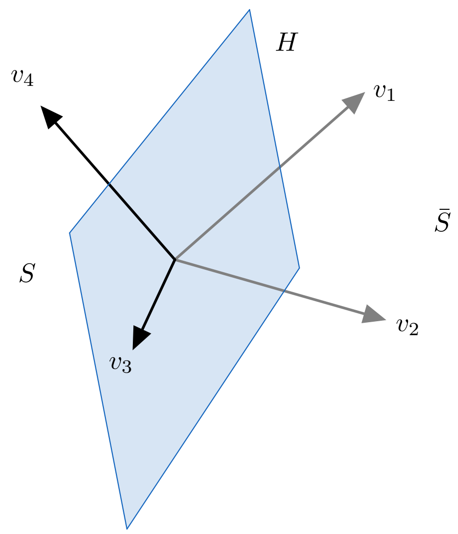

Figure 2: H divides the vectors to $S, \bar{S}$.

2.3 Analysis of the Algorithm

In order to find the approximation ratio of the algorithm we have to find the probability each edge is cut, because

$$\mathbb{E}(|E(S, \bar{S})|) = \sum_{e \in E} Pr(e \text{ } is \text{ } cut)$$

Let $(i, j) \in E$. Lets look at the plane spanned by v_i^*, v_j^* and denote r' the projection of r on this plane. since r is spherically symmetric distributed, so is its projection on any plane. The edge (i, j) is cut if both vectors are on different sides of a line perpendicular to r'. Denote the angle between v_i^*, v_j^* by $\theta_{i,j}$, so we get $Pr((i, j) \text{ } is \text{ } cut) = \frac{\theta_{i,j}}{\pi}$, as illustrated in figure 3.

Now, since $OPT_{SDP} \geq MAXCUT(G)$, if we divide the value of the rounded solution with the optimal value of the SDP, we will get a lower bound on the approximation ratio. Since $\langle v_i, v_j \rangle = cos(\theta_{i,j})$ we get:

$$\frac{rounded\, solution\, value}{OPT_{SDP}} = \frac{\sum\limits_{(i,j) \in E} \frac{\theta_{i,j}}{\pi}}{\sum\limits_{(i,j) \in E} \frac{1 - cos\theta_{i,j}}{2}} \geq \inf_{\theta \in [0,\pi]} \frac{\frac{\theta}{\pi}}{\frac{1 - cos\theta}{2}} \backsimeq 0.878567 \tag{2.1}$$

Therefore, the approximation ratio of the GW algorithm is $\alpha_{GW} := 0.878567$. In 2002 Feige and Schechtman [FS02] showed that this ratio is also the integrality gap of the given SDP, which means the randomized hyperplane rounding method is optimal.

3 Hardness of MAXCUT

Now we want to check if the algorithm presented above is optimal. It is only known to be NP-hard to approximate MAXCUT with ration better than 0.92, but [KKMO07] showed that if we assume that the UGC is true than the optimal ratio is indeed α_{GW}.

Definition 3.1 (Unique Games)**.** *Let $G = (U \cup V, E)$ be a bipartite graph and $L \in \mathbb{N}$. A unique game is a label cover instance for which each vertex has the label set $[L] = \{1, 2, ..., L\}$, and for each edge*

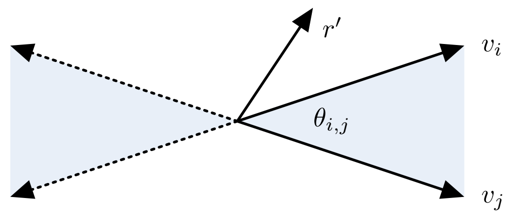

Figure 3: The polar angle of r is uniform in $[0, 2\pi)$, and the angles that will divide the vectors to different sets belong to two intervals of size $\theta_{i,j}$ each.

the allowed labeling $\pi_e : [L] \to [L]$ is a permutation. The optimization problem is to label each vertex $f : U \cup V \to [L]$ such that the labeling will satisfy the maximum amount of edges. An edge (i, j) is satisfied if $\pi_{(i,j)}(f(i)) = f(j)$.

Definition 3.2 (Unique Games Conjecture)**.** *$\forall \epsilon > 0, \exists L$ such that the problem gap$(\epsilon, 1-\epsilon)$-UG is NP-hard.*

It is unknown weather the UGC is true, however if we replace it with gap$(\epsilon, 1)$-UG it is easy to solve by guessing the labeling of a single vertex in every connected component, which determines its neighbors labeling and etc. If we get stuck we replace the value of the first vertex and try again.

We are going to prove the following theorem:

Theorem 3.3 (UGC-hardness of MAXCUT)**.** *Assuming UGC is true, it is NP-hard to approximate $MAXCAT$ better than α_{GW}.*

3.1 Gap Problems Reduction

In order to show the hardness of approximation result we first show that $\forall \rho \in (-1, 0], \epsilon' > 0$ $\exists \epsilon > 0$ such that there is a reduction from gap$(\epsilon, 1-\epsilon)$-UG to gap$\left(\frac{cos^{-1}\rho}{\pi} + \epsilon', (1-2\epsilon)(\frac{1-\rho}{2})\right)$-MAXCUT.

We start the reduction with a gap$(\epsilon, 1-\epsilon)$-UG instance $G = (U \cup V, E)$. For each vertex in V we introduce 2^L vertices to the MAXCUT instance. So the set of vertices is $V' \coloneqq V \times \{0, 1\}^L$.

Let $\rho \in (-1, 0]$. We define the edges of the graph using the following distribution D:

- Select $u \in U$ uniformly at random.
- Select $v, w \in N(u)$ independently uniformly at random.
- Select $x \in \{0, 1\}^L$ uniformly at random (represents a vertex from the cloud of v).
- Select $y \in \{0, 1\}^L$ (represents a vertex from the cloud of w) as follows: For any i, $Pr[y_i = x_i] = \frac{1+\rho}{2}$ and $Pr[y_i \neq x_i] = \frac{1-\rho}{2}$

We define for $x \in \{0, 1\}^L$ and permutation π the following:

$$x \circ \pi \in \{0, 1\}^L, \ \forall i : (x \circ \pi)_i = x_{\pi(i)}$$

Let π^v be the permutation from labels of v to labels of u, and π^w from w to u. So the edge we output according to the sampling is $(v \times (x \circ \pi^v), w \times (y \circ \pi^w))$. This process defines a distribution over the edges of G', so we can think of G' as a weighted graph where the weights are equal to the relevant probabilities. Figure 4 partially illustrates the reduction.

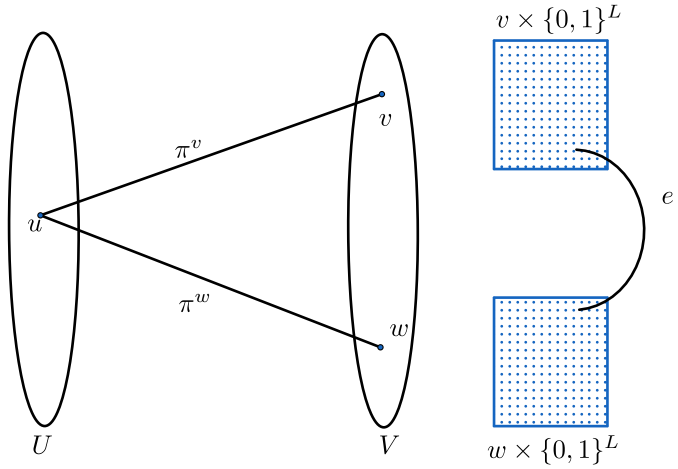

Figure 4: The reduction from a UG instance to a MAXCUT instance

3.1.1 Completeness

Let $l : U \cup V \to L$ be a labeling for the UG instance with value at least $1-\epsilon$. So the probability an edge is not satisfied is at most ϵ. We define the partition in the MAXCUT problem by assigning each vertex $v \times x$ in the cloud of a vertex $v \in V$ according to $f_v : \{0,1\}^L \to \{0,1\}$ which is defined as follows: $f_v(x) \coloneqq x_{l(v)}$.

The value of the cut is equal to

$$\sum_{e \in E(S,\bar{S})} weight(e) = \Pr_{e \sim D}(e\, is\, in\, the\, cut)$$

Let $e \sim D$. e is defined by u, v, w, x, y as defined above: $e = (v \times (x \circ \pi^v), w \times (y \circ \pi^w))$. suppose that both edges relevant to e of the label cover are satisfied by l. By the union bound this happens with probability at least $1-2\epsilon$. We look at the assignment to the endpoints of e.

$$f_v(x \circ \pi^v) = (x \circ \pi^v)_{l(v)} = x_{\pi^v(l(v))}$$

$$f_w(y \circ \pi^w) = (y \circ \pi^w)_{l(w)} = y_{\pi^w(l(w))}$$

Since u is a common neighbor to v, w we have $\pi^v(l(v)) = l(u) = \pi^w(l(w))$, So according to D the endpoints are on different sides with probability $\frac{1-\rho}{2}$.

In conclusion we get $val(cut) \geq (1-2\epsilon)(\frac{1-\rho}{2})$.

3.1.2 Soundness

Let $f_v : \{0,1\}^L \to \{+1,-1\}$ be assignments to the MAXCUT vertices with cut value at least $\frac{cos^{-1}\rho}{\pi} + \epsilon'$. We will use it to get a labeling for the UG with value at least ϵ (will be defined later). The size of the cut according to the definition of the reduction is as follows:

$$\begin{aligned}
val(cut) &= \mathop{\mathbb{E}}_{u\in U}\ \mathop{\mathbb{E}}_{v,w\in N(u)}\ \mathop{\mathbb{E}}_{x,y}\left[\frac{1-f_v(x\circ\pi^v)f_w(y\circ\pi^w)}{2}\right]\\
&= \mathop{\mathbb{E}}_{u\in U}\ \mathop{\mathbb{E}}_{x,y}\left[\frac{1-\mathop{\mathbb{E}}_{v\in N(u)}[f_v(x\circ\pi^v)]\mathop{\mathbb{E}}_{w\in N(u)}[f_w(y\circ\pi^w)]}{2}\right]\\
&(define\ h_u(x) := \mathop{\mathbb{E}}_{v\in N(u)}[f_v(x\circ\pi^v)])\\
&= \mathop{\mathbb{E}}_{u\in U}\ \mathop{\mathbb{E}}_{x,y}\left[\frac{1-h_u(x)h_u(y)}{2}\right]\\
&(define\ \mathbb{S}_\rho(g_u) := \mathop{\mathbb{E}}_{x,y}[h_u(x)h_u(y)])\\
&= \frac{1}{2}-\frac{1}{2}\mathop{\mathbb{E}}_{u\in U}[\mathbb{S}_\rho(g_u)]
\end{aligned}$$

$\mathbb{S}_\rho(g_u)$ defined above is the *stability* of g_u which measures the noise sensitivity of the function. This means the correlation of the function with itself after inserting noise as we did when we defined y as x with entries flipped with probability $\frac{1-\rho}{2}$. We will use now the Majority is Stablest Theorem.

Theorem 3.4 (Majority is Stablest)**.** *Let $\rho\in[0,1)$. For any $\epsilon>0$ there is a small enough $\delta>0$ such that if $f:\{-1,1\}^n\to[-1,1]$ satisfies:*

$$\mathbb{E}[f]=0$$

$$\forall i\in[n]:\ Inf_i(f)\le\delta$$

which means f is far from a dictator function, then

$$\mathbb{S}_\rho(f)\le 1-\frac{2}{\pi}arccos(\rho)+\epsilon$$

We are going to use a slightly different version of the theorem:

Theorem 3.5 (Majority is Stablest - different version)**.** *Let $\rho\in(-1,0]$. For any $\epsilon>0$ there is a small enough $\delta>0$ and large enough $k>0$ such that if $f:\{-1,1\}^n\to[-1,1]$ satisfies:*

$$\forall i\in[n]:\ Inf_i^k(f)\le\delta$$

which means f is far from a dictator function, then

$$\mathbb{S}_\rho(f)\ge 1-\frac{2}{\pi}arccos(\rho)-\epsilon$$

Assuming $val(cut)\ge\frac{cos^{-1}\rho}{\pi}+\epsilon'$ we get by Markov inequality that at least $\frac{\epsilon'}{2}$ fraction of $u\in U$ satisfy:

$$\mathbb{S}_\rho(g_u)\le 1-\frac{2}{\pi}arccos(\rho)-\epsilon'$$

We call such vertices U-good. By the Majority is Stablest theorem, for any U-good u there is a coordinate j that satisfies $Inf_j^k(g_u)\ge\delta$. We label u as $l(u)=j$, and the other not good vertices arbitrarily. We now get:

$$\delta\le\sum_{j\in S,|S|\le k}\widehat{g_u}(S)^2=\sum_{j\in S,|S|\le k}\mathop{\mathbb{E}}_{v\in N(u)}[\widehat{f_v}((\pi^v)^{-1}(S))]^2\le\sum_{j\in S,|S|\le k}\mathop{\mathbb{E}}_{v\in N(u)}[\widehat{f_v}((\pi^v)^{-1}(S))^2]=\mathop{\mathbb{E}}_{v\in N(u)}[Inf^k_{(\pi^v)^{-1}(j)}(f_v)]$$

From the equation above we conclude that for at least $\frac{\delta}{2}$ fraction of the neighbors v of u we have $Inf^k_{(\pi^v)^{-1}(j)}(f_v) \geq \frac{\delta}{2}$. We call those vertices V-good. In order to define labels for any $v \in V$ we define a set of candidates:

$$cand[v] = \{i \in [L] : Inf^k_i(f_v) \geq \frac{\delta}{2}\}$$

Then for any V-good v, $(\pi^v)^{-1}(j) \in cand[v]$. Moreover, since $\sum_{i \in [L]} Inf^k_i(f_v) \leq k$ we get that $|cand[v]| \leq \frac{2k}{\delta}$. Now we assign each $v \in V$ uniformly at random a label in $cand[v]$.

For any U-good u and V-good v, if the candidate that was chosen for v is $(\pi^v)^{-1}(l(u))$ then the edge (u, v) is satisfied, and therefore at least $\epsilon := (\frac{\epsilon'}{2})(\frac{\delta}{2})(\frac{2k}{\delta})$ fraction of the edges are satisfied. This completes the soundness proof.

3.2 Hardness of approximation ratio

In the previous section, we proved that assuming UGC, $\forall \rho \in (-1, 0]$,$\epsilon' > 0$,$\exists \epsilon > 0$ such that gap($\frac{cos^{-1}\rho}{\pi} + \epsilon', (1 - 2\epsilon)(\frac{1-\rho}{2})$)-MAXCUT is NP-hard. Therefore, we get a hardness of approximation factor of

$$\frac{\frac{cos^{-1}\rho}{\pi} + \epsilon'}{(1 - 2\epsilon)(\frac{1-\rho}{2})}$$

By slightly modifying ρ to move the completeness correction $(1 - 2\epsilon)$ into the soundness correction ϵ' we get $\forall \rho \in [-1, 0], \epsilon' > 0$ the factor:

$$\frac{\frac{cos^{-1}\rho}{\pi}}{\frac{1-\rho}{2}} + \epsilon'$$

Choosing $\rho = cos\theta$ for the θ that achieves α_{GW} we get a hardness factor of $\alpha_{GW} + \epsilon'$.

References

[FS02] Uriel Feige and Gideon Schechtman. On the optimality of the random hyperplane rounding technique for max cut. *Random Structures & Algorithms*, 20(3):403–440, 2002.

[GW95] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan ODonnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM Journal on Computing*, 37(1):319–357, 2007.