Advanced Approximation Algorithms **(CMU 18-854B, Spring 2008)**

Lecture 24: Hardness of approximation of *MAX-CUT*

April 10, 2008

Lecturer: Ryan O'Donnell *Scribe: S. Harsha Vardhan*

1 Known Results

- The 1 *vs* 1 decision problem is easy for *MAX-CUT* as the problem is equivalent to testing if the graph is bipartite.
- Goemans-Williamson's SDP based algorithm for *MAX-CUT* yields a $\frac{1}{2}(1-\rho)$ *vs* $(\arccos(\rho))/\pi$ approximation for all $\rho < -0.69$. Replacing ρ with $2\delta - 1$ shows that the $1-\delta$ *vs* $1-\Theta(\sqrt{\delta})$ problem is easy. If we think about *MAX-CUT* as *MIN-UNCUT*, this would mean that the δ *vs* $\Theta(\sqrt{\delta})$ problem is easy. From here on, we prefer to view *MAX-CUT* as *MIN-UNCUT*.
- Garg, Vazirani and Yannakakis[3] give an $O(\log n)$ factor approximation for the multi-cut problem which can be used to give and $O(\log n)$ factor approximation for the *MIN-UNCUT* problem, as any α factor approximation for the multi-cut problem translates in to an α factor approximation for the *MIN-UNCUT* problem.
- Charikar et. al. [1] give a $O(\sqrt{\log n})$ factor approximation for *MIN-UNCUT*.

We will now describe a 2-query long code test for testing dictator functions and use it in conjunction with the Unique Games Conjecture and the "Majority is Stablest" theorem to prove $\frac{1}{2} - \frac{1}{2}\rho - \eta$ *vs* $\frac{\arccos(\rho)}{\pi} + \eta$ hardness for the *MAX-CUT* problem.

2 2-query long code test and the MIS theorem

In order to test if the variables encoding the label for a vertex encode a valid label, we describe a 2-query long-code test for dictator functions. The test is as follows:

- pick $x \in \{0,1\}^k$ randomly.
- Construct y by copying x in to y and flipping each bit in y independently with probability $(1-\rho)/2$ (x and y are ρ-correlated).
- Check if $f(x)f(y) = -1$.

Completeness: If f is a dictator function, say $f(x) = x_a$ (thus encoding a valid label), then the test passes with probability:

$$Pr[f(x)f(y) = -1] = Pr[x_a \neq y_a] = (1-\rho)/2. \tag{1}$$

Soundness: Define the stability of a function f at noise ρ to be $\text{Stability}_\rho(f) = E_{x,y}[f(x)f(y)]$, where the expectation is over random x, and y constructed in the manner described earlier. The probability that an arbitrary f passes the test is

$$Pr[f \text{ passes test}] = E\left[\frac{1 - f(x)f(y)}{2}\right] \tag{2}$$

$$= \frac{1}{2} - \frac{1}{2}\text{Stability}_\rho(f). \tag{3}$$

$$\tag{4}$$

We would like this test to have the property that if the function that the long code encodes is *far* from a dictator function, then the probability that test succeeds is low. The "Majority is stablest" theorem, proved in [5] states exactly this:

Theorem 2.1. *For all* $\epsilon > 0, -1 < \rho < 0, \exists\tau(\rho,\epsilon) > 0$ *such that, if* $f : \{0,1\}^k \to \{0,1\}$ *has* $Inf_i < \tau$ *for all* $i \in [k]$*, then* $Pr[f \text{ passes}] < \frac{\arccos(\rho)}{\pi} + \epsilon$.

This theorem implies that a function that passes a test with probability greater than a certain value has at least one index j with high influence, and thus is close to the dictator function x_j.

Proof sketch for threshold function: Consider the function $f(x) = \text{sgn}(\sum_i a_i.x_i)$, with coefficients a_i scaled such that $\sum_i a_i^2 = 1$. It can be shown that if f has $\text{Inf}_i < \tau$ for all i, then all a_i are smaller than $O(\tau)$. The test described earlier consists of choosing a random k bit string x and a random $\frac{1-\rho}{2}$-biased k bit string μ and testing if $f(x)$ and $f(x\mu)$ have different signs.

Given a fixed μ, the test asks if a randomly chosen a has the property $\text{sgn}(a.x) \neq \text{sgn}(b.x)$, where $b = a\mu$. In the case where $a_i = 1/\sqrt{k}$ for all i, $a.x$ is just a sum of k independent variables drawn from the $\{-1,1\}$, the distribution of which is very similar to a Gaussian random variable. More generally, given that all a_is are small, the Central Limit Theorem implies that the distributions of $a.x$ and $b.x$ are *close* to that of a Gaussian. Since $a.g$ is also a Gaussian when g is a size k vector with independent Gaussian components, the distribution of $a.x$ ($b.x$) is *close* to that of $a.g$ ($b.g$). The test now simply asks if $a.g$ and $b.g$ have opposite signs, where g is a vector of Gaussian. From the analysis of Goemans-Williamson's SDP relaxation, we know that this probability is:

$$Pr[f \text{ passes}] \approx Pr[a.g \neq b.g] = \frac{\angle(a,b)}{\pi} = \frac{arccos(a.b)}{\pi} = \frac{\sum_i \mu_i}{\pi}. \tag{5}$$

Since $\sum_i \mu_i$ is close to ρ with high probability, f passes with probability approximately $arccos(\rho)$. $\square$.

The following stronger version of the MIS theorem states that functions which pass the test with high enough probability have at least one variables $j \in [k]$ that has a *large* low degree influence which is defined as: $\text{Inf}_j^{\leq C}(f) = \sum_{|S| \leq C, S \ni j} \hat{f}(S)^2$.

Theorem 2.2. *For all* $\epsilon > 0, \exists\tau > 0, C < \infty$ *such that if* $f : \{0,1\}^k \to \{0,1\}$ *and* $Pr[f$ *passes*$] > \frac{\arccos(\rho)}{\pi} + \epsilon$*, then* $Sugg(f)$ *defined as* $\{i \in [k] : Inf_i^{\leq C}(f)\}$ *is non-empty.*

Remark 2.3. Since

$$\sum_{j=1}^{k} \mathrm{Inf}_j^{\leq C}(f) = \sum_{j=1}^{k} \sum_{|S|\leq C, S\in j} \hat{f}(S)^2 \tag{6}$$

$$= \sum_{|S|\leq C} |S|\hat{f}(S)^2 \tag{7}$$

$$\leq C \sum_{|S|} |S|\hat{f}(S)^2 = C, \tag{8}$$

$$\tag{9}$$

$|Sugg(f)| \leq C/\tau$. This allows us to decode f in to a suggestion set of small size.

Remark 2.4. The MIS theorem is still true for $f : \{-1,1\}^k \to [-1,1]$, if we replace $Pr[f$ passes$]$ by $E[\frac{1}{2} - \frac{1}{2}f(x)f(y)]$.

3 An attempt that fails

4 Unique Games Conjecture

We first define the Unique Label Cover Problem.

Unique Label-Cover: An instance of the Unique Label-Cover problem is the same as the instance of the label cover problem except that the instance satisfies the following added restrictions:

- $K = L$ (the label set for both the partitions is the same) and
- The constraints on the edges, π_{vu}s are bijections.

The aim, as is in the case of the Label-Cover problem is to label the edges so as to maximize the fraction of the edges whose constraints are satisfied.

These restrictions make the 1 *vs* η decision problem for Unique Label-Cover is easy. In order to test if the given instance is 1 satisfiable, fix a label for any one of the vertices, say u. Since the edge constraints are bijections, the choice of label for u decides the labels of all neighbors of u. These in turn decide the value of labels for their neighbors and so on. If an edge constraint is violated some where, we keep running through all the choices of labels for u until we find one which induces a labeling that satisfies all the constraints. If one such is not found, we declare the instance to be not 1 satisfiable.

However, it is not known whether the $1-\eta$ vs η decision problem is hard for the Unique Label-Cover problem. Khot conjectured that this decision problem is hard:

Unique Games Conjecture (UGC) [4]: For arbitrarily small constants $\eta > 0$, there exists a constant $k = k(\eta)$ such that it is NP-hard to determine whether a unique Label Cover instance with the label sets of size k has optimum at least $1 - \eta$ or at most η.

This conjecture implies several other hardness results including the hardness of the $1-\delta$ *vs* $1-\delta^{1/2+o(1)}$ decision problem for MAX-2LIN. Since this conjecture proves hardness of problems whose hardness has not been proved otherwise, the validity of this conjecture is of great interest. One of the results that come close to deciding this conjecture [2] proves that the $1 - \epsilon$ *vs* $k^{\epsilon/(2-\epsilon)}$ and $1-\epsilon$ *vs* $1-O(\sqrt{\epsilon \log k})$ decision problems are easy for the Unique Label Cover problem with k labels.

5 Hardness of approximation of Max-2CSP

Theorem 5.1. *If we assume the validity of UGC, a c vs s dictator test translates into an associated $c - \eta$ vs $s + \zeta$ hardness for any CSP problem.*

Corollary: For all $\rho > 0, \eta > 0$, the $\frac{1}{2} - \frac{1}{2}\rho - \eta$ *vs* $(\arccos(\rho))/\pi + \eta$ decision problem is hard for MAX-CUT assuming UGC. This leads us to following corollary:

Corollary: Assuming UGC, $\forall$ constant $\delta > 0$, $1 - \delta$ *vs* $1 - \Theta(\sqrt{\delta})$ is hard for MAX-CUT, or equivalently, δ *vs* $\Theta(\sqrt{\delta})$ is hard for MIN-UNCUT.

This implies that a c-factor approximation is hard for MIN-UNCUT for all constants $c > 0$. Thus a constant factor approximation is also hard for MULTICUT and SPARSEST-CUT.

Proof. (of hardness of approximation of *MAX-CUT*) To start with, note that UGC holds *iff* UGC holds on regular graphs. This allows a regular bipartite graph instance for the Unique Label-Cover problem in our reduction.

- Randomly pick any $u \in U$ with equal probability.
- Randomly pick any two neighbors v, v' of u independently from the uniform distribution.
- Randomly pick ρ-correlated $x, y \in \{0,1\}^k$.
- Check if $g_v(x_{\pi_{vu}})g_{v'}(y_{\pi_{v'u}}) = -1]$.

Completeness: Suppose that the Unique Label-Cover instance has a label assignment $h : U \cup V \rightarrow K$ that satisfies $1-\eta$ fraction of the edges. Assign the $2^{|K|}$ variables at vertex v values that represent $g_v = h(v)$-th dictator function.

Since we have a chosen a regular graph for the Unique Label-Cover instance, the choices uv and uv' are uniformly random over the set of edges. Therefore, with probability at least $1 - 2\eta$, the constraints on both the edges uv and uv' are satisfied. Given that both the edge constraints are satisfied, we want to estimate the probability that the test succeeds. Since constraints π_{vu} and $\pi_{v'u}$ are both satisfied,

$$\pi_{vu}(h(v)) = h(u) = \pi_{v'u}(h(v')). \tag{10}$$

$$\text{Therefore,} \quad g_v(x_{\pi_{vu}}) = x_{\pi_{vu}}(h(v)) = x_{h(u)} \tag{11}$$

Similarly,

$$g_{v'}(y_{\pi_{v'u}}) = y_{h(u)}. \tag{12}$$

Since x and y are ρ-correlated, $x_{h(u)}$ and $y_{h(u)}$ are equal with probability $(1-\rho)/2$. Probability that the test passes is $(1-2\eta)(\frac{1-\rho}{2}) \geq \frac{1}{2} - \frac{1}{2}\rho - 2\eta$.

Soundness: To prove soundness property, we will prove it's contrapositive: given a set of long codes, f_u for $u \in U$ and g_v for $v \in V$ which pass the test with probability at least $(\arccos(\rho))/\pi + 2\epsilon$, it is possible to decode the long codes in to a label assignment, $\sigma : (U \cup V) \rightarrow K$ for the Unique Label-Cover instance that satisfies at least η fraction of edge constraints.

There exists at least an ϵ factor of fraction of *good* vertices $u \in U$ for which the test passes with probability at least $(\arccos(\rho))/\pi + \epsilon$ (If not, then the test will succeed will probability $<\leq 1.\epsilon + (\arccos(\rho))/\pi + \epsilon)(1-\epsilon/2) < (\arccos(\rho))/\pi + 2\epsilon$).

If u is good, then the probability of the test succeeding if we picked u in the first step is:

$$p_u = E_{v,v'}\left[E_{x,y}\left[\frac{1}{2} - \frac{1}{2} g_v(x_{\pi_{vu}}).g_{v'}(y_{\pi_{v'u}})\right]\right] \tag{13}$$

$$= E_{x,y}\left[\frac{1}{2} - \frac{1}{2} E_{v,v'}\left[g_v(x_{\pi_{vu}}).g_{v'}(y_{\pi_{v'u}})\right]\right] \tag{14}$$

Define $h_u : \{-1,1\}^{|K|} \rightarrow [-1,1]$ to be $h_u(z) = E_{w\sim u}[g_w(z_{\pi_{wu}})]$, the expected value of $g_u(z_{\pi_{wu}})$ over the choice of neighbors of u. The value of p_u is:

$$p_u = \frac{1}{2} - \frac{1}{2} E_{x,y}[h_u(x)h_u(y)] = \frac{1}{2} - \frac{1}{2}\text{Stability}_\rho(h_u) \tag{15}$$

Since u is a *good* node, $p_u > (\arccos(\rho))/\pi + \epsilon$. We can use the MIS theorem to deduce that h_u must have a coordinate j_u such that $\text{Inf}_{j_u}^{\leq C}(h_u) \geq \tau$. We use this coordinate as the label for *good* vertex u: $\sigma(u) = j$. Also, since $g_v = \sum_S \hat{g}_v(S)\chi_S$,

$$g_v \circ \pi = \sum_S \hat{g}_v(S)(\chi_S \circ \pi) \tag{16}$$

$$= \sum_T \hat{g}_v(\pi^{-1}(T))\chi_T. \tag{17}$$

Since $h_u(z) = E_{w\sim u}[g_u(z_{\pi_{wu}})]$, the Fourier coefficients of h_u are as follows:

$$h_u = \sum_T \left(E_{w\sim u}\left[\hat{g}_w(\pi_{w,u}^{-1}(T))\right]\right)\chi_T. \tag{18}$$

We now use these Fourier coefficients:

$$\tau \leq \text{Inf}_j^{\leq C}(h_u) \tag{19}$$
$$= \sum_{|S|\leq C, S\ni j} \hat{h}_u(S)^2 \tag{20}$$
$$= \sum_{|S|\leq C, S\ni j} \left(E_{w\sim u}\left[\hat{g}_w(\pi_{w,u}^{-1}(T))\right]\right)^2 \tag{21}$$
$$\leq \sum_{|S|\leq C, S\ni j} E_{w\sim u}\left[\hat{g}_w(\pi_{w,u}^{-1}(T))^2\right] \quad \text{Cauchy-Schwarz} \tag{22}$$
$$= E_{w\sim u}\left[\sum_{|S|\leq C, S\ni j} \hat{g}_w(\pi_{w,u}^{-1}(T))^2\right] \tag{23}$$
$$= E_{w\sim u}\left[\text{Inf}_{\pi_{w,u}^{-1}(j)}^{\leq C}(g_w)\right]. \tag{24}$$

We conclude that at least a τ fraction of the *good* node u's neighbors are *good-neighbors*, i.e., have $\text{Inf}_{\pi_{w,u}^{-1}(j)}^{\leq C} \geq \tau/2$. We can now decode g_w in to a set of suggestions consisting of indices with *large* low degree influences:

$$S_w = \{j : \text{Inf}_{\pi_{w,u}^{-1}(j)}^{\leq C} \geq \tau/2\}. \tag{25}$$

As we noted earlier, the size of set S_w is smaller than $2C/\tau$. We assign a *good-neighbor* a random label from it's small suggestion set. For all nodes that are neither *good* nor *good-neighbors*, we assign a random label.

The expected value of the fraction of satisfied constraints is at least $\epsilon.\frac{\tau}{2}\cdot\frac{1}{2C/\tau} = \epsilon\tau^2/4C$. If we use the value $\eta = \epsilon\tau^2/4C$ in the statement of the Unique Games conjecture ($c - \eta$ *vs* $s + \eta$ hardness), we have just shown how to decode a set a long codes that pass the test with probability greater than 2ϵ in to an assignment of labels that satisfies at least ζ fraction of edge constraints in the Unique Label instance. This completes the soundness part of the proof.

□

References

[1] A. Agarwal, M. Charikar, K. Makarychev, and Y. Makarychev. O(sqrt(log n)) approximation algorithms for min uncut, min 2cnf deletion, and directed cut problems. In *ACM Symposium of Theory of Computation*, pages 573–581, 2005.

[2] M. Charikar, K. Makarychev, and Y. Makarychev. Near-optimal algorithms for unique games. In *ACM Symposium of Theory of Computation*, pages 205–214, 2006.

[3] N. Garg, V. V. Vazirani, and M. Yannakakis. Approximate max-flow min-(multi)cut theorems and their application. In *ACM Symposium on Theory of Computing*, pages 698–707, 1993.

[4] S. Khot. On the power of unique 2-prover 1-round games. In *ACM Symposium of Theory of Computation*, pages 767–775, 2002.

[5] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low-influences - invariance and optimality. In *Proceedings of the Annual Symposium on Foundation of Compute Science*, pages 21–30, 2005.