CS 880: Advanced Complexity Theory	3/10/2008

Lecture 19: Inapproximability of Max-Cut

Instructor: Dieter van Melkebeek — Scribe: Matt Anderson

Last lecture we proved that MAX-3-SAT cannot be approximated to within any constant factor greater than $\frac{7}{8}$ unless P = NP. We reached this result by first showing that MAX-3-LIN cannot be approximated to within any factor greater than $\frac{1}{2}$ unless P = NP. Both of these results are tight in the sense that we have efficient deterministic algorithms that achieve the corresponding approximation factors.

Our approach for MAX-3-LIN was based on a formulation of the PCP Theorem in terms of constraint graph games (CGGs). We exploited a variation of the dictatorship test from Lecture 5 to test whether a given string satisfies any given property by probing three positions of the string's purported long code. By associating variables to each of the bits in the long code encodings of the labels in the constraint graph game, we were able to transform a CGG into a system of linear equations over three variables each, in a gap-preserving way.

In order to obtain the tight inapproximability result for MAX-3-LIN, we exploited special properties of the dictatorship from Lecture 5 and had to massage it. Today, we first give an alternate gap-preserving reduction to MAX-3-LIN which uses the dictatorship test in a blackbox fashion. The advantage is that we can obtain tight inapproximability results for other interesting problems like MAX-CUT by simply plugging in another dictatorship test. The disadvantage is that the reduction only works starting from unique games, and therefore only gives inapproximability results under the so-called unique games conjecture.

We first discuss the unique games conjecture and then develop the blackbox reduction.

1 The Unique Games Conjecture

Recall we have constraint graph games $G = (L, R, E, [l], [r], C)$, represented by a bipartite graph on the vertex sets L and R. $[l]$ and $[r]$ are the sets of values we can assign to nodes in the sets L and R, respectively. C denotes a collection of constraints c_e, one for each edge $e = (u, v) \in E$, where c_e puts a constraint on the values assigned to u and v. Note, previously we allowed the constraints to be of any function type. Today, we only consider unique games, where all constraints are of permutation type. This forces $l = r$ since all of the constraints are permutations, i.e., c_e is satisfied iff $\pi_e(j) = i$ for some permutation $\pi_e : [l] \to [r]$, where $i = \text{val}(u)$ denotes the label assigned to u, and similarly $j = \text{val}(v)$.

Remember that $\nu(G)$ denotes the maximum fraction of the constraints that can be satisfied by any assignment of values to the nodes in G. We also make use of the following quantity.

Definition 1.

$$\nu^*(G) = \textit{maximum fraction of } R \textit{ s.t. } (\forall v \in R)(\forall e = (u, v) \in E)\, c_e \textit{ is satisfied},$$

where the maximum is over all labelings of L and R.

In plain English, $\nu^*(G)$ is the maximum fraction, over all assignments, of nodes in R that have all of their constraints satisfied.

As we have seen previously, the PCP theorem implies that for every constant $\gamma > 0$ it is NP-hard to distinguish between games with $\nu(G) = 1$ and $\nu(G) \leq \gamma$ for some l and r (depending on γ). Notice that for unique games is it easy to decide whether $\nu(G) = 1$. The algorithm is as follows. Consider some node u and a possible label i for it. Since the constraints are permutations and must all be satisfied, there is no choice for the values on other nodes in the same connected component as u. The label i works for u if all constraints involving nodes int he connected component of u are satisfied. We try all possibilities for $i \in [l]$ and see whether there is one that works. $\nu(G) = 1$ iff the latter is the case for all connected components of G. Since l is finite, this procedure runs in polynomial time.

In an attempt to make our distinguishing task harder, we relax the condition $\nu(G) = 1$ to $\nu(G) \geq 1 - \gamma$. This leads to the Unique Games Conjecture.

Conjecture 1 (Unique Games Conjecture)**.** *For all $\gamma > 0$, there exists $l > 0$ such that it is* NP*-hard to distinguish between unique games with that l and $\nu(G) \geq 1 - \gamma$ or $\nu(G) \leq \gamma$.*

This is only one version of the Unique Games Conjecture. Another statement of the conjecture does not require it to be NP-hard to distinguish, rather that it should take more than polynomial time to distinguish between the two cases. If P = NP, the first statement implies the second one but possibly not the other way around. We will not bother about the difference, as we will always construct reductions from distinguishing the two cases to the problem at hand. We refer to this situation as the problem at hand being UG-hard.

We obtain another variation of the Unique Games Conjecture using the following lemma, which allows us to replace the condition $\nu(G) \geq 1 - \gamma$ by a condition of the form $\nu^*(G) \geq 1 - \gamma'$:

Lemma 1. *For all $\gamma' > 0$ and l there exists a $\gamma > 0$ such that one can transform in polynomial time a unique game G over $[l]$ into a unique game G' also over $[l]$ such that*

$$\nu(G) \geq 1 - \gamma \Rightarrow \nu^*(G') \geq 1 - \gamma',$$
$$\nu(G) \leq \gamma \Rightarrow \nu(G') \leq \gamma.$$

Proving this lemma is left to the reader. It is also possible to transform a unique game G into another unique game G' such that all of the vertices in R are regular (they have the same degree) while maintaining the other interesting properties. This also is left to the reader. Combining these two transforms we can restate the Unique Games Conjecture in modified form:

Conjecture 2 (Unique Games Conjecture')**.** *For all $\gamma > 0$ there exists $l > 0$ such that it is* NP*-hard to distinguish between unique games with that l and R regular and $\nu^*(G) \geq 1 - \gamma$ or $\nu(G) \leq \gamma$.*

2 The Dictatorship Test

We now start the development of our generic approach for deriving UG-hardness of approximation results, which uses a dictatorship as a blackbox. We first describe the test we used in the case of MAX-3-LIN, and then abstract the properties we need for the generic approach to apply.

2.1 The Test Used Before

In Lecture 5, we presented a dictatorship test T^g_α. Recall that before the final modification T^g_α accepted both dictators and the constant function with high probability. We made an addition to the test to make it reject the constant function with high probability. In fact, we mentioned two ways to do this: testing for balancedness or using self-reducibility. Last lecture we saw a third way to reject the constant function, namely to force $\hat{g}(\emptyset) = 0$ by considering only the odd part of the function and replacing $g(x)$ with $f(x) = \mathrm{E}_{b\in\{-1,1\}}[b \cdot f(b \cdot x)]$. Modifying T^g_α in this way:

- Dictatorship test T^g_α:
  - Pick $x, y \in \{-1,1\}^l$ and $b_x, b_y, b_z \in \{-1,1\}$ uniformly at random.
  - Pick $z \sim_\epsilon x$ with $\epsilon = \frac{1-\alpha}{2}$.
  - Accept iff $g(b_x x)g(b_y y)g(b_z z) = b_x b_y b_z$.

Using analysis similar to the original analysis of T^g_α one can show:

Lemma 2. $\Pr[T^g_\alpha \text{ } accepts] = \frac{1}{2} + \frac{1}{2}\sum_{S\subseteq[l],|S| \text{ } odd} \alpha^{|S|}(\hat{g}(S))^3$.

If g is not a Boolean function, but is instead an expectation over Boolean functions, Lemma 2 still holds. More formally, if $g = \mathrm{E}[f]$ over some distribution of $f : \{-1,1\}^l \to \{-1,1\}$ then Lemma 2 holds for g, where running the test on y means that whenever g needs to be evaluated on some point x, we take an independent sample f from g and return $f(x)$. By linearity of expectation, the original analysis carries through. Note that we have already implicitly used this for $f(x) = \mathrm{E}_{b\in\{-1,1\}}[b \cdot g(b \cdot x)]$, which is a distribution over two functions.

2.2 Required Properties

Our intent is to use T_α as a black box. The strength of the inapproximability result we obtain will depend on certain parameters of the test. Here are the properties we need for our analysis to follow through.

- Completeness: For all α and all dictators g: $\Pr[T^g_\alpha \text{ accepts}] \geq c_\alpha$.

  Claim 1. *The completeness property holds for the above test T_α with $c_\alpha = 1 - \epsilon$, where $\epsilon = \frac{1-\alpha}{2}$.*

  Proof. This follows since since we flip the dictator variable z_i with probability ϵ. □

- Soundness: For all α, and $\delta > 0$, there exists $\tau > 0$ and $d > 0$ such that if $\Pr[T^g_\alpha \text{ accepts}] \geq s_\alpha + \delta$, then there exists $j \in [l]$ with $I^{\leq d}_j(g) \geq \tau$, where $I^{\leq d}_j(g) = \sum_{S\ni i,|S|\leq d}(\hat{g}(S))^2$. This property should hold whenever g is the average of a distribution of Boolean functions on l variables.

  Claim 2. *The soundness property holds for the above test T_α with $\tau = \delta^2$, $d = \frac{\log 1/\delta}{\log 1/\alpha}$, and $s_\alpha = \frac{1}{2}$.*

  Proof. If $\Pr[T^g_\alpha \text{ accepts}] \geq \frac{1}{2} + \delta$, then using Lemma 2 and Parseval's equality we have

$$2\delta \leq \sum_{|S| \text{ odd}} \alpha^{|S|}(\hat{g}(S))^3 \leq \left(\max_{|S| \text{ odd}} (\alpha^{|S|}\hat{g}(S))\right) \sum_{|S| \text{ odd}} (\hat{g}(S))^2 \leq \max_{|S| \text{ odd}} (\alpha^{|S|}\hat{g}(S)).$$

This means there exists a Fourier coefficient over a set $S \neq \emptyset$ with $\hat{g}(S) \geq 2\delta$. Moreover, since $\hat{g}(S) \leq 1$, we have that $\alpha^{|S|} \geq 2\delta$, so

$$|S| \leq \frac{\log 1/\delta}{\log 1/\alpha} = d.$$

Since $|\alpha| \leq 1$, this also gives $(\hat{g}(S))^2 \geq \delta^2 = \tau$. If we pick a $j \in S$ the large coefficient must be included in the influence, which means that we have $I_j^{\leq d}(g) \geq \tau$. □

3 The Reduction

Let's start from a unique constraint graph game $G = (L, R, E, [l], [r], C)$, and consider a labeling of L. As in last lecture, let the function $f_u : \{-1,1\}^l \to \{-1,1\}$ represent the long code encoding of $\mathrm{val}(u)$ for $u \in L$. For $v \in R$ define $g_v(y)$ as:

$$g_v(y) = \mathrm{E}_{e=(u,v)\in E}[f_u(y \circ \pi_e)].$$

Recall that $y \circ \pi_e$ can be thought of as the permutation π_e applied to y, that is to say $(y \circ \pi_e)_i = y_{\pi_e(i)}$. It is useful to think of $f_u(y \circ \pi_e)$ as the opinion of u about what the function g_v should be, namely the dictator corresponding to the variable $\pi_e(\mathrm{val}(u))$.

The idea behind the reduction is the following: g_v is close to a dictator iff for most of the neighbors u of v, f_u is close to a dictator *and* for most of those dictator the dictating variable is the same. Thus, we simply run our dictatorship test T_α on g_v for a random $v \in R$. More formally, we consider the following test S_α^f:

- Test S_α^f:
  - Pick $v \in R$ at random.
  - Accept iff T_α on $g_v(y) = \mathrm{E}_{e=(u,v)\in E}[f_u(y \circ \pi_e)]$ accepts.

Note that the result of picking a random $v \in R$ at random and then a random neighbor u of v, results in a distribution of $e = (u, v)$ which is uniform over E. This is because R is regular.

3.1 Analysis of S_α^f

We proceed by analyzing the completeness and soundness of S_α^f. The completeness argument is fairly straightforward. Remember that $\nu^*(G)$ is the maximum fraction of $v \in R$ with all constraints satisfied. If $v \in R$ has all constraints satisfied, then v follows all the opinions of its neighbors in L. This implies that all of the f_u with $(u, v) \in E$ are the same dictator, so g_v in turn is a dictator. Thus, we have the following.

Claim 3. *There exists an f such that* $\Pr[S_\alpha^f$ *accepts*$] \geq c_\alpha \cdot \nu^*(G)$.

Proof. Consider a labeling of G realizing $\nu^*(G)$. For each $u \in L$, let f_u be the valid long code of the label of u. By the definition of $\nu^*(G)$, there exists a subset $R^* \subseteq R$ of relative size $\nu^*(G)$ with all constraints satisfied. For each of these $v \in R^*$, g_v is a dictator thus $\Pr[T_\alpha^{g_v}$ accepts$] \geq c_\alpha$. Therefore, $\Pr[S_\alpha^f$ accepts $] \geq \Pr[v \in R^*] \cdot \Pr[T_\alpha^{g_v}$ accepts $] \geq \nu^*(G) \cdot c_\alpha$. □

The soundness argument is somewhat more involved.

Claim 4. *If there exists an f such that* $\Pr[S_\alpha^f$ *accepts*$] \geq s_\alpha + 2\delta$*, then* $\nu(G) \geq \frac{\delta\tau^2}{4d}$*.*

Proof. By Markov's inequality, for a fraction at least δ of the $v \in R$, $\Pr[T_\alpha^{g_v}$ accepts$] \geq s_\alpha + \delta$. We call such v good. By the soundness property of T_α, for any good v there exists $j \in [l]$ such that $I_j^{\leq d}(g_v) \geq \tau$. Let us label v with such a value j.

We can upper bound the the influence of g_v in terms of the influences of the f_u's for the neighbors u of v as follows.

$$\begin{aligned} I_j^{\leq d}(g_v) &= \sum_{S \ni j, |S| \leq d} (\hat{g}_v(S))^2 && (1) \\ &= \sum_{S \ni j, |S| \leq d} \Big(\mathrm{E}_u[\hat{f}_u(\pi_e^{-1}(S))] \Big)^2 && (2) \\ &\leq \sum_{S \ni j, |S| \leq d} \mathrm{E}_u[(\hat{f}_u(\pi_e^{-1}(T)))^2] && (3) \\ &= \mathrm{E}_u \left[\sum_{T \ni \pi_e^{-1}(j), |T| \leq d} (\hat{f}_u(S))^2 \right] && (4) \\ &= \mathrm{E}_u[I_{\pi_e^{-1}(j)}^{\leq d}(f_u)]. && (5) \end{aligned}$$

Line 1 follows by the definition of influence, and line 2 by the definition of g_v. Note that we can use π_e^{-1} because π_e is a permutation. In particular, there is no need to refer to the projection π_e' from the last lecture, as $pi_e' = \pi_e$ so by last lecture $\chi_S(y \circ \pi_e) = \chi_{\pi_e'(S)}(y) = \chi_{\pi_e(S)}(y)$. Line 3 follows from Cauchy-Schwarz. Line 4 follows by setting $T = \pi^{-1}(S)$. Line 5 follows again by the definition of influence. Together this gives us:

$$\mathrm{E}_u[I_{\pi_e^{-1}(j)}^{\leq d}(f_u)] \geq \tau.$$

Applying Markov again we have that for a fraction at least $\frac{\tau}{2}$ of neighbors $u \in L$ of v that $I_{\pi_e^{-1}(j)}^{\leq d}(f_u) \geq \frac{\tau}{2}$. This means that a large fraction of v's neighbors have variables with high influence. These u's are "good" neighbors of v. Let $L_u = \{i \in [l] | I_i^{\leq d}(f_u) \geq \frac{\tau}{2}\}$. Notice that $|L_u| \leq \frac{d}{\tau/2}$, since

$$|L_u|\frac{\tau}{2} \leq \sum_{i \in L_u} I_i^{\leq d}(f_u) \leq \sum_{|S| \leq d, i \in S} (\hat{f}_u(S))^2 \leq d \cdot \sum_{|S| \leq d} (\hat{f}_u(S))^2 \leq d.$$

If we pick a value for u at random from L_u then the probability that c_e is satisfied for $e = (u, v)$ is at least $\frac{1}{|L_u|} \geq \frac{\tau}{2d}$. This is because v has value j so $c_{(u,v)}$ is satisfied iff u gets the value $\pi_e^{-1}(j)$, which lies in L_u since since $I_{\pi_e^{-1}(j)}^{\leq d}(f_u) \geq \frac{\tau}{2}$. Thus, the expected number of satisfied constraints is at least

$$\begin{aligned} \mathrm{E}[\#\{c_e\text{'s satisfied}\}] &\geq \Pr[v \text{ good}] \cdot \Pr[u \text{ good for } v \,|\, v \text{ good}] \cdot \Pr[\text{right value picked} \,|\, u \text{ and } v \text{ good}] \\ &= \delta \cdot \frac{\tau}{2} \cdot \frac{\tau}{2d} = \frac{\delta\tau^2}{4d}. \end{aligned}$$

This directly gives $\nu(G) \geq \frac{\delta\tau^2}{4d}$. □

By picking the positive constant γ small enough, we can make $\frac{\delta\tau^2}{4d} > \gamma$. By the soundness property of S_α (Claim 4), the former means that $\Pr[S_\alpha^f \text{ accepts }] < s_\alpha + 2\delta$. On the other hand, if $\nu^*(G) \geq 1 - \gamma$, the completeness property of the test T_α shows that $\Pr[S_\alpha^f \text{ accepts }] \geq c_\alpha(1-\gamma)$.

Let MAX-SAT(S_α) denote the problem of computing an assignment f such that $\Pr[S_\alpha^f \text{ accepts }]$ is maximized. The above shows that, if we had an approximation algorithm for MAX-SAT(S_α) that guarantees a factor of $\rho > \frac{s_\alpha+2\delta}{c_\alpha(1-\gamma)}$, we would be able to efficiently distinguish between the case where $\nu(G) < \gamma$ and $\nu^*(G) > 1-\gamma$. This is because if we run such an approximation algorithm on an instance where $\nu(G^*) > 1-\gamma$, then it returns a feasible solution f that satisfies more than a fraction $\rho(1-\gamma) \geq s_\alpha + 2\delta$ of the test constraints, whereas on instances where $\nu(G) < \gamma$, no feasible solution can satisfy that many of the test constraints. Since we can make δ and γ arbitrarily small, we conclude the following.

Theorem 1. *Let T_α, $\alpha \in A$, denote a family of dictatorship tests satisfying the completeness and soundness conditions stated in Section 2.2, and S_α the constraint graph game test described at the beginning of Section 3. Let h be any optimization problem that includes all problems of the form MAS-SAT(S_α) for $\alpha \in A$. Then h is UG-hard to approximate to within any constant factor $\rho > \inf_{\alpha\in A}(\frac{s_\alpha}{c_\alpha})$.*

If we plug in the test T_α from Section 2, the problem MAX-SAT(S_α) becomes a special case of MAX-3-LIN. Using the parameters c_α from Claim 1 and s_α from Claim 2, we conclude that it is UG-hard to approximate MAX-3-LIN to within any factor $\rho > \inf_{-1<\alpha<1}(\frac{s_\alpha}{c_\alpha}) = \inf_{0<\epsilon<1}(\frac{1/2}{1-\epsilon}) = \frac{1}{2}$.

4 Next Time

The above result is actually weaker than the inapproximability we derived last lecture for MAX-3-LIN, as we showed there that the same approximition factors are NP-hard rather than just UG-hard. However, next lecture we will pick the fruits of the generic approach we developed this lecture. We will see that when we plug in the noise sensitivity test for T_α in Theorem 1, then the problem MAX-SAT(S_α) becomes a special case of MAX-CUT, showing that approximating MAX-CUT to within any constant factor that is better than the current record of $\rho_{GW} \approx .878$ is UG-hard. Thus, we get a tight inapproximability result for MAX-CUT under the Unique Games Conjecture.