Inapproximability of Vertex Cover and Independent Set in Bounded Degree Graphs

Per Austrin
KTH – Royal Inst. of Technology
Stockholm, Sweden

Subhash Khot
New York University
New York, U.S.

Muli Safra
Tel-Aviv University
Tel-Aviv, Israel

***Abstract*—We study the inapproximability of Vertex Cover and Independent Set on degree d graphs. We prove that:**

- **Vertex Cover is Unique Games-hard to approximate to within a factor $2-(2+o_d(1))\frac{\log\log d}{\log d}$. This exactly matches the algorithmic result of Halperin [1] up to the $o_d(1)$ term.**
- **Independent Set is Unique Games-hard to approximate to within a factor $O(\frac{d}{\log^2 d})$. This improves the $\frac{d}{\log^{O(1)}(d)}$ Unique Games hardness result of Samorodnitsky and Trevisan [2]. Additionally, our result does not rely on the construction of a query efficient PCP as in [2].**

I. Introduction

Vertex Cover and Independent Set are two of the most well-studied NP-complete problems. On general graphs (i.e. with unbounded degree), it is a notoriously difficult problem even to approximate the solutions to these problems, and there is a strong evidence that indeed no *good* approximation is feasible. However for graphs whose degree is bounded by a constant, significantly better approximation guarantees are known. In this paper, we investigate whether one can obtain a tight inapproximability result for graphs with bounded degree d as a function of d. We present a randomized reduction from the Unique Games problem to each of these two problems, giving UG-hardness results close to the approximation ratio of the best algorithms known[1].

Our Results

For the Vertex Cover problem, we prove:

Theorem I.1. *For every sufficiently large integer d, it is* UG*-hard (under randomized reductions) to approximate vertex cover in a degree d graph to within factor $2-(2+o_d(1))\frac{\log\log d}{\log d}$.*

We note that Halperin [1] presents an efficient algorithm that approximates vertex cover in a degree d graph to within essentially the same factor, up to the $o_d(1)$ term. This improves on the general well-known 2-approximation ratio for graphs of unbounded degree. On the inapproximability side, Khot and Regev [3] showed $2-\varepsilon$ UG-hardness result for any constant $\varepsilon > 0$, whereas Dinur and Safra [4] showed 1.36 NP-hardness result.

For the Independent Set problem, we prove:

Theorem I.2. *For every sufficiently large integer d, it is* UG*-hard (under randomized reductions) to approximate independent set in a degree d graph to within factor $O(\frac{d}{\log^2 d})$.*

This result is close to the best known algorithm for this problem that achieves $O(\frac{d\log\log d}{\log d})$ approximation (see Halperin [1], or Halldórsson [5]). It is an intriguing question whether one can improve the approximation algorithm, or improve on this inapproximability result (or both). Previously, Samorodnitsky and Trevisan [2] showed $\frac{d}{\log^{O(1)} d}$ UG-hardness for the problem (doing optimistic estimates, it seems the best possible result their proof could yield is $\frac{d}{\log^3 d}$). The same authors, in an earlier work [6], gave $\frac{d}{2^{O(\sqrt{\log d})}}$ NP-hardness result. For graphs with unbounded degree, the best algorithm known, due to Feige [7], achieves an approximation ratio of $O\left(\frac{n(\log\log n)^2}{(\log n)^3}\right)$, whereas the problem was shown to be hard to approximate within $n^{1-\varepsilon}$ for any constant $\varepsilon > 0$ by Håstad [8]. Håstad's result has been further improved and the current best inapproximability result is $\frac{n}{2^{(\log n)^{3/4+\varepsilon}}}$ by Khot and Ponnuswami [9].

Techniques

Showing an inapproximabilty result for Vertex Cover essentially amounts to showing the Independent Set problem is hard to approximate even when the size of the Independent Set is very large. For an inapproximability ratio close to 2, this calls for showing that it is hard to distinguish between a graph with an independent set of roughly half of the vertices, and a graph in which every independent set has negligible size. Consequently, both our results follow from the same randomized reduction from the Unique Games problem, albeit with different choices of parameters.

The reduction produces an n-vertex degree d graph, which, in case the Unique-Game instance was al-

[1]See Section II-A for definitions of Unique Games problem and UG-hardness of a problem.

most completely satisfiable — the completeness case — has a *large* independent set. Here *large* refers to $\left(\frac{1}{2} - \Theta(\frac{\log\log d}{\log d})\right) \cdot n$ for Theorem I.1, and $\Theta(\frac{1}{\log d}) \cdot n$ for Theorem I.2. In contrast, if one can satisfy only a small fraction of the constraints of the Unique-Game instance — the soundness case — there is no independent set of size even βn for an appropriately small constant β, where $\beta = \frac{1}{\log d}$ for Theorem I.1 and $\beta = \Theta\left(\frac{\log d}{d}\right)$ for Theorem I.2.

The reduction proceeds in two steps: (1) the first step produces a graph G with unbounded degree and (2) in the second step, we sparsify the graph so as to have all degrees bounded by d, yielding the final graph G'. The sparsification step simply picks $d \cdot n$ edges from G at random so that the average degree (and hence the maximum degree after removing a small fraction of edges) is bounded by d.

The second step clearly can only increase the size of the independent-set, hence the completeness proof is fine. For the soundness proof, we must show that the size of the independent set can only be slightly increased. We prove that if G had no independent set of size βn, G' does not have independent set of size βn either. In order to prove this, we actually need the graph G to have a stronger property. In the soundness case, we show not only that G has no independent set of size βn, we have a much stronger *density* property: every set of βn vertices contains a $\Gamma(\beta)$ fraction of the edges for an appropriate function $\Gamma(\cdot)$. This stronger property allows us to prove the correctness of the sparsification step by a simple union bound over all sets of size βn.

Now, let us elaborate on the first step of the reduction. This step is almost the same as in Khot and Regev's paper [3]. Their reduction produces an n-vertex graph that has no independent set of size βn. We show that one can in fact define an appropriate probability distribution on the edges of their graph and prove the density property that every set of βn vertices contains $\Gamma(\beta)$ fraction of the edges. The analysis of this step departs from that of Khot and Regev, and is instead inspired by that of Dinur et al. [10] for showing UG-hardness for coloring problems. The density property follows from a quite straightforward application of a *Thresholds are Stablest* type theorem [11], giving precise bounds on the function $\Gamma(\cdot)$. Note that we also obtain an alternate proof of the $2 - \varepsilon$ inapproximability result for vertex cover that is arguably simpler than the Khot-Regev proof.

It is interesting that we obtain a $2-(2+o_d(1))\frac{\log\log d}{\log d}$ inapproximability result for vertex cover where even the constant in front of the $\frac{\log\log d}{\log d}$ term is optimal. Finally, we remark that the earlier result of Samorodnitsky and Trevisan [2] that gave $\frac{d}{\log^{O(1)} d}$ inapproximability for independent set problem, requires the construction of a sophisticated query-efficient PCP. We instead prove an improved result without relying on such a PCP.

II. Preliminaries

We will consider graphs that are both vertex weighted and edge weighted. We will assume that the sum of the vertex weights equals 1 and so does the sum of the edge weights so that the weights can be thought of as probability distributions. For a weighted graph G and a subset of its vertices S, let $w(S)$ denote the weight of vertex set S and $G(S)$ denote the induced subgraph on S. For vertex sets S and T, let $w(S,T)$ denote the weight of edges between vertex sets S and T. As a convention, an unweighted graph would refer to a graph with uniform probability distributions over its vertices and edges.

Definition II.1. A graph G is (δ, ε)-dense if for every $S \subseteq V(G)$ with $w(S) \geq \delta$, the total weight $w(S,S)$ of edges inside S is at least ε.

A. Unique Games

In this section, we state the formulation of the Unique Games Conjecture that we will use.

Definition II.2. An instance $\Lambda = (U, V, E, \Pi, [L])$ of *Unique Games* consists of an unweighted bipartite multigraph $G = (U \cup V, E)$, a set Π of *constraints*, and a set $[L]$ of *labels*. For each edge $e \in E$ there is a constraint $\pi_e \in \Pi$, which is a permutation on $[L]$. The goal is to find a *labelling* $\ell : U \cup V \to [L]$ of the vertices such that as many edges as possible are satisfied, where an edge $e = (u,v)$ is said to be satisfied by ℓ if $\ell(v) = \pi_e(\ell(u))$.

Definition II.3. Given a Unique Game instance $\Lambda = (U, V, E, \Pi, [L])$, let $\mathrm{Opt}(\Lambda)$ denote the maximum fraction of simultaneously satisfied edges of Λ by any labelling, i.e.

$$\mathrm{Opt}(\Lambda) := \frac{1}{|E|} \max_{\ell : U \cup V \to [L]} |\{\, e : \ell \text{ satisfies } e \,\}|.$$

Let $\mathrm{IndOpt}(\Lambda)$ denote the maximum value α such that there is a subset $V' \subseteq V$, $|V'| \geq \alpha |V|$ and a labeling $\ell : U \cup V' \mapsto [L]$ such that every edge in the induced subgraph $G(U \cup V', E)$ is satisfied by the labeling ℓ.

The Unique Games Conjecture of Khot [12] states that:

Conjecture II.4. *For every constant $\gamma > 0$, there is a sufficiently large constant L such that, for Unique Games instances Λ with label set $[L]$ it is* NP*-hard to distinguish between*

- $\mathrm{Opt}(\Lambda) \geq 1 - \gamma$
- $\mathrm{Opt}(\Lambda) \leq \gamma$.

Khot and Regev [3] proved that the following stronger version of UGC is in fact equivalent to UGC (the stronger version is necessary for proving the inapproximability of Vertex Cover and Independent Set):

Conjecture II.5. *For every constant $\gamma > 0$, there is a sufficiently large constant L such that, for Unique Games instances Λ with label set $[L]$ it is* NP*-hard to distinguish between*

- $\mathrm{IndOpt}(\Lambda) \geq 1 - \gamma$
- $\mathrm{Opt}(\Lambda) \leq \gamma$.

Moreover Λ is regular, i.e. all the left (resp. right) vertices have the same degree.

Now we define what we mean by a problem being UG-hard. We present a definition for a maximization problem; a similar definition can be made for a minimization problem.

Definition II.6. For a maximization problem $\mathcal{P}$, let $Gap\mathcal{P}_{c,s}$ denote its promise version where a given instance $\mathcal{I}$ is guaranteed to satisfy either $\mathrm{Opt}(\mathcal{I}) \geq c$ or $\mathrm{Opt}(\mathcal{I}) \leq s$. We say that $Gap\mathcal{P}_{c,s}$ is UG-hard if there is a polynomial time reduction from a Unique Games instance Λ to a $Gap\mathcal{P}_{c,s}$ instance $\mathcal{I}$, such that for some $\gamma > 0$,

- $\mathrm{Opt}(\Lambda) \geq 1 - \gamma \implies \mathrm{Opt}(\mathcal{I}) \geq c$.
- $\mathrm{Opt}(\Lambda) \leq \gamma \implies \mathrm{Opt}(\mathcal{I}) \leq s$.

In this case, we also say that $\mathcal{P}$ is UG-hard to approximate to within ratio better than c/s.

B. Influence, Noise, and Stability

For $p \in [0, 1]$, we use $\{0,1\}^n_{(p)}$ to denote the n-dimensional boolean hypercube with the p-biased product distribution, i.e., if x is a sample from $\{0,1\}^n_{(p)}$ then the probability that the i'th coordinate $x_i = 1$ is p, independently for each $i \in [n]$. Whenever we have a function $f : \{0,1\}^n_{(p)} \to \mathbb{R}$ we think of it as a random variable and hence expressions like $\mathbb{E}[f]$, $\mathrm{Var}[f]$, etc, are interpreted as being with respect to the p-biased distribution.

Definition II.7. The *influence* of the i'th variable on $f : \{0,1\}^n_{(p)} \to \mathbb{R}$ is given by

$$\mathrm{Inf}_i(f) = \mathop{\mathbb{E}}_{(x_j)_{j \neq i}} \left[\mathop{\mathrm{Var}}_{x_i}[f(x) \,|\, (x_j)_{j \neq i}] \right]$$

Definition II.8. Let $p \in (0, 1/2]$ and $\rho \in [-p/(1-p), 1]$. The *Beckner operator* T_ρ acts on functions $f : \{0,1\}^n_{(p)} \to \mathbb{R}$ by

$$T_\rho f(x) = \mathop{\mathbb{E}}_{y}[f(y)],$$

where each bit y_i of y has the following distribution, independently of the other bits: If $x_i = 1$, then $y_i = 1$ with probability $p + \rho(1-p)$. If $x_i = 0$, then $y_i = 1$ with probability $p - \rho p$.

We will use the following basic fact about the number of influential variables of $T_\rho f$.

Fact II.9. Let $f : \{0,1\}^n_{(p)} \to \mathbb{R}$ and $\rho \in [-p/(1-p), 1]$. Then, the number of $i \in [n]$ such that

$$\mathrm{Inf}_i(T_\rho f) \geq \tau$$

is at most $\frac{\mathrm{Var}[f]}{\tau e \ln(1/|\rho|)}$.

Finally, we have the notion of noise stability.

Definition II.10. Let $f : \{0,1\}^n_{(p)} \to \mathbb{R}$ for $p \leq 1/2$, and $\rho \in [-p/(1-p), 1]$. The *noise stability* of f at ρ is given by

$$\mathbb{S}_\rho(f) = \mathbb{E}[f \cdot T_\rho f].$$

Alternatively, one can write $\mathbb{S}_\rho(f) = \mathbb{E}[f(x)f(y)]$, where the distribution of the pair of bits (x_i, y_i) is given by $\Pr[x_i = 1] = \Pr[y_i = 1] = p$, and $\Pr[x_i = y_i = 1] = p \cdot (p + \rho(1-p)) \in [0, p]$, independently for each $i \in [n]$.

C. Gaussian Stability Bounds

Definition II.11. Let $\rho \in [-1, 1]$. We define $\Gamma_\rho : [0,1] \to [0,1]$ by

$$\Gamma_\rho(\mu) = \Pr\left[X \leq \Phi^{-1}(\mu) \wedge Y \leq \Phi^{-1}(\mu)\right]$$

where X and Y are jointly normal random variables with mean 0 and covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

We will use the following "Thresholds are Stablest" type of corollary of the MOO Theorem [11]. The formulation that we use here is equivalent to e.g. the formulation that is used in [13].

Theorem II.12. *For every $p \in (0, 1/2)$, $\rho \in [-p/(1-p), 1)$ and $\varepsilon > 0$ there exist $\tau > 0$ and $\delta > 0$ such that the following holds for every n: let $f : \{0,1\}^n_{(p)} \to [0,1]$ be a function with*

$$\mathrm{Inf}_i(T_{1-\delta} f) \leq \tau$$

for each $i \in [n]$. Then

$$\mathbb{S}_\rho(f) \geq \Gamma_\rho(\mathbb{E}[f]) - \varepsilon.$$

We will need asymptotic estimates of $\Gamma_\rho(\mu)$ for small μ, in particular good lower bounds. Several such estimates can be found in the literature (see e.g. [14], [15]), but we need bounds for the case where ρ tends to -1 as μ tends to 0, whereas the bounds we are aware of are stated only for fixed $\rho \in (-1, 1)$ or ρ tending to 1 with μ. Thus, for the sake of completeness, we give a proof of the following Lemma in Appendix A.

Lemma II.13. *Let $\rho := \rho(\mu)$ be such that $-1 < \rho < 0$. Then, there is a $\mu_0 > 0$ such that for all $\mu < \mu_0$*

$$\Gamma_\rho(\mu) \geq \frac{1}{2}\mu^{2/(1+\rho)}(1+\rho)^{3/2}.$$

III. Main Theorem

In this section, we give the main theorem upon which our results are based.

Theorem III.1. *Fix $p \in (0, 1/2)$, $\beta \in (0,1)$, and $\varepsilon > 0$. Then for all sufficiently small $\gamma > 0$, there is an algorithm which, on input a Unique Games instance $\Lambda = (U, V, E, \Pi, L)$ outputs a weighted graph G with the following properties:*

- *Completeness: If $\mathrm{IndOpt}(\Lambda) \geq 1 - \gamma$, G has an independent set of weight $p - \gamma$.*
- *Soundness: If $\mathrm{Opt}(\Lambda) \leq \gamma$ and Λ is regular, then G is $(\beta, \Gamma_\rho(\beta) - \varepsilon)$-dense where $\rho = -p/(1-p)$.*

The running time of the algorithm is polynomial in $|U|$, $|V|$, $|E|$ and exponential in L.

Proof: Let $\nu : \{0,1\}^2 \to \mathbb{R}$ be the probability distribution on $\{0,1\}^2$ such that $\Pr[x_1 = 1] = \Pr[x_2 = 1] = p$, and $\Pr[x_1 = x_2 = 1] = 0$. Let D be the degree of every vertex in U in the Unique Games instance Λ.

The reduction is as follows: the vertex set of G is $V \times \{0,1\}^L$. For every pair of edges $e_1 = (u, v_1)$, $e_2 = (u, v_2)$ with the same endpoint u in U, and every $x, y \in \{0,1\}^L$, there is an edge in G between (v_1, x) and (v_2, y) with weight

$$\frac{1}{|U| \cdot D^2} \cdot \prod_{i=1}^{L} \nu(x_{\pi_{e_1}(i)}, y_{\pi_{e_2}(i)}),$$

whenever this quantity is non-zero. Note that the sum of all edge weights equals 1. Let the weight of a vertex (v, x) be $\frac{1}{|V|}$ times the probability mass of $x \in \{0,1\}^L_{(p)}$ under the p-biased distribution. Thus the sum of all vertex weights equals 1. Note also that the marginal of the the distribution $\nu(,)$ (used to define the weights on edges) on either co-ordinate is the p-biased distribution on $\{0,1\}$ (used to define the weights on vertices). Therefore, the weight of every vertex is exactly $\frac{1}{2}$ times the sum of the weights of the edges incident on it. It is clear that the running time of the reduction is as stated, so it remains to see that the the reduction has the desired completeness and soundness properties.

Completeness: Suppose there is a subset $V' \subseteq V$ with relative size $1 - \gamma$ and a labeling $\ell : U \cup V' \to [L]$ that satisfies every edge between U and V' in the Unique Games instance Λ.

Consider the set of vertices $S = \{(v, x) : v \in V', x_{\ell(v)} = 1\} \subseteq V(G)$. Its weight is $w(S) \geq (1-\gamma) \cdot p \geq p - \gamma$. We claim that S is an independent set. To see this, assume for contradiction that G has an edge between $(v_1, x) \in S$ and $(v_2, y) \in S$. Then there is a $u \in U$ and edges $e_1 = (u, v_1)$, $e_2 = (u, v_2)$ such that $\ell(v_1) = \pi_{e_1}(\ell(u))$ and $\ell(v_2) = \pi_{e_2}(\ell(u))$. But then $\nu(x_{\pi_{e_1}(\ell(u))}, y_{\pi_{e_2}(\ell(u))}) = \nu(x_{\ell(v_1)}, x_{\ell(v_2)}) = \nu(1,1) = 0$, contradicting the assumption that (v_1, x) and (v_2, y) are connected by an edge in G.

Soundness: Let $S \subseteq V(G)$ have $w(S) = \beta$. We will prove that if $w(S, S)$ is even slightly smaller than $\Gamma_\rho(\beta)$, then $\mathrm{Opt}(\Lambda)$ must be significantly large.

For $v \in V$, let $S_v : \{0,1\}^L_{(p)} \to \{0,1\}$ be the indicator function of S restricted to v, i.e., $S_v(x) = 1$ if $(v, x) \in S$, and $S_v(x) = 0$ otherwise. For $u \in U$, define $S_u : \{0,1\}^L_{(p)} \to [0,1]$ by

$$S_u(x) = \mathop{\mathbb{E}}_{e=(u,v) \in E(u)}[S_v(x \circ \pi_e)],$$

where $E(u)$ are the set of edges incident upon u in Λ. Now, the weight $w(S, S)$ can be written as

$$\begin{aligned} w(S,S) & \\ &= \mathop{\mathbb{E}}_{\substack{u \in U \\ e_1, e_2 \in E(u)}}\left[\mathop{\mathbb{E}}_{(x,y) \sim \nu^{\otimes L}}[S_{v_1}(x \circ \pi_{e_1}) S_{v_2}(y \circ \pi_{e_2})]\right] \\ &= \mathop{\mathbb{E}}_{u \in U}\left[\mathop{\mathbb{E}}_{x,y}[S_u(x) S_u(y)]\right] \\ &= \mathop{\mathbb{E}}_{u \in U}[\mathbb{S}_\rho(S_u)], \end{aligned} \tag{1}$$

where $\rho := \rho(p) = -p/(1-p)$ (since this is the correlation coefficient between the bits x_i and y_i under the distribution ν).

Let $\mu_u = \mathbb{E}_x[S_u(x)]$. The regularity of Λ implies that

$$\mathop{\mathbb{E}}_{u \in U}[\mu_u] = \beta.$$

Suppose that for a fraction $\geq 1 - \varepsilon/2$ of all $u \in U$ it is the case that $\mathbb{S}_\rho(S_u) \geq \Gamma_\rho(\mu_u) - \varepsilon/2$. If this holds, we have that

$$w(S,S) \geq \mathop{\mathbb{E}}_{u \in U}[\Gamma_\rho(\mu_u)] - \varepsilon \geq \Gamma_\rho(\mathop{\mathbb{E}}_{u}[\mu_u]) - \varepsilon = \Gamma_\rho(\beta) - \varepsilon,$$

where the second inequality follows from the fact that Γ_ρ is convex[2].

Hence, if $w_G(S, S) \leq \Gamma_\rho(\beta) - \varepsilon$, there must be a set $U^* \subseteq U$ of size at least $|U^*| \geq \varepsilon|U|/2$, such that for every $u \in U^*$ it holds that $\mathbb{S}_\rho(S_u) < \Gamma_\rho(\mu_u) - \varepsilon/2$. By Theorem II.12 (with parameters p, $\rho(p)$ and $\varepsilon/2$, applied to the function S_u) we conclude that for each $u \in U^*$ there exists an $i \in [L]$ such that $\mathrm{Inf}_i(T_{1-\delta} S_u) \geq \tau$ for some $\tau > 0$, $\delta > 0$ depending only on p and ε. Since S_u is the average of functions $\{S_v \mid e = (u, v) \in E(u)\}$ (via appropriate π_e), for at least $\tau/2$ fraction of neighbors v of u, there must be $j = j(u, v) \in [L]$ such that $\pi_e(i) = j$ and $\mathrm{Inf}_j(T_{1-\delta} S_v) \geq \tau/2$.

[2]See e.g. the full version of [13] — the definition of Γ_ρ there differs slightly from the one used here, but only by an affine transformation of the input argument, and this does not affect convexity.

Now, define for every $v \in V$, a candidate set of labels to be the set of all $b \in [L]$ such that $\mathrm{Inf}_b(T_{1-\delta}S_v) \geq \tau/2$. By Fact II.9, this set has size at most $\frac{1}{\frac{\tau}{2}e\ln(1/(1-\delta))}$. Finally, pick one label at random from this set to be the label of $v \in V$, and for every $u \in U$, let its label be the projection of one of its randomly selected neighbor. From the preceding discussion, it follows that this randomized labeling satisfies, in expectation, at least $\Omega\left(\varepsilon\tau^4\ln^2(1/(1-\delta))\right)$ fraction of the edges of the Unique Games instance. This is a contradiction if the soundness γ of the Unique Games instance was chosen to be sufficiently small to begin with.

■

IV. Postprocessing

Note that in the soundness case of Theorem III.1, we obtain a graph that is $(\beta, \Gamma_\rho(\beta)-\varepsilon)$-dense. In particular, there is no independent set of weight β as long as $\Gamma_\rho(\beta) > \varepsilon$. The graph is both vertex-weighted as well as edge-weighted. In this section, we show that we can make the graph unweighted (in other words, weights are uniform) and then sparsify it so that the degree is bounded by d, preserving the maximum size of the independent set during the process. In particular, we have:

Theorem IV.1. *Let $p \in (0, 1/2)$. Then, for every sufficiently small $\beta > 0$ it is* UG*-hard (under randomized reductions) to distinguish graphs with an independent set of size $p-\beta$ from graphs with no independent set of size 2β, even on graphs of maximum degree $\frac{32\beta\log(1/\beta)}{\Gamma_\rho(\beta)}$, where $\rho = -p/(1-p)$.*

The proof of this theorem follows by the three steps outlined in the following three sections.

A. Removing Vertex and Edge Weights

First, we apply Theorem III.1 with parameter $\varepsilon = \Gamma_\rho(\beta)/2$ giving a weighted graph G_0. We assume w.l.o.g. that $\gamma < \beta$. In the completeness case, G_0 has an independent set of size $p-\gamma \geq p-\beta$. In the soundness case,

1) G_0 is $(\beta, \Gamma_\rho(\beta)/2)$-dense.
2) The sum of weights of edges incident upon any vertex is proportional to the weight of that vertex.

We now remove the vertex weights. We replicate every vertex so that the number of its copies is proportional to its weight. If $\{u_i\}_{i=1}^r$ and $\{v_j\}_{j=1}^s$ are copies of vertices u and v respectively, and (u, v) is an edge of the original graph, then we introduce an edge between every pair (u_i, v_j) and distribute the weight of the edge (u, v) evenly among the new $r \cdot s$ edges. It can easily be verified that if the original graph is $(\beta, \Gamma_\rho(\beta)/2)$-dense, then so is the new graph. Property (2) above implies that in the new graph, the weight of edges incident on every vertex is exactly the same. We then remove edge-weights, by simply replacing each edge by a number of parallel edges proportional to its weight. This yields an un-weighted graph G_1 with the same density properties as G_0 except that it is un-weighted and regular (though its degree is unbounded).

B. Sparsification

Let n be the number of vertices of the graph G_1 constructed in the previous section. We now construct a new graph G_2 by picking dn edges of G_1 at random (with repetition). If G_1 is (β, α)-dense (in our application, $\alpha = \Gamma_\rho(\beta)/2$), then the probability that G_2 has an independent set of size βn is bounded by

$$\binom{n}{\beta n}(1-\alpha)^{dn} \leq e^{n(2\beta\ln(1/\beta)-d\alpha)},$$

so that if $d > \frac{2\beta\ln(1/\beta)}{\alpha}$ (say, $d = 4\beta\log(1/\beta)/\alpha$), w.h.p. G_2 does not have any independent set of size βn.

C. Small Average Degree To Bounded Degree

In the sparsification step, we pick dn edges of G_1 at random. This yields a graph G_2 with average degree $2d$. Call a vertex bad if it has degree more than $4d$. It can be easily shown, using the regularity of G_1 and Chernoff bounds, that the probability of a vertex being bad is $2^{-\Omega(d)}$, and hence with constant probability the fraction of bad vertices is at most $2^{-\Omega(d)}$. In our choice of parameters, it holds that $2^{-\Omega(d)} \ll \beta$. We remove all edges of G_2 that are incident upon a bad vertex, giving a graph G_3. It is clear that the maximum degree of G_3 is bounded by $4d$, that the independence number of G_3 is at least that of G_2, and that, with constant probability the independence number of G_3 is at most $2^{-\Omega(d)}$ larger than that of G_2. In particular, if G_0 was $(\beta, \Gamma_\rho(\beta)/2)$-dense, then with constant probability it holds that G_3 does not contain any independent set of size 2β, whereas if G_0 had an independent set of weight $p-\beta$, then so does G_3.

V. Choice of Parameters

In this section, we show how to choose the parameters appropriately, so as to achieve Theorems I.1 and I.2.

A. Vertex Cover

We will use Theorem IV.1 with parameters chosen as follows. Let $p = 1/2 - \delta$, where δ is chosen so that $(2\delta)^{-1} = \frac{\log d}{\log\log d} - c$ for a sufficiently large constant c (e.g. $c = 10$ suffices) and $\beta = 1/\log d$. The inapproximability we get for Vertex Cover is then

$$\begin{aligned}\frac{1-2\beta}{1-(p-\beta)} &= \frac{2-4\beta}{1+2\delta+2\beta} \\ &\leq 2-4\delta+O(\beta+\delta^2) \\ &= 2-(2+o_d(1))\frac{\log\log d}{\log d},\end{aligned}$$

in graphs with maximum degree $32\beta\log(1/\beta)/\Gamma_\rho(\beta)$. It remains to see that this maximum degree is at most d. Using Lemma II.13 to approximate $\Gamma_\rho(\beta)$, we have that

$$\begin{aligned}\Gamma_\rho(\beta) &\geq \frac{1}{2}\beta^{2/(1+\rho)}(1+\rho)^{3/2} \\ &= 4\beta^{\frac{1}{2\delta}+1}(\delta/(1+2\delta))^{3/2} \\ &\geq \beta^{\frac{1}{2\delta}+1}\delta^{3/2},\end{aligned}$$

The maximum degree is then bounded by

$$\frac{32\log(1/\beta)}{\beta^{1/(2\delta)}\delta^{3/2}} = d\cdot(\log d)^{-c}\cdot\operatorname{poly}\log d,$$

which is at most d if c is a sufficiently large constant.

B. Independent Set

For Independent Set, we use Theorem IV.1 with the following choices of parameters: $p=\Theta(1/\log d)$ and $\beta=\Theta(\log d/d)$. We then get a hardness of approximating Independent set within

$$\frac{p-\beta}{2\beta} = \Theta\left(\frac{d}{\log^2 d}\right),$$

in graphs of maximum degree $32\beta\log(1/\beta)/\Gamma_\rho(\beta)$. Again using Lemma II.13 to estimate this quantity, we have

$$\begin{aligned}\Gamma_\rho(\beta) &\geq \frac{1}{2}\beta^{2/(1+\rho)}(1+\rho)^{3/2} \\ &= \beta^{2+\Theta(1/\log(d))}\cdot\Theta(1) \\ &= \Theta(\beta^2).\end{aligned}$$

Hence, the maximum degree is at most

$$\frac{32\beta\log(1/\beta)}{\Gamma_\rho(\beta)} \leq \Theta\left(\frac{\log(1/\beta)}{\beta}\right).$$

Making sure that β is a sufficiently large multiple of $\log(d)/d$, we see that the maximum degree becomes bounded by d.

ACKNOWLEDGMENT

Per Austrin was funded by ERC Advanced investigator grant 226203 and a grant from the Mittag-Leffler Institute. Subhash Khot was funded by NSF CAREER and NSF Expedition grants.

REFERENCES

[1] E. Halperin, “Improved Approximation Algorithms for the Vertex Cover Problem in Graphs and Hypergraphs,” *SIAM J. Comput.*, vol. 31, no. 5, pp. 1608–1623, 2002.

[2] A. Samorodnitsky and L. Trevisan, “Gowers Uniformity, Influence of Variables, and PCPs,” in *ACM Symposium on Theory of Computing (STOC)*, 2006, pp. 11–20.

[3] S. Khot and O. Regev, “Vertex Cover Might be Hard to Approximate to Within $2-\varepsilon$,” *Journal of Computer and System Sciences*, vol. 74, no. 3, pp. 335–349, 2008.

[4] I. Dinur and S. Safra, “On the Hardness of Approximating Minimum Vertex Cover,” *Annals of Mathematics*, vol. 162, no. 1, pp. 439–485, 2005.

[5] M. M. Halldórsson, “Approximations of Weighted Independent Set and Hereditary Subset Problems,” in *COCOON*, 1999, pp. 261–270.

[6] A. Samorodnitsky and L. Trevisan, “A PCP characterization of NP with optimal amortized query complexity,” in *ACM Symposium on Theory of Computing (STOC)*, 2000, pp. 191–199.

[7] U. Feige, “Approximating Maximum Clique by Removing Subgraphs,” *SIAM J. Discrete Math.*, vol. 18, no. 2, pp. 219–225, 2004.

[8] J. Håstad, “Clique is hard to approximate within $n^{1-\varepsilon}$,” *Acta Mathematica*, vol. 48, no. 4, pp. 105–142, 1999.

[9] S. Khot and A. K. Ponnuswami, “Better Inapproximability Results for MaxClique, Chromatic Number and Min-3Lin-Deletion,” in *International Colloquium on Automata, Languages and Programming (ICALP)*, 2006, pp. 226–237.

[10] I. Dinur, E. Mossel, and O. Regev, “Conditional Hardness for Approximate Coloring,” in *ACM Symposium on Theory of Computing (STOC)*, 2006, pp. 344–353.

[11] E. Mossel, R. O’Donnell, and K. Oleszkiewicz, “Noise Stability of Functions With Low Influences: Invariance and Optimality,” 2009, to appear in Annals of Mathematics.

[12] S. Khot, “On the Power of Unique 2-prover 1-round Games,” in *ACM Symposium on Theory of Computing (STOC)*, 2002, pp. 767–775.

[13] P. Austrin, “Balanced Max 2-Sat Might Not be the Hardest,” in *ACM Symposium on Theory of Computing (STOC)*, 2007, pp. 189–197.

[14] E. de Klerk, D. V. Pasechnik, and J. P. Warners, “On Approximate Graph Colouring and MAX-k-CUT Algorithms Based on the ϑ-Function,” *Journal of Combinatorial Optimization*, vol. 8, no. 3, pp. 267–294, 2004.

[15] S. Khot, G. Kindler, E. Mossel, and R. O’Donnell, “Optimal Inapproximability Results for MAX-CUT and Other 2-variable CSPs?” *Siam Journal on Computing*, vol. 37, pp. 319–357, 2007.

APPENDIX

In this section, we use $A(x) \overset{x \to y}{\sim} B(x)$ to denote that the ratio between $A(x)$ and $B(x)$ tends to 1 as x tends to y.

Proof of Lemma II.13: Using the standard bound $\Phi(x) \overset{x \to -\infty}{\sim} -\phi(x)/x$, Lemma A.1 below implies that

$$\begin{aligned}
\Gamma_\rho(\mu) &\overset{\mu \to 0}{\sim} \frac{(1+\rho)\cdot\phi(t)\cdot\phi\left(t\sqrt{\frac{1-\rho}{1+\rho}}\right)}{(-t)^2\sqrt{\frac{1-\rho}{1+\rho}}} \\
&= \sqrt{\frac{(1+\rho)^3}{1-\rho}}\left(\frac{\phi(t)}{-t}\right)^{\frac{2}{1+\rho}}\left(-\sqrt{2\pi}t\right)^{-\frac{2\rho}{1+\rho}} \\
&\geq \sqrt{\frac{(1+\rho)^3}{1-\rho}}\mu^{2/(1+\rho)} \\
&\geq \frac{1}{\sqrt{2}}\mu^{2/(1+\rho)}(1+\rho)^{3/2}
\end{aligned}$$

where the first inequality used the bound $-\phi(t)/t > \Phi(t) = \mu$ and simply discarded the last factor as it is larger than 1, and the second inequality used that $\frac{1}{1-\rho} \geq 1/2$. It follows that for sufficiently small μ, $\Gamma_\rho(\mu) \geq \frac{1}{2}\mu^{2/(1+\rho)}(1+\rho)^{3/2}$. ■

We will use the following Lemma to prove Lemma II.13, which is well-known in the case of fixed $\rho \in (-1,1)$, but as we mentioned earlier, we are not aware of any reference for the case when ρ is not bounded away from -1, and hence we also give a (straightforward but slightly tedious) proof.

Lemma A.1. *For any $-1 < \rho := \rho(\mu) \leq 0$, it holds that*

$$\Gamma_\rho(\mu) \overset{\mu \to 0}{\sim} (1+\rho)\frac{\phi(t)}{-t}\Phi\left(t\sqrt{\frac{1-\rho}{1+\rho}}\right),$$

where $t := t(\mu) = \Phi^{-1}(\mu)$.

Proof: We can write

$$\Gamma_\rho(\mu) = \int_{x=-\infty}^{t}\phi(x)\Phi\left(\frac{t-\rho x}{\sqrt{1-\rho^2}}\right)dx.$$

Since $\frac{t-\rho x}{\sqrt{1-\rho^2}}$ tends to $-\infty$ as μ tends to 0, we have

$$\Gamma_\rho(\mu) \overset{\mu \to 0}{\sim} \int_{x=-\infty}^{t}\phi(x)\frac{\phi\left(\frac{t-\rho x}{\sqrt{1-\rho^2}}\right)}{(\rho x - t)/\sqrt{1-\rho^2}}dx.$$

But $\phi(x)\phi\left(\frac{t-\rho x}{\sqrt{1-\rho^2}}\right) = \phi(t)\phi\left(\frac{x-\rho t}{\sqrt{1-\rho^2}}\right)$ and hence

$$\Gamma_\rho(\mu) \overset{\mu \to 0}{\sim} \phi(t)\sqrt{1-\rho^2}\int_{x=-\infty}^{t}\frac{\phi\left(\frac{x-\rho t}{\sqrt{1-\rho^2}}\right)}{\rho x - t}dx \quad (2)$$

Let us denote the integral in (2) by $f(\mu)$. Performing the change of variables $y = \frac{x-\rho t}{\sqrt{1-\rho^2}}$, we can simplify and obtain

$$\begin{aligned}
f(\mu) &= \int_{y=-\infty}^{t\sqrt{\frac{1-\rho}{1+\rho}}}\frac{\phi(y)}{\rho y - t\sqrt{1-\rho^2}}dy \\
&= \int_{y=-\infty}^{t'}\frac{\phi(y)}{\rho y - t'(1+\rho)}dy,
\end{aligned}$$

where we defined $t' = t\sqrt{\frac{1-\rho}{1+\rho}}$. We will show that $f(\mu) \overset{\mu \to 0}{\sim} \frac{\Phi(t')}{-t'}$. It is easy to see that this is an upper bound on $f(\mu)$ (by using the simple lower bound on the denumerator of the integrand given by $y = t'$), so let us focus on the lower bound.

Pick $\varepsilon > 0$. We then have

$$\begin{aligned}
f(\mu) &\geq \int_{y=t'(1+\varepsilon)}^{y=t'}\frac{\phi(y)}{\rho y - (1+\rho)t'}dy \\
&\geq \frac{\Phi(t') - \Phi(t'(1+\varepsilon))}{-t'(1-\rho\varepsilon)}
\end{aligned}$$

Using the fact that for $\alpha > 1$ and sufficiently small x, $\Phi(\alpha x) \leq \Phi(x)^\alpha$, we see that

$$f(\mu) \geq \frac{1-\Phi(t')^\varepsilon}{1-\rho\varepsilon}\cdot\frac{\Phi(t')}{-t'}.$$

As $\varepsilon > 0$ was arbitrary it follows that $f(\mu) \overset{\mu \to 0}{\sim} -\Phi(t')/t'$ (using that since $t' \leq t$, $t' \to -\infty$ as $\mu \to 0$). Plugging this into (2), we obtain

$$\begin{aligned}
\Gamma_\rho(\mu) &\overset{\mu \to 0}{\sim} \phi(t)\sqrt{1-\rho^2}\frac{\Phi(t'}{-t'} \\
&= (1+\rho)\frac{\phi(t)}{-t}\Phi\left(t\sqrt{\frac{1-\rho}{1+\rho}}\right),
\end{aligned}$$

which concludes the proof. ■