18.218 Topics in Combinatorics Spring 2021 – Lecture 17

Dor Minzer

In this lecture, we will discuss the Vertex-Cover problem in greater detail, and prove that the simple 2-approximation algorithm we have seen for it is essentially optimal assuming the Unique-Games Conjecture.

1 Preliminaries

1.1 The independent set problem

For technical reasons, it is more convenient to work with the independent set problem. Given a graph $G = (V, E)$, an independent set in G is a set of vertices $I \subseteq V$ that contains no edge from E. Noting that a set I is an independent set iff $V \setminus I$ is a vertex cover, it follows that $\mathsf{IS}(G) = n - \mathsf{VC}(G)$, where $\mathsf{IS}(G)$ denotes the size of the maximum independent set in G, and $\mathsf{VC}(G)$ denotes the size of the smallest vertex cover in G.

The main result we present here is the following hardness result:

Theorem 1.1. *Assuming UGC, for all $\varepsilon > 0$, given a graph $G = (V, E)$ it is NP-hard to distinguish between the cases:*

1. **YES case***: $\mathsf{IS}(G) \geqslant \left(\frac{1}{2} - \varepsilon\right) n$.*
2. **NO case***: $\mathsf{IS}(G) \leqslant \varepsilon n$.*

Using the observation above about the relationship between independent sets and vertex-covers, we have:

Corollary 1.2. *Assuming UGC, for all $\varepsilon > 0$, given a graph $G = (V, E)$ it is NP-hard to distinguish between the cases:*

1. **YES case***: $\mathsf{VC}(G) \leqslant \left(\frac{1}{2} + \varepsilon\right) n$.*
2. **NO case***: $\mathsf{VS}(G) \geqslant (1 - \varepsilon)n$.*

In particular, it follows from the corollary that assuming UGC, it is NP-hard to approximate the size of the minimum vertex cover up to factor $\frac{1-\varepsilon}{1/2+\varepsilon} = 2 - O(\varepsilon)$, for all $\varepsilon > 0$. We will henceforth focus our discussion on proving Theorem 1.1.

The proof of the theorem follows the dictatorship testing framework introduced in the previous lecture, and we begin by introducing the two ingredients.

1.2 The p-biased Kneser graph

Recall that last time, we used the ρ-noisy cube to encode dictatorships using cuts. Here, we want to use independent sets in order to encode dictatorships, and we want that a dictatorship will correspond to a large independent set whereas functions that are very far from being dictatorships will correspond to either sets that are not independent or small independent sets. Towards this end, we introduce the p-biased Kneser graph.

Definition 1.3. *Let $0 \leqslant p < 1/2$. The p-biased Kneser graph is a graph whose vertex set is $P([n])$, and the weight of a set A is $\mu_p(A) = p^{|A|}(1-p)^{n-|A|}$. The edges of the graph are $E = \{ (A, B) \mid A \cap B = \emptyset \}$.*

Looking at the Kneser graph, we see that independent sets in it are creatures we have already encountered: they are simply intersecting families. In this language, a dictatorship family is a family of the form $\mathcal{F} = \{ A \mid A \ni i \}$, and it has p-biased measure p (and it is also the heaviest independent set in the graph). Moreover, we have proved that any independent set in the graph, which is an intersecting family, is nearly contained in a junta. Thus, if it is of non-negligible measure, it must have an influential coordinate. Therefore, at least intuitively, if we have a family of non-negligible measure with no influential variables, we automatically get it is not independent. This observation is what will ultimately give us the gap.

1.3 A stronger form of the Unique-Games Conjecture

The second ingredient we need is a stronger form of the Unique-Games Conjecture. Recall that an assignment to a Unique-Game is a labeling $A\colon V \to \Sigma$, and we say an assignment satisfies a constraint if it gives the two endpoints of the edge matching labels. A t-assignment is a labeling which assigns to each vertex t possible assignments, i.e. $A\colon V \to \binom{\Sigma}{t}$. We say A satisfies an edge (u, v) if $A(u), A(v)$ contain pairs of labels that are compatible with the constraint of u, v, i.e. if $A(v) \cap \phi_{u,v}(A(u)) \neq \emptyset$.

Definition 1.4 (A strongish form of UGC)**.** *For $\eta > 0$, $t \in \mathbb{N}$ we are given a non-bipartite Unique-Games instance $\Psi = (X, E, \Sigma, \Phi)$, and we wish to distinguish between the following two cases:*

1. **YES case***: there exists $X' \subseteq X$ of size at least $(1-\eta)|X|$, and an assignment $A\colon X' \to \Sigma$, such that A satisfies all of the constraints inside X'.*
2. **NO case***: for all $X' \subseteq X$ of size at least $\eta|X|$ and a t-assignment, i.e. $A\colon X' \to \binom{\Sigma}{t}$, not all of the constraints in X' are satisfied by A.*

It turns out that there is a reduction from general Unique-Games to strongish unique games. Namely, one can prove the following result.

Theorem 1.5. *Assuming UGC, for all $t \in \mathbb{N}$, $\eta > 0$, the problem gap-StrongishUG$_t[1-\eta, \eta]$ is NP-hard.*

We will not prove this theorem here. Instead, we will show a reduction from Strongish Unique-Games to the Indepednent Set problem.

2 The reduction

Fix $\varepsilon > 0$, and denote $p = \frac{1}{2} - \varepsilon$. Starting with an instance of strongish unique games $\Psi = (X, E, \Sigma, \Phi)$, our goal is to produce a (weighted) graph $G = (V, w, E')$ such that: if Ψ was in the YES case, then G has an independent set of eight at least $p-\varepsilon$, and if Ψ was in the NO case, then the largest indepednent set in G has weight at most ε.

For each $x \in X$, we produce a copy of the Kneser graph over Σ. Namely, our vertices are $V = X \times P(\Sigma)$ as for the edges, we create edges inside the Kneser graph of each vertex $x \in X$, i.e. we insert the edges

$$E_1 = \{\{(x,A),(x,B)\} \mid x \in X, A, B \subseteq \Sigma, A \cap B = \emptyset\}.$$

We also add edges across the Kneser graphs; the idea is similar to the idea in the last lecture, since we want to ensure compatibility between the encodings given by each Kneser graph. More precisely, we add the edges

$$E_2 = \left\{ \{(x,A),(x',B)\} \quad x, x' \in X, (x,x') \in E, A, B \subseteq \Sigma, B \cap \phi_{x,x'}(A) = \emptyset \quad . \right.$$

In words, we add edges between the vertices in the Kneser graphs of x and x' if the two sets do not contain a pair of labels that is compatible with the constraint $\phi_{x,x'}$. The set of edges E' is $E_1 \cup E_2$.

Finally, the weight of a vertex (x,A) is $w(x,A) = \frac{1}{|X|}\mu_p(A)$. This completes the description of the reduction.

We prove the following lemma which summarizes the properties of the reduction.

Lemma 2.1. *For all $\varepsilon > 0$, $p = \frac{1}{2} - \varepsilon$ there are $t \in \mathbb{N}$ and $\eta > 0$ such that the following holds.*

1. *If Ψ is from the* **YES case** *of Strongish Unique-Games, then G has an independent set of weight at least $p - \varepsilon$.*

2. *If Ψ is from the* **NO case** *of Strongish Unique-Games, then the heaviest independent set in G has weight at most ε.*

This lemma, together with Theorem 1.5, immediately implies Theorem 1.1.

3 Analysis of the reduction

We now analyze the reduction.

3.1 Completeness

Suppose there is $X' \subseteq X$ of size at least $(1-\eta)|X|$ and an assignment $H \colon X' \to \Sigma$ satisfying all of the constraints in X'. Define

$$I = \left\{ (x,A) \mid x \in X', A \ni H(x) \quad . \right.$$

We claim that I has weight $p - \eta$ and that I is an independent set. Indeed,

$$w(I) = \sum_{x \in X'} \frac{1}{|X|}\mu_p(\{A \subseteq \Sigma \mid A \ni H(x)\}) = \sum_{x \in X'} \frac{1}{|X|}p = \frac{|X'|}{|X|}p \geqslant (1-\eta)p \geqslant p - \eta \geqslant p - \varepsilon.$$

Now, to see that I is an independent set, assume for contradiction it contains an edge between say (x,A) and (x',B). Clearly we must have that $x \neq x'$ (otherwise A, B both contain $H(x)$), so H satisfies the constraint between x and x'. It thus follows that A contains $H(x)$, and B contains $H(x') = \phi_{x,x'}(H(x))$, and so $\phi_{x,x'}(A) \cap B \neq \emptyset$, so this is actually not an edge in the graph G. Contradiction.

3.2 Soundness

We now move on to the soundness of the construction. We will not give the entire argument, and instead convey the spirit of the matter.

Suppose we have an independent set I in G of weight at least ε. Consider the upwards closure of the family

$$I\uparrow = \{\, (x, A) \mid \exists B \subseteq A, (x, B) \in I\,\}\,.$$

It is easy to see that $I\uparrow$ is also independent set. Abusing notations, we drop the $\uparrow$ notation and denote this family by I.

For each $x \in X$, let

$$I_x = \{\, A \subseteq \Sigma \mid (x, A) \in I\,\}\,.$$

Note that

$$\mathop{\mathbb{E}}_{x\in X}\left[\mu_p(I_x)\right] = w(I) \geqslant \varepsilon,$$

so letting $X_1 = \{x \mid \mu_p(I_x) \geqslant \varepsilon\}$, we have that $|X'| \geqslant \frac{\varepsilon}{2}\,|X|$.

Next, define $f\colon [p, p+\varepsilon/2] \to [0,1]$ by $f(q) = \mathbb{E}_{x\in X}\left[\mu_q(I_x)\right]$. Then by Lagrange we may find $q' \in (p, p+\varepsilon/2)$ such that

$$f'(q') = \frac{f(q+\varepsilon/2) - f(q)}{q+\varepsilon/2-q} \leqslant \frac{2}{\varepsilon}.$$

Thus,

$$\mathop{\mathbb{E}}_{x}\left[\frac{d}{dq}\mu_q(I_x)|_{q=q'}\right] \leqslant \frac{2}{\varepsilon},$$

and letting $X_2 = \left\{x \mid \frac{d}{dq}\mu_q(I_x)|_{q=q'} \leqslant \frac{8}{\varepsilon^2}\right\}$ we have by Markov's inequality that $|X_2| \geqslant \left(1-\frac{\varepsilon}{4}\right)|X|$. We fix q' and X_2 henceforth.

Thus, we take $X' = X_1 \cap X_2$, and note that $|X'| \geqslant \frac{\varepsilon}{4}\,|X|$.

3.2.1 Applying Friedgut's Junta Theorem

Note that here, for each $x \in X'$ we have that $\frac{d}{dq}\mu_q(I_x)|_{q=q'} \leqslant \frac{8}{\varepsilon^2}$, hence by Friedgut's Theorem we get that I_x is close to a junta, i.e. to a family that depends only on coordinates $J_x \subseteq \Sigma$ where $|J_x| = O_\varepsilon(1) = t$.

Here, we will make a simplifying assumption. This assumption is not necessary and there are ways to circumvent it, but it adds an additional technical difficulty which we wish to avoid. Instead of assuming that I_x is close to a junta, we will simply assume it is a junta. Namely, we will assume that there is $\mathcal{J}_x \subseteq J_x$ such that

$$I_x = \{\, A \subseteq \Sigma \mid A \cap J_x \in \mathcal{J}_x\,\}\,.$$

We are now ready to define a t-assignment. Indeed, define $H\colon X' \to \binom{X}{t}$ by $H(x) = J_x$. To finish the proof, we show that H satisfies all of the constraints inside X', which gives us contradiction to the fact Ψ was from the no case.

3.2.2 The Juntas are intersecting

Lemma 3.1. *H defined above satisfies all of the constraints inside X'.*

Proof. Assume towards contradiction this is not the case, and let $x, x' \in X'$ be such that H fails to satisfy the constraint between them. Then $J_{x'} \cap \phi_{x,x'}(J_x) = \emptyset$. As $\mu_q(I_x), \mu_q(I_{x'}) > 0$, we may find $A_x \in \mathcal{J}_x$, $A_{x'} \in \mathcal{J}_{x'}$, and as $J_{x'} \cap \phi_{x,x'}(J_x) = \emptyset$, we have that $A_{x'} \cap \phi_{x,x'}(A_x) = \emptyset$. Thus, there is an edge between (x, A_x) and $(x', A_{x'})$ in our graph G, and we next argue that both of these points are in I. This yields a contradiction to the fact that I is an independent set.

Indeed, as I_x is a J_x junta and $A_x \in \mathcal{J}_x$, we get that $A_x \in I_x$, and similarly for $A_{x'}$. This completes the proof. $\square$

18.218 Topics in Combinatorics: Analysis of Boolean Functions
Spring 2021