$(2-\varepsilon)$-hardness of approximating Vertex Cover[1]

We want to show that (weighted) Vertex Cover is hard to approximate to within $2-\varepsilon$ for any $\varepsilon > 0$ assuming the Unique games conjecture. This was first proved by Khot and Regev [KR08], building on the work of Dinur and Safra [DS05], who showed that it is NP-hard to approximate the (weighted) Vertex Cover to within $10\sqrt{5} - 21 - \varepsilon = 1.3606 - \varepsilon$ for any $\varepsilon > 0$.

1 The UG variant we use

We will use the following variant of the UG problem, introduced by Khot and Regev [KR08]. For parameters $t, r \in \mathbb{N}$ and $\nu \in (0,1)$, we define the (t,r,ν)-Label Cover (also denoted (t,r,ν)-LC) problem as follows: an instance of this problem is a 4-tuple (X, E, L, Ψ), where (X, E) is an undirected graph (not necessarily bipartite), L is a set of r labels, and $\Psi = \{\psi_e \mid e \in E\}$ is a set of permutations of L (i.e. each $\psi_e \subseteq L \times L$ is the graph of a permutation). The YES and NO instances of this problem are defined as follows:

- YES case: $\exists X_0 \subseteq X$ such that $|X_0| \geq (1-\nu)|X|$ and a labeling $\sigma : X \to L$ s.t. for all $(u,v) \in E \cap (X_0 \times X_0)$, the labeling σ satisfies $\psi_{(u,v)}$: that is, $(\sigma(u), \sigma(v)) \in \psi_{(u,v)}$.
- NO case: $\forall X_0 \subseteq X$ such that $|X_0| \geq \nu|X|$ and for all labelings $\sigma : X_0 \to \binom{L}{\leq t}$, there exists an edge $e = (u,v) \in E \cap (X_0 \times X_0)$ such that $(\sigma(u) \times \sigma(v)) \cap \psi_e = \emptyset$. I.e. even when we are allowed to assign up to t labels to each of the vertices in X_0, there is always an edge e with both endpoints in X_0 such that no pair of (up to t^2) labels satisfies the constraint corresponding to edge e.

We use the following result of Khot and Regev [KR08] without proof:

Theorem 1. *Assume the Unique Games Conjecture. Then $\forall t \in \mathbb{N}$, $\forall \nu \in (0,1)$, $\exists r \in \mathbb{N}$ such that (t,r,ν)-Label Cover is NP-hard.*

2 Reduction from the UG variant to VC

The reduction we outline will produce, given a (t,r,ν)-LC instance I, a node weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ with the total weight of all nodes being 1 and such that:

- If I is a YES instance, then there exists an Independent set in $\mathcal{G}$ of weight at least $\frac{1}{2} - \alpha$. (And hence, there is a vertex cover of size at most $1/2 + \alpha$.)
- If I is a NO instance, then every independent set in $\mathcal{G}$ has weight at most β. (And hence every vertex cover is of size at least $1 - \beta$.)

[1] Prahladh: Thanks to Srikanth Srinivasan for scribing these notes. The proof given here is essentially the proof of Khot-Regev and Dinur-Safra, with a simplification due to Harsha, Håstad and Sachdeva and a further simplification due to Srinivasan while scribing the notes.

Putting the above together, we obtain the result that Vertex Cover is UG-hard to approximate within $2 - O(\alpha + \beta)$. Since α and β will be parameters we can make arbitrarily small, this will prove the hardness result we want.

For now, we leave t, ν as parameters. They will be fixed at the end of the proof. The label size r will then be the one guaranteed by Theorem 1.

2.1 Description of $\mathcal{G}$

Let $\varepsilon > 0$ be a small constant parameter and let $p = \frac{1}{2} - \varepsilon$.

Starting from an instance $I = (X, E, L, \Psi)$ of (t, r, ν)-LC, the corresponding instance $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ of Vertex Cover is defined as follows:

- $\mathcal{V} = X \times 2^L$,
- $\mathcal{E} = \{((u, F), (v, G)) \mid (u, v) \in E \text{ and } (F \times G) \cap \psi_{(u,v)} = \emptyset\}$, and
- $w(u, F) = \frac{1}{|X|} p^{|F|}(1-p)^{r-|F|}$. In other words, the weight function defines a probability distribution on $\mathcal{V}$ where to pick a random element (u, F) of $\mathcal{V}$, we choose $u \in X$ uniformly at random and F according to the p-biased measure on $\{0,1\}^L = 2^L$.[2]

Let $\mathrm{Ind}(\mathcal{G})$ denote the weight of the largest independent set in $\mathcal{G}$.

2.2 Completeness

Claim 2. *If I is a YES instance of (t, r, ν)-LC, then* $\mathrm{Ind}(\mathcal{G}) \geq (1 - \nu) \cdot (\frac{1}{2} - \varepsilon)$.

Proof sketch. Let $\sigma : X \to L$ be a labeling that witnesses the fact that I is a YES instance. There is a subset $X_0 \subseteq X$ of size at least $(1 - \nu)|X|$ such that σ satisfies all the edges with both endpoints in X_0. The set $\{(u, F) \mid u \in X_0, F \ni \sigma(v)\} \subseteq \mathcal{V}$ is an independent set in $\mathcal{G}$ of the required weight. ☐

2.3 Soundness

Claim 3. *There exists $t = t(\nu, \varepsilon) \in \mathbb{N}$ such that if I is a NO instance of (t, r, ν)-LC, then* $\mathrm{Ind}(\mathcal{G}) \leq 2\nu$.

Proof. Assume for the sake of contradiction that there is an independent set $\mathcal{I}$ of $\mathcal{G}$ of weight at least 2ν. Without loss of generality, assume that $\mathcal{I}$ is a maximal independent set.

We treat $\mathcal{I}$ as a Boolean function with domain $\mathcal{V} = X \times 2^L$. For each $v \in X$, we have a natural restriction $\mathcal{I}_v : \{0,1\}^L \to \{0,1\}$. Let μ_p denote the p-biased measure over $\{0,1\}^L$.

Since $\mathbf{E}_{v \in X, F \sim \mu_p}[\mathcal{I}(v, F)] \geq 2\nu$, we see that with probability at least ν over the choice of $v \in X$, we have

$$\mathop{\mathbf{E}}_{F}[\mathcal{I}_v(F)] \geq \nu.$$

[2] Throughout, we blur the distinction between $\{0,1\}^L$ and 2^L.

Let X_0 denote the set of $v \in X$ satisfying the above condition. We have $|X_0| \geq \nu|X|$.

Since $\mathcal{I}$ is an independent set in $\mathcal{G}$, we see that for each $(u, F), (v, G) \in \mathcal{I}$, we must have either $(u, v) \notin E$ or $(F \times G) \cap \psi_{(u,v)} \neq \emptyset$. In particular, this easily implies that for any $F' \supseteq F$ and $G' \supseteq G$, the vertices (u, F') and (v, G') continue to be non-adjacent. Since $\mathcal{I}$ is a maximal independent set, we therefore see that $\mathcal{I}_u(F') = \mathcal{I}_v(G') = 1$. That is, the functions $\mathcal{I}_u$ for $u \in X$ are monotone.

A digression into Extremal Combinatorics. Given a function $f : \{0,1\}^L \to \{0,1\}$ and $i \in L$, we define the influence of the ith variable w.r.t. the measure μ_p as follows:

$$\mathrm{Inf}_i^p(f) := \mathop{\mathbb{E}}_{x_{-i} \sim \mu_p^{\otimes(n-1)}} \left[\mathop{\mathrm{Var}}_{x_i \sim \mu_p} [f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)] \right] = p(1-p) \cdot \Pr_{x \sim \mu_p^{\otimes n}} [f(x) \neq f(x \oplus e_i)]$$

The total influence of f w.r.t. μ_p is defined to be

$$\mathrm{I}^p(f) = \sum_{i \in L} \mathrm{Inf}_i^p(f)$$

For $C \subseteq L$, we say that f is a C-junta if it only depends on the inputs indexed by C. Further, we say that f is δ-approximated by a C-junta if there is a C-junta g such that

$$\Pr_{x \sim \mu_p} [f(x) \neq g(x)] \leq \delta.$$

The following is a slight specialization of an important theorem of Friedgut, which allows us to say that functions of low total influences are well approximated by juntas.

Theorem 4 (Friedgut's theorem)**.** *Fix any $p \in (1/4, 1/2)$, $\delta \in (0,1)$ and $k \in \mathbb{N}$. Let $f : \{0,1\}^L \to \{0,1\}$ be a function such that $\mathrm{I}^p(f) \leq k$. Then, f is δ-approximated by a C-junta where $|C| \leq \exp(O_p(k/\delta))$.*

However, the above theorem is not applicable unless f has low total influence. Our functions $\mathcal{I}_v$ do not necessarily have that property. However, it turns out that we can turn the functions into those of low total influence by tweaking the bias. Here, we use crucially the monotonicity of these functions. Details follow.

We use the following lemma about monotone functions.

Lemma 5 (Russo's lemma (Margulis, Russo))**.** *If f is a monotone Boolean function, then $\frac{d\mu_p(f)}{dp} = I^p(f)/p(1-p)$ where $\mu_p(f)$ denotes the expectation of f w.r.t. μ_p.*

Note that if f is a monotone Boolean function then $\mu_p(f)$ is a quantity that is bounded between 0 and 1 and increases monotonically with p. In particular, the derivative of this quantity w.r.t. p cannot be uniformly large. Specifically, in any interval of length $\varepsilon' > 0$, there must be a point where it is at most $1/\varepsilon'$. Putting this observation together with Friedgut's theorem, we obtain the following monotone version of Friedgut's theorem due to Dinur and Safra.

Theorem 6 (Friedgut's theorem for monotone functions (Dinur-Safra))**.** *Let $p = \frac{1}{2} - \varepsilon$ for $\varepsilon \in (0, 1/4)$. Fix any $\delta \in (0,1)$ and $k \in \mathbb{N}$. Let $f : \{0,1\}^L \to \{0,1\}$ be a monotone function. Then, there is a $q \in (p, p + \varepsilon/2)$ and an $\exp(O(1/\delta\varepsilon))$-junta f' such that*

$$\Pr_{x \sim \mu_q} [f(x) \neq f'(x)] \leq \delta.$$

We now return to the proof of Claim 3.

Using the above theorem for $\delta := \nu^2/16$, for each $v \in X_0$, we can choose a $q_v \in (p, p+\varepsilon/2)$, $C_v \subseteq L$ with $|C_v| \leq h(\delta,\varepsilon) = \exp(O(1/\delta\varepsilon))$, and a C_v-junta f_v such that f_v δ-approximates $\mathcal{I}_v$ w.r.t. the measure μ_{q_v}. Also note that $\frac{1}{2} - \varepsilon \leq q_v \leq p + \varepsilon/2 = \frac{1}{2} - \frac{\varepsilon}{2}$.

We define $t = h(\delta, \varepsilon)$. Let $\sigma : X_0 \to \binom{L}{\leq t}$ be defined by $\sigma(v) = C_v$. We would like to show that for all edges $e = (u,v) \in E \cap (X_0 \times X_0)$, we have $(\sigma(u) \times \sigma(v)) \cap \psi_{(u,v)} \neq \emptyset$.

Fix any edge $(v_1, v_2) \in E \cap (X_0 \times X_0)$. For brevity, we use $\mathcal{I}_1, \mathcal{I}_2, f_1, f_2$, etc. instead of $\mathcal{I}_{v_1}, \mathcal{I}_{v_2}, f_{v_1}, f_{v_2}$, etc.. Without loss of generality, assume that $\psi_{(v_1,v_2)}$ is the identity permutation. Then we need to show that $C_1 \cap C_2 \neq \emptyset$.

We argue by contradiction. Assume that $C_1 \cap C_2 = \emptyset$. We then show that $\mathcal{I}$ is not an independent set by finding $(v_1, F_1), (v_2, F_2) \in \mathcal{I}$ such that $((v_1, F_1), (v_2, F_2)) \in \mathcal{E}$. Consider the following probabilistic process of picking F_1, F_2. For each $i \in L$, we place it in F_1 with probability q_1, in F_2 with probability q_2 and in neither with probability $1 - (q_1 + q_2)$ (which is positive since $q_1, q_2 < 1/2$). Note that by definition $F_1 \cap F_2 = \emptyset$ and hence we have $((v_1, F_1), (v_2, F_2)) \in \mathcal{E}$ with probability 1. We only need to argue that with positive probability these elements belong to $\mathcal{I}$: equivalently, we need to show that with positive probability, $\mathcal{I}_j(F_j) = 1$ for each $j \in \{1,2\}$.

Consider first the probability that $f_j(F_j)$ is 1. Note that by the monotonicity of $\mathcal{I}_j$ and the fact that $q_j \geq p$, we have $\mu_{q_j}(\mathcal{I}_j) \geq \mu_p(\mathcal{I}_j) \geq \nu$. As f_j is a δ-approximation to $\mathcal{I}_j$ w.r.t. μ_{q_j}, we see that $\mu_{q_j}(f_j) \geq \nu - \delta \geq \nu/2$. As F_j is distributed according to μ_{q_j}, we see that $\Pr_{F_j}[f_j(F_j) = 1] \geq \nu/2$.

Now, since $C_1 \cap C_2 = \emptyset$ by assumption, we see that the events $f_j(F_j)$ depend on disjoint sets of co-ordinates in L and hence, the events are mutually independent. Thus, we have

$$\Pr_{(F_1,F_2)}[f_1(F_1) = 1 \wedge f_2(F_2) = 1] \geq \nu^2/4.$$

Finally, we have

$$\begin{aligned}\Pr_{(F_1,F_2)}[\mathcal{I}_1(F_1) = 1 \wedge \mathcal{I}_2(F_2) = 1] &\geq \Pr_{(F_1,F_2)}[f_1(F_1) = 1 \wedge f_2(F_2) = 1] \\ &\quad - \Pr_{F_1}[f_1(F_1) = 1 \wedge \mathcal{I}_1(F_1) \neq 1] - \Pr_{F_2}[f_2(F_2) = 1 \wedge \mathcal{I}_2(F_2) \neq 1] \\ &\geq \nu^2/4 - 2\delta \geq \nu^2/8 > 0\end{aligned}$$

where the second inequality follows from the fact that f_j is a δ-approximation to $\mathcal{I}_j$ w.r.t. μ_{q_j}.

□

References

[DS05] IRIT DINUR and SHMUEL SAFRA. *On the hardness of approximating minimum vertex cover*. Ann. of Math., 162(1):439—485, 2005. (Preliminary version in *34th STOC*, 2002). doi:10.4007/annals.2005.162.439. 1

[KR08] SUBHASH KHOT and ODED REGEV. *Vertex cover might be hard to approximate to within 2-ϵ*. J. Comput. Syst. Sci., 74(3):335–349, 2008. (Preliminary version in *18th IEEE Conference on Computational Complexity*, 2003). doi:10.1016/j.jcss.2007.06.019. 1