Hardness Of Approximation

Lecture 5: Finishing NP-hardness of $gap3Lin(\frac{1}{2}+\delta, 1-\delta)$

Instructor: Irit Dinur and Amey Bhangale *Scribe: Roie Salama*

1 Reminder from Last Week

Last week we described a reduction to gap-$3Lin$. The sketch is as follows:

1. Start with an instance I_1 of gap-$3Lin(0.99, 1-\epsilon)$.
2. Construct a "Smooth linear Label Cover" by doing smooth parallel repetition on I_1.
3. Construct the final instance I_2 of gap-$3Lin$.

Check the previous lecture notes for more details. In this lecture, we prove the completeness and soundness of the reduction, thereby finishing the proof of NP-hardness of gap-$3Lin(\frac{1}{2}+\delta, 1-\delta)$ for every constant $\delta > 0$.

2 Completeness and Soundness

2.1 Completeness

Claim 2.1. *If* $val(I_1) = 1-\epsilon \Rightarrow val(I_2) \geq 1 - k\epsilon$

Proof. In this case, the instance I_1 is $(1-\epsilon)$ satisfiable. We fix this assignment σ to the variables and set f_u to be the Hadamard encoding of σ restricted to the variables in u. Clearly, this assignment satisfies the folding property as it is coming from one fixed global assignment. Now, by union bound for at least $1-k\epsilon$ fraction of U, all the constraints are satisfied by σ. Thus, for these f_u's, the $3Lin$ test passes with probability 1. Hence, overall the acceptance probability is $1-k\epsilon$ and thus $val(I_2) \geq 1-k\epsilon$. □

2.2 Soundness

Claim 2.2. *Assume* $\delta \gg O(\beta\sqrt{k})$. *If* $Val(I_2) \geq \frac{1}{2}+\delta \Rightarrow$ *the value of the Label Cover instance* $\geq \frac{\delta^4}{128}$

Proof. By averaging, for at least $\frac{\delta}{2}$ of the $u's$, f_u passes the test with probability greater than $\frac{1}{2}+\frac{\delta}{2}$. Now, using Claim 4.1 from the previous lecture, for every such u we have $\mid \hat{f}_u(S) \mid \geq \delta$ for some $S \subseteq [3k]$ such that the vector 1_S satisfies the clauses $C_1^u, C_2^u, \ldots, C_k^u$ (constraints corresponding to u). Call these u's "good" , and also these $S's$ "good".

Note that good $S's$ are "rare" , in the sense that for every function f_u , at most $\frac{1}{\delta^2}$ of it's coefficients are greater than δ , that follows from the fact that $\sum_S \hat{f}_u(S)^2 = 1$.

For by the sub-code covering property, we have

$$\begin{aligned}
\hat{f}_u(S) &:= \mathop{\mathbb{E}}_{x\in\{0,1\}^{3k}}[f_u(x)\chi_S(x)] \\
\text{(Claim 3.2, lecture 4)} \quad &= \mathop{\mathbb{E}}_{v\in N(u)} \mathop{\mathbb{E}}_{y\in\{0,1\}^{3k-2\beta k}}[f_u(y^{\uparrow uv})\chi_S(y^{\uparrow uv}) \pm O(\beta\sqrt{k}) \\
&= \mathbb{E}_{v\in N(u)}[\hat{f}_v(S^{\downarrow uv})] \pm O(\beta\sqrt{k})
\end{aligned}$$

Here, $S^{\downarrow_{uv}}$ is an operation which takes a string from the domain of f_u to the domain of f_v in a natural way by deleting the extra coordinates. Thus, for a good S, we have

$$\mathbb{E}_{v\in N(u)}[\hat{f}_v(S^{\downarrow_{uv}})] \geq \delta - O(\beta\sqrt{k}) \geq \frac{\delta}{2},$$

assuming $\delta \gg O(\beta\sqrt{k})$. Note that for an edge (u,v) the label pair $(1_S, 1_{S^{\downarrow_{uv}}})$ satisfies the constraint on the edge (u,v) for every good S. Thus, the above inequality says that for a typical neighbor of a good u, the Fourier coefficient corresponding to the projected label is also heavy. More formally, by averaging argument, at least $\frac{\delta}{4}$ fraction of $v \in N(u)$, $|\hat{f}_v(S^{\downarrow_{uv}})|$ is $\frac{\delta}{4}$.

This gives a randomized strategy to get a non trivial satisfying assignment to the Label Cover instance. For every good u, assign a random label from the following set

$$\mathcal{L}_u := \{1_S \mid \hat{f}_u(S) \geq \delta \text{ and } 1_S \text{ satisfies } C_1^u, C_2^u \ldots, C_k^u\}.$$

Also, for $v \in V$, assign a random label from the following list

$$\mathcal{L}_v := \left\{1_T \mid \hat{f}_v(T) \geq \frac{\delta}{4}\right\}.$$

(Here f_v is embedded in f_u and it does not matter which u one chooses, as we get the same function because of the folding.) We know $|\mathcal{L}_v|$ is of size at most $\frac{16}{\delta^2}$. For a good u, $|\mathcal{L}_u| \geq 1$ and suppose we pick a label 1_S. Using the above argument, for at least $\frac{\delta}{4}$ fraction of $v \in N(u)$, the label $1_{S^{\downarrow_{uv}}}$ is present in $\mathcal{L}_v$. For such neighbors, the probability that the constraint on (u,v) is satisfied by the randomized labeling is at least $\frac{\delta^2}{16}$ (this is because we have to pick the correct label from a list of size at most $\frac{16}{\delta^2}$). Thus in expectation, for at least $\frac{\delta}{2}$ fraction of u (these are the good u's), for at least $\frac{\delta}{4}$ fraction of the neighbors of u, the constraint is satisfied with probability at least $\frac{\delta^2}{16}$. Overall, at least $\frac{\delta^4}{128}$ fractions of the constraints are satisfied by the randomized labeling in expectation. □

Recall that we are setting $k = \frac{1}{\sqrt{\epsilon}}$ and $\beta = \frac{1}{k^{0.6}}$. Thus, the in the soundness case, the value of the Label Cover instance is at most $2^{-\Omega(\beta k)} = 2^{-\Omega(k^{0.4})}$ and hence from the above claim $val(I_2)$ is at most $\frac{1}{2} + O(\beta\sqrt{k})$ in the soundness case for large enough k. Therefore we get that $gap\text{-}3Lin\left(\frac{1}{2} + O\left(\frac{1}{k^{0.1}}\right), 1 - \frac{1}{k^2}\right)$ is NP-hard. Setting k a large enough constant proves NP-hardness of $gap\text{-}3Lin(\frac{1}{2} + \delta, 1 - \delta)$ for every constant $\delta > 0$.

3 Boolean functions and influence

Definition 3.1. *for $f : \{-1,+1\}^n \to \{-1,+1\}$ define its i^{th} influence by*

$$Inf_i(f) = Pr_{x\sim\{-1,+1\}^n}[f(x) \neq f(x^{(i)})],$$

where $x^{(i)}$ means flipping the i-th bit of x .

For example, for the parity function

$$Parity(x_1, ..., x_n) = \prod_{i=1}^{n} x_i$$

$$\forall i : Inf_i(Parity) = 1$$

It is easy to see that the i^{th} influence is precisely the following:

$$Inf_i(f) = \sum_{i \in S} \hat{f}(S)^2$$

Next, we define

$$Inf_i^{(1-\delta)}(f) = \sum_{i \in S} (1-s)^{|S|-1} \hat{f}(S)^2$$

called the δ-*attenuated* influence. While there can be as many as n coordinates of a Boolean function whose influence is large (eg. Parity has all the influences 1), it is bounded in case of δ-attenuated influences of *any* Boolean function.

Claim 3.2. *for every* $f : \{-1,+1\}^n \rightarrow \{-1,+1\}$

$$| \{ i \in [n] \mid Inf_i^{(1-\delta)}(f) \geq \epsilon \} | \leq \frac{1}{\epsilon\delta}$$

Proof. $\sum_i Inf_i^{(1-\delta)}(f) = \sum_{i \in S} (1-\delta)^{|S|-1} \hat{f}(S)^2 = \sum_{S \subset [n]} \mid S \mid (1-\delta)^{|S|-1} \hat{f}(S)^2 \leq \sum_S \frac{1}{\delta} \hat{f}(S)^2 \leq \frac{1}{\delta}$

and the result follows (because otherwise the sum on the ϵ-heavy influences alone would exceed $\frac{1}{\delta}$). $\square$

References