THEORY OF COMPUTING
www.theoryofcomputing.org

Inapproximability of Vertex Cover and Independent Set in Bounded Degree Graphs

Per Austrin* Subhash Khot† Muli Safra‡

September 23, 2010

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 [10] 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 [21]. Additionally, our result does not rely on the construction of a query efficient PCP as in [21].

1 Introduction

Vertex Cover and Independent Set are two of the most well-studied NP-complete problems. Recall that an independent set of a graph is a set of vertices no two of which are connected by an edge, and that

*Work done while at KTH – Royal Institute of Technology in Stockholm, funded by ERC Advanced investigator grant 226203 and a grant from the Mittag-Leffler Institute.

†Research funded by NSF CAREER and NSF Expedition grants.

‡Research partially supported by a bi-national US-Israeli BSF grant, and by an ISF grant.

ACM Classification: F.2

AMS Classification: 68Q17

Key words and phrases: vertex cover, independent set, bounded degree, hardness, approximation, unique games conjecture

a vertex cover is a set of vertices that contains at least one endpoint of every edge (or equivalently, the complement of an independent set). 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 (see Section 2.1 for definitions of Unique Games problem and UG-hardness of a problem).

Our Results

For the Vertex Cover problem, we prove:

Theorem 1.1. *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 [10] presents an efficient algorithm that approximates vertex cover in a degree d graph to within essentially the same factor, up to the value of the $o_d(1)$ error term. This improves on the general well-known 2-approximation ratio for graphs of unbounded degree. In general graphs, the best current approximation algorithm, due to Karakostas [12], has an approximation ratio of $2-\Omega(1/\sqrt{\log n})$. On the inapproximability side, Khot and Regev [16] showed $2-\varepsilon$ UG-hardness result for any constant $\varepsilon>0$, whereas Dinur and Safra [6] showed a 1.36 NP-hardness result.

For the Independent Set problem, we prove:

Theorem 1.2. *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 [10], or Halldórsson [9]). It is an intriguing question whether one can improve the approximation algorithm, or improve on this inapproximability result (or both). Previously, Samorodnitsky and Trevisan [21] 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 [20], 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 [11] (assuming $\mathsf{NP} \not\subseteq \mathsf{ZPP}$). 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 [15] (assuming NP does not have randomized quasi-polynomial time algorithms).

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. Also, we remark that the earlier result of Samorodnitsky and Trevisan [21], which 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.

Finally, we note that all known algorithms for the bounded degree case make no assumption that d is constant and work also for the case $d = n$. In particular the best algorithms for bounded degree Vertex Cover and Independent Set give algorithms for the general case with approximation ratios $2 - (2 - o(1))\frac{\log\log n}{\log n}$ and $O\left(\frac{n\log\log n}{\log n}\right)$, respectively. It is natural to ask whether this is inherent, i. e., whether the approximability for the unbounded degree case equals the approximability for the bounded degree case with the degree bound d set to n. Using the best current algorithms for the two problems, we see that our hardness results do not hold for $d = n$ (unless the UGC is false). Previously known hardness results were not strong enough to rule out this possibility.

Techniques

Showing an inapproximability 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 almost 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 1.1, and $\Theta(\frac{1}{\log d}) \cdot n$ for Theorem 1.2. 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 1.1 and $\beta = \Theta\left(\frac{\log d}{d}\right)$ for Theorem 1.2.

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 that not only does G have no independent set of size βn, but we also 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 construction is almost the same as in Khot and Regev's paper [16]. 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. [5] for showing UG-hardness for coloring problems. The density property follows from a quite straightforward application of a *Thresholds are Stablest* type theorem [19], 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.

2 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 2.1. 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.

2.1 Unique Games

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

Definition 2.2. 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 *labeling* $\ell : U \cup V \to [L]$ 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 2.3. 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 labeling, 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 $\ell : U \cup V' \to [L]$ 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 [13] can be stated states as follows:

Conjecture 2.4. For every $\gamma > 0$, there is an 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.

This formulation of the UGC differs from Khot's original formulation, but was proved to be equivalent by Khot and Regev [16] (this version is necessary for the known proofs of inapproximability of Vertex Cover and Independent Set).

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 2.5. For a maximization problem $\mathcal{P}$, let $Gap\mathcal{P}_{c,s}$ denote its promise version where every instance $\mathcal{I}$ is guaranteed to satisfy either $\mathrm{Opt}(\mathcal{I}) \geq c$ or $\mathrm{Opt}(\mathcal{I}) \leq s$ and the goal is to distinguish between the two. We say that $Gap\mathcal{P}_{c,s}$ is UG-hard if for some $\gamma > 0$ there is a polynomial time reduction mapping Unique Games instances Γ to $Gap\mathcal{P}_{c,s}$ instances $\mathcal{I}$ such that

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

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

Note that if the UGC holds then a problem being UG-hard is equivalent to the problem being NP-hard.

2.2 Influence, Noise, and Stability

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

Definition 2.6. The *influence* of the ith variable on $f : \{0,1\}^n_{(q)} \to \mathbb{R}$ is given by

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

Definition 2.7. Let $q \in (0,1/2]$ and $\rho \in [-q/(1-q),1]$. The *Beckner operator* T_ρ acts on functions $f : \{0,1\}^n_{(q)} \to \mathbb{R}$ 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 $q+\rho(1-q)$. If $x_i = 0$, then $y_i = 1$ with probability $q-\rho q$. In other words, the joint distribution of (x_i, y_i) is such that both coordinates are q-biased and that their correlation coefficient equals ρ.

Throughout the paper, we will in fact only be using the case $\rho = -q/(1-q)$. Note that for this choice of ρ, the joint distribution of (x_i, y_i) is as follows:

$$\Pr[x_i = y_i = 0] = 1-2q, \quad \Pr[x_i = 0 \wedge y_i = 1] = \Pr[x_i = 1 \wedge y_i = 0] = q, \quad \Pr[x_i = y_i = 1] = 0.$$

In particular the probability of having $x_i = y_i = 1$ equals 0.

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

Fact 2.8. Let $f : \{0,1\}^n_{(q)} \to \mathbb{R}$ and $\rho \in [-q/(1-q), 1]$. Then for every $\tau > 0$, the number of $i \in [n]$ such that

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

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

For a proof see e. g., Lemma 3.4 in [8]. That statement is for a somewhat different setting but the proof in our setting is identical.

Finally, we have the notion of noise stability.

Definition 2.9. Let $f : \{0,1\}^n_{(q)} \to \mathbb{R}$ for $q \leq 1/2$, and $\rho \in [-q/(1-q), 1]$. 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 $\Pr[x_i = 1] = \Pr[y_i = 1] = q$, and $\Pr[x_i = y_i = 1] = q \cdot (q + \rho(1-q)) \in [0, q]$, independently for each i.

2.3 Gaussian Stability Bounds

We use $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ and $\Phi(t) = \int_{x=-\infty}^{t} \phi(x) dx$ to denote the pdf and cdf of the standard normal distribution, respectively, and $\Phi^{-1} : [0,1] \to [-\infty, \infty]$ to denote the inverse of Φ.

Definition 2.10. 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 [19]. The formulation that we use here is equivalent to e. g., the formulation that is used in [1].

Theorem 2.11. *For every $q \in (0, 1/2)$, $\rho \in [-q/(1-q), 0)$ and $\varepsilon > 0$ there exist $\tau > 0$ and $\delta > 0$ such that the following holds for every n: let $f : \{0,1\}^n_{(q)} \to [0,1]$ 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.

Lemma 2.12. *For every sufficiently small $\mu > 0$ and every $-1 < \rho < 0$ it holds that*

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

Several similar estimates can be found in the literature (see e. g., [4, 14]), but we need bounds for the case where ρ is not bounded away from -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 now provide a proof of the lemma.

In the remainder of 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. In what follows we shall repeatedly use the standard bound $\Phi(x) \overset{x\to -\infty}{\sim} -\phi(x)/x$.

We use the following lemma which is well-known in the case of fixed $\rho \in (-1,1)$, but as with Lemma 2.12, 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 2.13. *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

$$\Gamma_\rho(\mu) \overset{\mu\to 0}{\sim} \phi(t)\sqrt{1-\rho^2}\int_{x=-\infty}^{t} \frac{\phi\left(\frac{x-\rho t}{\sqrt{1-\rho^2}}\right)}{\rho x - t}dx \tag{2.1}$$

Let us denote the integral in (2.1) by $f(\mu)$. Performing the change of variables $y = \frac{x-\rho t}{\sqrt{1-\rho^2}}$, we can simplify and obtain

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

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

Pick $\varepsilon > 0$. We then have

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

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 (2.1), we obtain

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

which concludes the proof. □

We are now ready to prove the lower bound on $\Gamma_\rho(\mu)$.

Proof of Lemma 2.12. Lemma 2.13 implies that

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

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

3 Main Theorem

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

Theorem 3.1. *Fix $q \in (0,1/2)$ 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 $q-\gamma$.*

- *Soundness: If* $\mathrm{Opt}(\Lambda) \leq \gamma$ *and* Λ *is regular, then* G *is* $(\beta, \Gamma_\rho(\beta) - \varepsilon)$*-dense for every* $\beta \in [0,1]$*, where* $\rho = -q/(1-q)$*.*

Furthermore, the weight of every vertex in G *is proportional to the sum of weights of its incident edges. 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 $\Pr[x_1 = 1] = \Pr[x_2 = 1] = q$, and $\Pr[x_1 = x_2 = 1] = 0$. Note that this distribution is exactly the joint distribution of two coordinates (x_i, y_i) in the Definition 2.9 of $\mathbb{S}_\rho(f)$ for $\rho = -q/(1-q)$.

For a string $x \in \{0,1\}^L$ and permutation π on $[L]$, let $x \circ \pi$ denote the string in $\{0,1\}^L$ where the ith coordinate is $x_{\pi(i)}$.

The reduction is as follows: the vertex set of G is $V \times \{0,1\}^L$. To describe the edges, we describe how to sample a random edge of the graph (the probability distribution induced on pairs of vertices of G by this sampling procedure give the weights of the edges):

1. Pick a uniformly random vertex $u \in U$ and two independent uniformly random edges $e_1 = (u, v_1)$, $e_2 = (u, v_2)$ incident upon u.

2. Pick random x, y in $\{0,1\}^L$ where each pair of coordinates (x_i, y_i) is sampled from ν, independently.

3. Output edge between $(v_1, x \circ \pi_{e_1}^{-1})$ and $(v_2, y \circ \pi_{e_2}^{-1})$.

Let the weight of a vertex (v, x) be $\frac{1}{|V|}$ times the probability mass of $x \in \{0,1\}^L_{(q)}$ under the q-biased distribution. Thus the sum of all vertex weights equals 1. Note also that the marginal of the the distribution $\nu(\cdot,\cdot)$ (used to define the weights on edges) on either coordinate is the q-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 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 $w(S) \geq (1-\gamma) \cdot q \geq q - \gamma$. 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 Fix an arbitrary $S \subseteq V(G)$ and let $\beta = w(S)$. 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 this part, let $E_\Lambda(u)$ denote the set of neighbors of u in Λ.

For $v \in V$, let $S_v : \{0,1\}^L_{(q)} \to \{0,1\}$ 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 $S_u : \{0,1\}^L_{(q)} \to [0,1]$ by $S_u(x) = \mathbb{E}_{e=(u,v) \in E_\Lambda(u)}[S_v(x \circ$

$\pi_e^{-1})]$. Now, the weight $w(S,S)$ can be written as

$$\begin{aligned} w(S,S) &= \mathop{\mathbb{E}}_{\substack{u\in U\\ e_1,e_2\in E_\Lambda(u)}}\left[\mathop{\mathbb{E}}_{(x,y)\sim\nu^{\otimes L}}[S_{v_1}(x\circ\pi_{e_1}^{-1})S_{v_2}(y\circ\pi_{e_2}^{-1})]\right] \\ &= \mathop{\mathbb{E}}_{u\in U}\left[\mathop{\mathbb{E}}_{x,y}[S_u(x)S_u(y)]\right] \\ &= \mathop{\mathbb{E}}_{u\in U}[\mathbb{S}_\rho(S_u)], \end{aligned} \tag{3.1}$$

where $\rho := \rho(q) = -q/(1-q)$ (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

$$w(S,S) \geq \mathop{\mathbb{E}}_{u\in U}[\Gamma_\rho(\mu_u)] - \varepsilon \geq \Gamma_\rho(\mathop{\mathbb{E}}_u[\mu_u]) - \varepsilon = \Gamma_\rho(\beta) - \varepsilon, \tag{3.2}$$

where the second inequality follows from the fact that Γ_ρ is convex.[1]

Hence, if $w(S,S) \leq \Gamma_\rho(\beta) - \varepsilon$, 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 2.11 (with parameters q, $\rho(q)$ 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 q and ε. Since S_u is the average of functions $\{S_v \mid e = (u,v) \in E_\Lambda(u)\}$ (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 $\mathrm{Inf}_j(T_{1-\delta}S_v) \geq \tau/2$ (here we used the well-known fact that $\mathrm{Inf}_i(\cdot)$ is convex).

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 2.8, 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 the label of a randomly selected neighbor.

To analyze the value of this labelling, let $u \in U^*$ and let $i \in [L]$ be a label such that $\mathrm{Inf}_i(T_{1-\delta}S_u) \geq \tau$. As mentioned above, for a $\tau/2$ fraction of edges $e = (u,v) \in \Lambda(u)$ we have $\mathrm{Inf}_{\pi_e(i)}(T_{1-\delta}(S_v)) \geq \tau/2$. Thus, for a random edge $e = (u,v) \in \Lambda(u)$, the probability that v is assigned the label $\pi_e(i)$ is $\Omega(\tau^2\ln(1/(1-\delta)))$ (since each of the $\tau/2$ fraction of "good" neighbors has a $\Omega(\tau\ln(1/(1-\delta)))$ probability of getting the right label). Note that this is also the probability that u gets the label i.

As U^* constitutes an ε fraction of U, 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.

□

[1] See e. g., the full version of [1] — 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.

4 Post-Processing

Note that in the soundness case of Theorem 3.1, 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 4.1. *For every sufficiently small $\beta > 0$ and every $q \in (0, 1/2)$ it is* UG*-hard (under randomized reductions) to distinguish graphs with an independent set of size $q - \beta$ 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 $\rho = -q/(1-q)$.*

Proof. We begin with applying Theorem 3.1 with parameter $\varepsilon = \Gamma_\rho(\beta)/2$ and $\gamma = \beta$, giving a weighted graph G_0. In the completeness case, G_0 has an independent set of size $q - \gamma = q - \beta$. In the soundness case,

1. G_0 is $(\beta', \Gamma_\rho(\beta') - \varepsilon)$-dense for every $\beta' \in [0, 1]$.
2. The sum of weights of edges incident upon any vertex is proportional to the weight of that vertex.

The process of converting G_0 to a graph of bounded degree with similar properties is done in three steps.

Step 1: Removing Vertex and Edge Weights First we remove the vertex weights. Without loss of generality, we may assume that each edge weight $w(e)$ is of the form $w'(e)/W$ for some integers $w'(e)$ and W, and similarly for the vertex weights. To achieve this, let $W = \text{poly}|V(G_0)|$ be large enough and round the edge weights accordingly. Then update the vertex weights so that they are still proportional to the weight of incident edges. This rounding can cause a difference of order $\text{poly}|V(G_0)|/W$ in the weights of vertex and edge sets, but this arbitrarily small error is easily handled by making ε and γ slightly smaller in the invocation of Theorem 3.1.

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. Call the new graph G'_0.

We claim that if the original graph G_0 is $(\beta', \Gamma_\rho(\beta') - \varepsilon)$-dense for every $\beta' \in [0, 1]$, then so is G'_0. To see this, consider a subset S' of vertices of G'_0, and construct a random subset S of vertices of G_0 where if a δ fraction of the copies of v are included in S', we include it in S with probability δ, independently. Note that $\mathbb{E}[w(S)] = w(S')$ and that $\mathbb{E}[w(S, S)] = w(S', S')$. Hence,

$$w(S', S') = \mathbb{E}[w(S, S)] \geq \mathbb{E}[\Gamma_\rho(w(S)) - \varepsilon] \geq \Gamma_\rho(\mathbb{E}[w(S)]) - \varepsilon = \Gamma_\rho(w(S')) - \varepsilon,$$

where the first inequality used that G is $(\beta', \Gamma_\rho(\beta') - \varepsilon)$-dense for *every* $\beta' \in [0, 1]$, and the second inequality used the convexity of Γ_ρ as in (3.2).

Property (2) of G_0 implies that in G'_0, the weight of edges incident on every vertex is exactly the same. We now remove edge weights, by simply replacing each edge by a number of parallel edges proportional to its weight. This yields an unweighted graph G_1 with the same density properties as G_0' except that it is unweighted and regular (though its degree is unbounded).

From now on, the only density property of G_1 that we will use is that, in the soundness case, G_1 is $(\beta, \Gamma_\rho(\beta)/2)$-dense.

Step 2: 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.

Step 3: 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, we have that $2^{-\Omega(d)} = \beta^{\Omega(\beta/\Gamma_\rho(\beta))} = \beta^{\Omega(1/\beta)} \ll \beta$ (where the last equality used that for $\rho \leq 0$ we have $\Gamma_\rho(\beta) < \beta^2$). 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 $q - \beta$, then so does G_3.

This concludes the proof of Theorem 4.1.

□

5 Choice of Parameters

In this section, we show how to choose the parameters appropriately, so as to achieve Theorems 1.1 and 1.2.

5.1 Vertex Cover

We will use Theorem 4.1 with parameters chosen as follows. Let $q = 1/2 - \delta$, 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

$$\frac{1-2\beta}{1-(q-\beta)} = \frac{2-4\beta}{1+2\delta+2\beta} \leq 2 - 4\delta + O(\beta + \delta^2) = 2 - (2 + o_d(1))\frac{\log \log d}{\log d},$$

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 2.12 to approximate $\Gamma_\rho(\beta)$, we have that

$$\Gamma_\rho(\beta) \geq \frac{1}{2}\beta^{2/(1+\rho)}(1+\rho)^{3/2} = 4\beta^{\frac{1}{2\delta}+1}(\delta/(1+2\delta))^{3/2} \geq \beta^{\frac{1}{2\delta}+1}\delta^{3/2}.$$

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.

5.2 Independent Set

For Independent Set, we use Theorem 4.1 with the following choices of parameters: $q = \Theta(1/\log d)$ and $\beta = \Theta(\log d/d)$. We then get a hardness of approximating Independent Set within

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

in graphs of maximum degree $32\beta\log(1/\beta)/\Gamma_\rho(\beta)$. Again using Lemma 2.12 to estimate this quantity, we have

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

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.

6 Concluding Remarks

It would be interesting to determine whether our methods can also yield tight results for the Vertex Cover problem in bounded degree k-uniform hypergraphs. Here the best algorithm, by Halperin [10], has an approximation ratio of $k-k(k-1)(1+o_d(1))\frac{\log\log d}{\log d}$. The main crux would be to find a suitable probability distribution on $\{0,1\}^k$ with small "correlation" (in the sense of [18]).

Another interesting open question is the approximability of Vertex Cover (and Independent Set) in degree d bounded graphs for small, concrete values of d (whereas our results are asymptotic in d). For instance, what is the approximability of Vertex Cover in cubic graphs? It seems that answering this question would require new ideas.

In independent and subsequent papers, Bansal and Khot [2, 3] gave new hardness results Vertex Cover in graphs (and hypergraphs though we'll restrict attention to graphs for this discussion). Their results are stronger in that, in the Yes case, the graphs are almost bipartite. However their results are weaker in that, in the No case, they do not get the strong density property that we crucially rely on (in the first paper they get no density at all and in the subsequent paper they get some density bound but it is not clear whether this density is enough to yield the tight inapproximability that we get).

In a recent manuscript, Kumar et al. [17] gives hardness results and algorithms for a large class of problems called strict-CSPs, which contains both the Vertex Cover and Independent Set problems. They relate the approximability to the integrality gap of a certain canonical linear programming relaxation. In their result, the approximation ratio of both the algorithm and hardness result have a common explanation – the existence of an integrality gap. In our result on the other hand, the matching ratios for Vertex Cover appear somewhat accidental (even though both appear because of certain normal distribution estimates). However, the results of [17] do not yield anything for bounded degree graphs and it is somewhat doubtful whether this can be done. Specifically, Halperin's algorithm for bounded degree Vertex Cover uses (and appears to need) the additional power of semidefinite programming so it is quite possible that the integrality gap for the linear program does not capture the optimal approximation ratio for the bounded degree case. In general, it remains an interesting open question to come up with a characterization which captures approximability of (strict or non-strict) constraint satisfaction problems in the bounded occurrence setting.

6.1 Acknowledgments

We are grateful to the anonymous referees for their many helpful comments.

References

[1] Per Austrin. Balanced Max 2-Sat Might Not be the Hardest. In *ACM Symposium on Theory of Computing (STOC)*, pages 189–197, 2007. Full version available as ECCC Report TR06-088. 6, 10

[2] Nikhil Bansal and Subhash Khot. Optimal Long Code Test with One Free Bit. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 453–462, 2009. 13

[3] Nikhil Bansal and Subhash Khot. Inapproximability of Hypergraph Vertex Cover and Applications to Scheduling Problems. In *International Colloquium on Automata, Languages and Programming (ICALP)*, pages 250–261, 2010. 13

[4] Etienne de Klerk, Dmitrii V. Pasechnik, and Joost P. Warners. On Approximate Graph Colouring and MAX-k-CUT Algorithms Based on the ϑ-Function. *Journal of Combinatorial Optimization*, 8(3):267–294, 2004. 7

[5] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional Hardness for Approximate Coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009. 4

[6] Irit Dinur and Shmuel Safra. On the Hardness of Approximating Minimum Vertex Cover. *Annals of Mathematics*, 162(1):439–485, 2005. 2

[7] Uriel Feige. Approximating Maximum Clique by Removing Subgraphs. *SIAM J. Discrete Math.*, 18(2):219–225, 2004. 2

[8] Venkatesan Guruswami, Rajsekar Manokaran, and Prasad Raghavendra. Beating the Random Ordering is Hard: Inapproximability of Maximum Acyclic Subgraph. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 573–582, 2008. 6

[9] Magnús M. Halldórsson. Approximations of Weighted Independent Set and Hereditary Subset Problems. In *COCOON*, pages 261–270, 1999. 2

[10] Eran Halperin. Improved Approximation Algorithms for the Vertex Cover Problem in Graphs and Hypergraphs. *SIAM Journal on Computing*, 31(5):1608–1623, 2002. 1, 2, 13

[11] Johan Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Mathematica*, 48(4):105–142, 1999. 2

[12] George Karakostas. A Better Approximation Ratio for the Vertex Cover Problem. In *International Colloquium on Automata, Languages and Programming (ICALP)*, pages 1043–1050, 2005. 2

[13] Subhash Khot. On the Power of Unique 2-prover 1-round Games. In *ACM Symposium on Theory of Computing (STOC)*, pages 767–775, 2002. 4

[14] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal Inapproximability Results for MAX-CUT and Other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007. 7

[15] Subhash Khot and Ashok Kumar Ponnuswami. Better Inapproximability Results for MaxClique, Chromatic Number and Min-3Lin-Deletion. In *International Colloquium on Automata, Languages and Programming (ICALP)*, pages 226–237, 2006. 2

[16] Subhash Khot and Oded Regev. Vertex Cover Might be Hard to Approximate to Within $2-\varepsilon$. *Journal of Computer and System Sciences*, 74(3):335–349, 2008. 2, 3, 5

[17] Amit Kumar, Rajsekar Manokaran, Madhur Tulsiani, and Nisheeth K. Vishnoi. On the Optimality of a Class of LP-based Algorithms. arXiv:0912.1776v1 [cs.CC], 2009. 14

[18] Elchanan Mossel. Gaussian Bounds for Noise Correlation of Functions and Tight Analysis of Long Codes. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 156–165, 2008. 13

[19] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise Stability of Functions With Low Influences: Invariance and Optimality. *Annals of Mathematics*, 171(1):295–341, 2010. 4, 6

[20] Alex Samorodnitsky and Luca Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *ACM Symposium on Theory of Computing (STOC)*, pages 191–199, 2000. 2

[21] Alex Samorodnitsky and Luca Trevisan. Gowers Uniformity, Influence of Variables, and PCPs. In *ACM Symposium on Theory of Computing (STOC)*, pages 11–20, 2006. 1, 2

AUTHORS

Per Austrin
University of Toronto, Canada
austrin@cs.toronto.edu
http://www.cs.toronto.edu/~austrin

Subhash Khot
New York University, USA
khot@cs.nyu.edu
http://cs.nyu.edu/~khot/

Muli Safra
Tel-Aviv University, Israel
safra@post.tau.ac.il
http://www.cs.tau.ac.il/~safra/

ABOUT THE AUTHORS

PER AUSTRIN graduated from KTH – Royal Institute of Technology in Sweden, 2008; his advisor was Johan Håstad. His thesis focused on hardness of approximation for constraint satisfaction problems with connections to discrete harmonic analysis.

SUBHASH KHOT is an Associate Professor in the Computer Science Department at New York University, part of the Courant Institute of Mathematical Sciences. He completed his Ph. D. in the summer of 2003 at the Princeton CS Department under the supervision of Prof. Sanjeev Arora. He was in Princeton for another year as a member of School of Mathematics, IAS and then an Assistant Professor at the College of Computing at Georgia Tech during Fall '04-Summer '07.

MULI SAFRA is a Professor of Computer Science at Tel Aviv University. He has a Ph. D. from the Weizmann Institute, completed in 1990 under the supervision of Amir Pnueli. He is one of the recipients of the 2001 Gödel Prize.