Integrality gaps of semidefinite programs for Vertex Cover and relations to ℓ_1 embeddability of negative type metrics*

Hamed Hatami Avner Magen Evangelos Markakis†

Department of Computer Science
University of Toronto

Abstract

We study various SDP formulations for Vertex Cover by adding different constraints to the standard formulation. We show that Vertex Cover cannot be approximated better than $2 - O(\sqrt{\log \log n / \log n})$ even when we add the so-called pentagonal inequality constraints to the standard SDP formulation, and thus almost meet the best upper bound known due to Karakostas, of $2 - \Omega(\sqrt{1/\log n})$. We further show the surprising fact that by strengthening the SDP with the (intractable) requirement that the metric interpretation of the solution embeds into ℓ_1 with no distortion, we get an exact relaxation (integrality gap is 1), and on the other hand if the solution is arbitrarily close to being ℓ_1 embeddable, the integrality gap is $2 - o(1)$. Finally, inspired by the above findings, we use ideas from the integrality gap construction of Charikar to provide a family of simple examples for negative type metrics that cannot be embedded into ℓ_1 with distortion better than $8/7 - \epsilon$. To this end we prove a new isoperimetric inequality for the hypercube.

1 Introduction

A vertex cover in a graph $G = (V, E)$ is a set $S \subseteq V$ such that every edge $e \in E$ intersects S in at least one endpoint. Denote by $\mathrm{vc}(G)$ the size of the minimum vertex cover of G. It is well-known that the minimum vertex cover problem has a 2-approximation algorithm, and it is widely believed that for every constant $\epsilon > 0$, there is no $(2-\epsilon)$-approximation algorithm for this problem. Currently the best known hardness result for this problem, based on the PCP theorem, shows that 1.36-approximation is NP-hard [10]. If we were to assume the Unique Games Conjecture [18] the problem would be essentially settled as $2 - \Omega(1)$ would then be NP-hard [19].

In [14], Goemans and Williamson introduced semidefinite programming as a tool for obtaining approximation algorithms. Since then semidefinite programming has become an important technique, and for many problems the best known approximation algorithms are obtained by solving an SDP relaxation of them.

The best known algorithms for Vertex Cover compete in "how big is the little oh" in the $2 - o(1)$ factor. The best two are in fact based on SDP relaxations: Halperin [15] gives a $(2 -$

*A preliminary version of this work appears in [16].

†Contact author. Currently at the Center for Math and Computer Science (CWI), Amsterdam, the Netherlands, email: vangelis@cwi.nl.

$\Omega(\log\log\Delta/\log\Delta)$)-approximation where Δ is the maximal degree of the graph while Karakostas obtains a $(2-\Omega(1/\sqrt{\log n}))$-approximation [17]. As we later show, our lower bound almost meets the latter upper bound even in this resolution of the little oh.

The standard way to formulate the Vertex Cover problem as a quadratic integer program is the following:

$$\begin{array}{lll} \textbf{Min} & \sum_{i\in V}(1+x_0x_i)/2 & \\ \text{s.t.} & (x_i-x_0)(x_j-x_0)=0 & \forall\ ij\in E \\ & x_i\in\{-1,1\} & \forall\ i\in\{0\}\cup V, \end{array}$$

where the set of the vertices i for which $x_i=x_0$ correspond to the vertex cover. Relaxing this integer program to a semidefinite program, the scalar variable x_i becomes a vector $\mathbf{v}_i$ and we get:

$$\begin{array}{lll} \textbf{Min} & \sum_{i\in V}(1+\mathbf{v}_0\mathbf{v}_i)/2 & \\ \text{s.t.} & (\mathbf{v}_i-\mathbf{v}_0)\cdot(\mathbf{v}_j-\mathbf{v}_0)=0 & \forall\ ij\in E \\ & \|\mathbf{v}_i\|=1 & \forall\ i\in\{0\}\cup V. \end{array} \tag{1}$$

Kleinberg and Goemans [21] proved that SDP (1) has integrality gap of $2-o(1)$. Specifically, given $\epsilon>0$ they construct a graph G_ϵ for which $\mathrm{vc}(G_\epsilon)$ is at least $(2-\epsilon)$ times larger than the solution to SDP (1). They also suggested the following strengthening of SDP (1) and left its integrality gap as an open question:

$$\begin{array}{lll} \textbf{Min} & \sum_{i\in V}(1+\mathbf{v}_0\mathbf{v}_i)/2 & \\ \text{s.t.} & (\mathbf{v}_i-\mathbf{v}_0)\cdot(\mathbf{v}_j-\mathbf{v}_0)=0 & \forall\ ij\in E \\ & (\mathbf{v}_i-\mathbf{v}_k)\cdot(\mathbf{v}_j-\mathbf{v}_k)\geq 0 & \forall\ i,j,k\in\{0\}\cup V \\ & \|\mathbf{v}_i\|=1 & \forall\ i\in\{0\}\cup V. \end{array} \tag{2}$$

Charikar [6] answered this question by showing that the same graph G_ϵ but a different vector solution satisfies SDP (2)[1] and gives rise to an integrality gap of $2-o(1)$ as before. The following is an equivalent formulation to SDP (2):

$$\begin{array}{lll} \textbf{Min} & \sum_{i\in V}1-\|\mathbf{v}_0-\mathbf{v}_i\|^2/4 & \\ \text{s.t.} & \|\mathbf{v}_i-\mathbf{v}_0\|^2+\|\mathbf{v}_j-\mathbf{v}_0\|^2=\|\mathbf{v}_i-\mathbf{v}_j\|^2 & \forall\ ij\in E \\ & \|\mathbf{v}_i-\mathbf{v}_k\|^2+\|\mathbf{v}_j-\mathbf{v}_k\|^2\geq\|\mathbf{v}_i-\mathbf{v}_j\|^2 & \forall\ i,j,k\in\{0\}\cup V \\ & \|\mathbf{v}_i\|=1 & \forall\ i\in\{0\}\cup V \end{array} \tag{3}$$

Viewing SDPs as relaxations over ℓ_1: The above reformulation reveals a connection to metric spaces. The second constraint in SDP (3) says that $\|\cdot\|^2$ induces a metric on $\{\mathbf{v}_i : i\in\{0\}\cup V\}$, while the first says that $\mathbf{v}_0$ is on the shortest path between the images of every two neighbours. This suggests a more careful study of the problem from the metric viewpoint which is the purpose of this article. Such connections are also important in the context of the Sparsest Cut problem, where the natural SDP relaxation was analyzed in the breakthrough work of Arora, Rao and Vazirani [5] and it was shown that its integrality gap is at most $O(\sqrt{\log n})$. This later gave rise to some significant progress in the theory of metric spaces [7, 4].

Let $f:(X,d)\to(X',d')$ be an embedding of metric space (X,d) into another metric space (X',d'). The value $\sup_{x,y\in X}\frac{d'(f(x),f(y))}{d(x,y)}\times\sup_{x,y\in X}\frac{d(x,y)}{d'(f(x),f(y))}$ is called the distortion of f. For a

[1] To be more precise, Charikar's result was about a slightly weaker formulation than (2) but it is not hard to see that the same construction works for SDP (2) as well.

metric space (X, d), let $c_1(X, d)$ denote the minimum distortion required to embed (X, d) into ℓ_1. Notice that $c_1(X, d) = 1$ if and only if (X, d) can be embedded isometrically into ℓ_1, namely without changing any of the distances. Consider a vertex cover S and its corresponding solution to SDP (2), i.e., $\mathbf{v}_i = 1$ for every $i \in S \cup \{0\}$ and $\mathbf{v}_i = -1$ for every $i \notin S$. The metric defined by $\|\cdot\|^2$ on this solution (i.e., $d(i, j) = \|\mathbf{v}_i - \mathbf{v}_j\|^2$) is isometrically embeddable into ℓ_1. Thus we can strengthen SDP (2) by allowing any arbitrary list of valid inequalities in ℓ_1 to be added. The triangle inequality is one type of such constraints. The next natural inequality of this sort is the pentagonal inequality: A metric space (X, d) is said to satisfy the pentagonal inequality if for $S, T \subset X$ of sizes 2 and 3 respectively it holds that $\sum_{i\in S, j\in T} d(i, j) \geq \sum_{i,j\in S} d(i, j) + \sum_{i,j\in T} d(i, j)$. Note that this inequality does not apply to every metric, but it does hold for those that are ℓ_1-embeddable. This leads to the following natural strengthening of SDP (3):

$$\begin{array}{lll} \textbf{Min} & \sum_{i\in V} 1 - \|\mathbf{v}_0 - \mathbf{v}_i\|^2/4 & \\ \text{s.t.} & \|\mathbf{v}_i - \mathbf{v}_0\|^2 + \|\mathbf{v}_j - \mathbf{v}_0\|^2 = \|\mathbf{v}_i - \mathbf{v}_j\|^2 & \forall\, ij \in E \\ & \sum_{i\in S, j\in T} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \geq \sum_{i,j\in S} \|\mathbf{v}_i - \mathbf{v}_j\|^2 + & \forall\, S, T \subseteq \{0\} \cup V, \\ & \qquad\qquad \sum_{i,j\in T} \|\mathbf{v}_i - \mathbf{v}_j\|^2 & |S| = 2, |T| = 3 \\ & \|\mathbf{v}_i\| = 1 & \forall\, i \in \{0\} \cup V \end{array} \qquad (4)$$

In Theorem 5, we prove that SDP (4) has an integrality gap of $2 - o(1)$. It is important to point out that a-priori there is no reason to believe that local addition of inequalities such as these will not improve the integrality gap; indeed in the case of Sparsest Cut triangle inequality is necessary to achieve the $O(\sqrt{\log n})$ bound mentioned above. It is interesting to note that for Sparsest Cut, it is not known how to show a nonconstant integrality gap against pentagonal (or any other k-gonal) inequalities, although recently a nonconstant integrality gap was shown in [20] and later in [8], in the presence of the triangle inequalities[2].

A recent related result by Georgiou, Magen, Pitassi and Tourlakis [13] shows an integrality gap of $2 - o(1)$ for a nonconstant number of rounds of the so-called LS_+ system for Vertex Cover. It is not known whether this result subsumes Theorem 5 or not, since pentagonal inequalities are not generally implied by any number of rounds of the $LS+$ procedure. We elaborate on this in the Discussion section.

One can further impose any ℓ_1-constraint not only for the metric defined by $\{\mathbf{v}_i : i \in V \cup \{0\}\}$, but also for the one that comes from $\{\mathbf{v}_i : i \in V \cup \{0\}\} \cup \{-\mathbf{v}_i : i \in V \cup \{0\}\}$. Triangle inequalities for this extended set result in the constraints $\|\mathbf{v}_i - \mathbf{v}_j\|^2 + \|\mathbf{v}_i - \mathbf{v}_k\|^2 + \|\mathbf{v}_j - \mathbf{v}_k\|^2 \leq 2$. The corresponding tighter SDP is used in [17] to get integraility gap of at most $2 - \Omega(\frac{1}{\sqrt{\log n}})$. Karakostas [17] asks whether the integrality gap of this strengthening breaks the "$2 - o(1)$ barrier": we answer this negatively in Section 4.3. In fact we show that the above upper bound is almost asymptotically tight, exhibiting integrality gap of $2 - O(\sqrt{\frac{\log\log n}{\log n}})$.

Integrality gap with respect to ℓ_1 embeddability: At the extreme, strengthening the SDP with ℓ_1-valid constraints, would imply the condition that the metric defined by $\|\cdot\|$ on $\{\mathbf{v}_i : i \in \{0\} \cup V\}$, namely $d(i, j) = \|\mathbf{v}_i - \mathbf{v}_j\|^2$ is ℓ_1 embeddable. Doing so leads to the following intractable program:

[2] As Khot and Vishnoi note, and leave as an open problem, it is possible that their example satisfies some or all k-gonal inequalities.

$$\begin{array}{lll} \textbf{Min} & \sum_{i \in V} 1 - \|\mathbf{v}_0 - \mathbf{v}_i\|^2/4 & \\ \text{s.t.} & \|\mathbf{v}_i - \mathbf{v}_0\|^2 + \|\mathbf{v}_j - \mathbf{v}_0\|^2 = \|\mathbf{v}_i - \mathbf{v}_j\|^2 & \forall\ ij \in E \\ & \|\mathbf{v}_i\| = 1 & \forall\ i \in \{0\} \cup V \\ & c_1(\{\mathbf{v}_i : i \in \{0\} \cup V\}, \|\cdot\|^2) = 1 & \end{array} \tag{5}$$

In [1], it is shown that an SDP formulation of Minimum Multicut, even with the constraint that the $\|\cdot\|^2$ distance over the variables is isometrically embeddable into ℓ_1, still has a large integrality gap. For the Max Cut problem, which is more intimately related to our problem, it is easy to see that ℓ_1 embeddability does not prevent integrality gap of 8/9. It is therefore tempting to believe that there is a large integrality gap for SDP (5) as well. Surprisingly, SDP (5) has no gap at all: we show in Theorem 2, that the value of SDP (5) is exactly the size of the minimum vertex cover. A consequence of this fact is that any feasible solution to SDP (2) that surpasses the minimum vertex cover induces an ℓ_2^2 distance which is not isometrically embeddable into ℓ_1. This includes the integrality gap constructions of Kleinberg and Goemans, and that of Charikar's for SDPs (2) and (3) respectively. The construction of Charikar is more interesting in this context as the obtained ℓ_2^2 distance is also a negative type metric, that is an ℓ_2^2 metric that satisfies triangle inequality. See [9] for background and nomenclature.

In contrast to Theorem 2, we show in Theorem 3 that if we relax the embeddability constraint in SDP (5) to $c_1(\{\mathbf{v}_i : i \in \{0\} \cup V\}, \|\cdot\|^2) \leq 1 + \delta$ for any constant $\delta > 0$, then the integrality gap may "jump" to $2 - o(1)$. Compare this with a problem such as Sparsest Cut in which an addition of such a constraint immediately implies integrality gap at most $1 + \delta$.

Negative type metrics that are not ℓ_1 embeddable: Negative type metrics are metrics which are the squares of Euclidean distances of set of points in Euclidean space. Inspired by Theorem 2, we construct in Section 5 a simple negative type metric space $(X, \|\cdot\|^2)$ that does not embed well into ℓ_1. Specifically, we get $c_1(X) \geq \frac{8}{7} - \epsilon$ for every $\epsilon > 0$. In order to show this we prove a new isoperimetric inequality for the hypercube $Q_n = \{-1, 1\}^n$, which we believe is of independent interest. This theorem generalizes the standard one, and under certain conditions provides better guarantees for edge expansion.

Theorem 1 *(Generalized Isoperimetric inequality) For every set $S \subseteq Q_n$,*

$$|E(S, S^c)| \geq |S|(n - \log_2 |S|) + p(S).$$

where $p(S)$ denotes the number of vertices $\mathbf{u} \in S$ such that $-\mathbf{u} \in S$.

Khot and Vishnoi [20] constructed an example of an n-point negative type metric that for every $\delta > 0$ requires distortion at least $(\log \log n)^{1/6-\delta}$ to embed into ℓ_1. Krauthgamer and Rabani [22] showed that in fact Khot and Vishnoi's example requires a distortion of at least $\Omega(\log \log n)$. Later Devanur, Khot, Saket and Vishnoi [8] showed an example with distortion $\Omega(\log \log n)$ even on average when embedded into ℓ_1 (we note that our example is also "bad" on average). Although the above examples require nonconstant distortion to embed into ℓ_1, we believe that our result is still interesting because (i) our construction is much simpler than the ones in [20, 22, 8]; in comparison, showing that triangle inequality holds requires a lot of technical work in [20, 22, 8] whereas in our construction it is immediate (ii) very few examples are known of negative type metrics that do not embed isometrically into ℓ_1, and any such example reveals some underlying sructure. Prior to Khot and Vishnoi's result, the best known lower bounds (see [20]) were due to Vempala, 10/9 for a metric obtained by a computer search, and Goemans, 1.024 for a metric based on the Leech Lattice. We mention that by [4] every negative type metric embeds into ℓ_1 with distortion $O(\sqrt{\log n} \log \log n)$.

2 Preliminaries and notation

A vertex cover of a graph G is a set of vertices that touch all edges. An independent set in G is a set $I \subseteq V$ such that no edge $e \in E$ joins two vertices in I. We denote by $\alpha(G)$ the size of the maximum independent set of G. Vectors are always denoted in bold font (such as $\mathbf{v}$, $\mathbf{w}$, etc.); $\|\mathbf{v}\|$ stands for the Euclidean norm of $\mathbf{v}$, $\mathbf{u} \cdot \mathbf{v}$ for the inner product of $\mathbf{u}$ and $\mathbf{v}$, and $\mathbf{u} \otimes \mathbf{v}$ for their tensor product. Specifically, if $\mathbf{v}, \mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \otimes \mathbf{v}$ is the vector with coordinates indexed by ordered pairs $(i,j) \in [n]^2$ that assumes value $\mathbf{u}_i \mathbf{v}_j$ on coordinate (i,j). Similarly, the tensor product of more than two vectors is defined. It is easy to see that $(\mathbf{u} \otimes \mathbf{v}).(\mathbf{u}' \otimes \mathbf{v}') = (\mathbf{u} \cdot \mathbf{u}')(\mathbf{v} \cdot \mathbf{v}')$. For two vectors $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$, denote by $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{n+m}$ the vector whose projection to the first n coordinates is $\mathbf{u}$ and to the last m coordinates is $\mathbf{v}$.

Next, we give a few basic definitions and facts about finite metric spaces. A metric space (X, d_X) embeds with distortion at most D into (Y, d_Y) if there exists a mapping $\phi : X \mapsto Y$ so that for all $a, b \in X$ $\gamma \cdot d_X(a,b) \leq d_Y(\phi(a), \phi(b)) \leq \gamma D \cdot d_X(a,b)$, for some $\gamma > 0$. We say that (X, d) is ℓ_1 embeddable if it can be embedded with distortion 1 into $\mathbb{R}^m$ equipped with the ℓ_1 norm. An ℓ_2^2 distance on X is a distance function for which there there are vectors $\mathbf{v}_x \in \mathbb{R}^m$ for every $x \in X$ so that $d(x,y) = \|\mathbf{v}_x - \mathbf{v}_y\|^2$. If, in addition, d satisfies triangle inequality, we say that d is an ℓ_2^2 metric or negative type metric. It is well known [9] that every ℓ_1 embeddable metric is also a negative type metric.

3 ℓ_1 and integrality gap of SDPs for Vertex Cover – an "all or nothing" phenomenon

It is well known that for Sparsest Cut there is a tight connection between ℓ_1 embeddability and integrality gap. In fact the integrality gap is bounded above by the least ℓ_1 distortion of the SDP solution. At the other extreme stand problems like Max Cut and Multi Cut, where ℓ_1 embeddability does not provide any strong evidence for small integrality gap. In this section we show that Vertex Cover falls somewhere between these two classes of ℓ_1-integrality gap relationship witnessing a sharp transition in integrality gap in the following sense: while ℓ_1 embeddability implies no integrality gap, allowing a small distortion, say 1.001 does not prevent integrality gap of $2 - o(1)$!

Theorem 2 *For a graph $G = (V, E)$, the answer to the SDP formulated in SDP (5) is the size of the minimum vertex cover of G.*

Proof. Let d be the metric solution of SDP (5). We know that d is the result of an ℓ_2^2 unit representation (i.e., it comes from square norms between unit vectors), and furthermore it is ℓ_1 embeddable. By cut representations of ℓ_1 embeddable metrics (see e.g. [9]) we can assume that there exist $\lambda_t > 0$ and $f_t : \{0\} \cup V \to \{-1, 1\}$, $t = 1, \ldots, m$, such that

$$\|\mathbf{v}_i - \mathbf{v}_j\|^2 = \sum_{t=1}^{m} \lambda_t |f_t(i) - f_t(j)|, \tag{6}$$

for every $i, j \in \{0\} \cup V$. Without loss of generality, we can assume that $f_t(0) = 1$ for every t. For convenience, we switch to talk about Independent Set and its relaxation, which is the same as SDP (5) except the objective becomes Max $\sum_{\mathrm{i}\in \mathrm{V}} \|\mathbf{v}_0 - \mathbf{v}_{\mathrm{i}}\|^2/4$. Obviously, the theorem follows from showing that this is an exact relaxation.

We argue that (i) $I_t = \{i \in V : f_t(i) = -1\}$ is a (nonempty) independent set for every t, and (ii) $\sum \lambda_t = 2$. Assuming these two statements we get

$$\sum_{i\in V} \frac{\|\mathbf{v}_i - \mathbf{v}_0\|^2}{4} \quad = \quad \sum_{i\in V} \frac{\sum_{t=1}^m \lambda_t |1 - f_t(i)|}{4} = \sum_{t=1}^m \frac{\lambda_t |I_t|}{2} \le \max_{t\in[m]} |I_t| \le \alpha(G),$$

and so the relaxation is exact and we are done.

We now prove the two statements. The first is rather straightforward: For $i, j \in I_t$, (6) implies that $d(i, 0) + d(0, j) > d(i, j)$. It follows that ij cannot be an edge else it would violate the first condition of the SDP (we may assume that I_t is nonempty since otherwise the $f_t(\cdot)$ terms have no contribution in (6)). The second statement is more surprising and uses the fact that the solution is optimal. The falsity of such a statement for the problem of Max Cut (say) explains the different behaviour of the latter problem with respect to integrality gaps of ℓ_1 embeddable solutions. We now describe the proof.

Let $\mathbf{v}'_i = (\sqrt{\lambda_1/2} f_1(i), \ldots, \sqrt{\lambda_m/2} f_m(i), 0)$. From (6) we conclude that $\|\mathbf{v}'_i - \mathbf{v}'_j\|^2 = \|\mathbf{v}_i - \mathbf{v}_j\|^2$, hence there exists a vector $\mathbf{w} = (w_1, w_2, ..., w_{m+1}) \in \mathbb{R}^{m+1}$ and an orthogonal transformation T, such that

$$\mathbf{v}_i = T\left(\mathbf{v}'_i + \mathbf{w}\right)$$

Since the constraints and the objective function of the SDP are invariant under orthogonal transformations, without loss of generality we may assume that

$$\mathbf{v}_i = \mathbf{v}'_i + \mathbf{w},$$

for $i \in V \cup \{0\}$. We know that

$$1 = \|\mathbf{v}_i\|^2 = \|T(\mathbf{v}'_i + \mathbf{w})\|^2 = \|\mathbf{v}'_i + \mathbf{w}\|^2 = w_{m+1}^2 + \sum_{t=1}^m (\sqrt{\lambda_t/2} f_t(i) + w_t)^2. \tag{7}$$

Since $\|\mathbf{v}'_i\|^2 = \|\mathbf{v}'_0\|^2 = \sum_{t+1}^m \lambda_t/2$, for every $i \in V \cup \{0\}$, from (7) we get $\mathbf{v}'_0 \cdot \mathbf{w} = \mathbf{v}'_i \cdot \mathbf{w}$. Summing this over all $i \in V$, we have

$$|V|(\mathbf{v}'_0 \cdot \mathbf{w}) = \sum_{i\in V} \mathbf{v}'_i \cdot \mathbf{w} = \sum_{t=1}^m (|V| - 2|I_t|)\sqrt{\lambda_t/2}\, w_t,$$

or

$$\sum_{t=1}^m |V|\sqrt{\lambda_t/2}\, w_t = \sum_{t=1}^m (|V| - 2|I_t|)\sqrt{\lambda_t/2}\, w_t,$$

and therefore

$$\sum_{t=1}^m |I_t|\sqrt{\lambda_t/2}\, w_t = 0. \tag{8}$$

Now (7) and (8) imply that

$$\max_{t\in[m]} |I_t| \ge \sum_{t=1}^m (\sqrt{\lambda_t/2} f_t(0) + w_t)^2 |I_t| = \sum_{t=1}^m \left(\frac{\lambda_t |I_t|}{2} + w_t^2 |I_t|\right) \ge \sum_{t=1}^m \frac{\lambda_t |I_t|}{2}. \tag{9}$$

As we have observed before

$$\sum_{t=1}^{m} \frac{\lambda_t |I_t|}{2} = \sum_{i \in V} \frac{\|\mathbf{v}_i - \mathbf{v}_0\|^2}{4}$$

which means (as clearly $\sum_{i\in V} \frac{\|\mathbf{v}_i - \mathbf{v}_0\|^2}{4} \geq \alpha(G)$) that the inequalities in (9) must be tight. Now, since $|I_t| > 0$ we get that $\mathbf{w} = \mathbf{0}$ and from (7) we get the second statement, i.e., $\sum \lambda_t = 2$. This concludes the proof. ■

Now let δ be an arbitrary positive number, and let us relax the last constraint in SDP (5) to get

$$\begin{array}{lll} \textbf{Min} & \sum_{i\in V} 1 - \|\mathbf{v}_0 - \mathbf{v}_i\|^2/4 & \\ \text{s.t.} & \|\mathbf{v}_i - \mathbf{v}_0\|^2 + \|\mathbf{v}_j - \mathbf{v}_0\|^2 = \|\mathbf{v}_i - \mathbf{v}_j\|^2 & \forall\ ij \in E \\ & \|\mathbf{v}_i\| = 1 & \forall\ i \in \{0\} \cup V \\ & c_1(\{\mathbf{v}_i : i \in \{0\} \cup V\}, \|\cdot\|^2) \leq 1 + \delta & \end{array}$$

Theorem 3 *For every $\epsilon > 0$, there is a graph G for which $\frac{\mathrm{vc}(G)}{\mathrm{sd}(G)} \geq 2 - \epsilon$, where $\mathrm{sd}(G)$ is the solution to the above SDP.*

The proof appears in the next section after we describe Charikar's construction.

4 Integrality gap for stronger semidefinite formulations

In this section we discuss the integrality gap for stronger semidefinite formulations of vertex cover. In particular we show that Charikar's construction satisfies both SDPs (11) and (4). We start by describing this construction.

4.1 Charikar's construction

The graphs used in the construction are the so-called Hamming graphs. These are graphs with vertices $\{-1, 1\}^n$ and two vertices are adjacent if their Hamming distance is exactly an even integer $d = \gamma n$. A result of Frankl and Rödl [12] shows that $\mathrm{vc}(G) \geq 2^n - (2 - \delta)^n$, where $\delta > 0$ is a constant depending only on γ. In fact, when one considers the exact dependency of δ in γ it can be shown (see [13]) that as long as $\gamma = \Omega(\sqrt{\log n/n})$ then any vertex cover comprises $1 - O(1/n)$ fraction of the graph. Kleinberg and Goemans [21] showed that by choosing a constant γ and n sufficiently large, this graph gives an integrality gap of $2 - \epsilon$ for SDP (1). Charikar [6] showed that in fact G implies the same result for the SDP formulation in (2) too. To this end he introduced the following solution to SDP (2):

For every $\mathbf{u}_i \in \{-1, 1\}^n$, define $\mathbf{u}'_i = \mathbf{u}_i/\sqrt{n}$, so that $\mathbf{u}'_i \cdot \mathbf{u}'_i = 1$. Let $\lambda = 1 - 2\gamma$, $q(x) = x^{2t} + 2t\lambda^{2t-1}x$ and define $\mathbf{y}_0 = (0, \ldots, 0, 1)$, and

$$\mathbf{y}_i = \sqrt{\frac{1-\beta^2}{q(1)}} \left(\underbrace{\mathbf{u}'_i \otimes \ldots \otimes \mathbf{u}'_i}_{2t \text{ times}}, \sqrt{2t\lambda^{2t-1}}\mathbf{u}'_i, 0 \right) + \beta \mathbf{y}_0,$$

where β will be determined later. Note that $\mathbf{y}_i$ is normalized to satisfy $\|\mathbf{y}_i\| = 1$.

Moreover $\mathbf{y}_i$ is defined so that $\mathbf{y}_i \cdot \mathbf{y}_j$ takes its minimum value when $ij \in E$, i.e., when $\mathbf{u}'_i \cdot \mathbf{u}'_j = -\lambda$. As is shown in [6], for every $\epsilon > 0$ we may set $t = \Omega(\frac{1}{\epsilon}), \beta = \Theta(1/t), \gamma = \frac{1}{4t}$ to get that $(\mathbf{y}_0 - \mathbf{y}_i) \cdot (\mathbf{y}_0 - \mathbf{y}_j) = 0$ for $ij \in E$, while $(\mathbf{y}_0 - \mathbf{y}_i) \cdot (\mathbf{y}_0 - \mathbf{y}_j) \geq 0$ always.

Now we verify that all the triangle inequalities, i.e., the second constraint of SDP (2) are satisfied: First note that since every coordinate takes only two different values for the vectors in $\{\mathbf{y}_i : i \in V\}$, it is easy to see that $c_1(\{\mathbf{y}_i : i \in V\}, \|\cdot\|^2) = 1$. So the triangle inequality holds when $i, j, k \in V$. When $i = 0$ or $j = 0$, the inequality is trivial, and it only remains to verify the case that $k = 0$, i.e., $(\mathbf{y}_0 - \mathbf{y}_i) \cdot (\mathbf{y}_0 - \mathbf{y}_j) \geq 0$, which was already mentioned above. Now $\sum_{i \in V} (1 + \mathbf{y}_0 \cdot \mathbf{y}_i)/2 = \frac{1+\beta}{2} \cdot |V| = \left(\frac{1}{2} + O(\epsilon)\right)|V|$. In our application, we prefer to set γ and ϵ to be $\Omega(\sqrt{\frac{\log\log n}{\log n}})$ and since, by the above comment, $\mathrm{vc}(G) = (1 - O(1/n))|V|$ the integrality gap we get is

$$(1 - O(1/n))/(1/2 + O(\epsilon)) = 2 - O(\epsilon) = 2 - O\left(\sqrt{\frac{\log\log|V|}{\log|V|}}\right).$$

4.2 Proof of Theorem 3

We show that the negative type metric implied by Charikar's solution (after adjusting the parameters appropriately) requires distortion of at most $1 + \delta$. Let $\mathbf{y}_i$ and $\mathbf{u}'_i$ be defined as in Section 4.1. To prove Theorem 3, it is sufficient to prove that $c_1(\{\mathbf{y}_i : i \in \{0\} \cup V\}, \|\cdot\|^2) = 1 + o(1)$. Note that every coordinate of $\mathbf{y}_i$ for all $i \in V$ takes at most two different values. It is easy to see that this implies $c_1(\{\mathbf{y}_i : i \in V\}, \|\cdot\|^2) = 1$. In fact

$$f : \mathbf{y}_i \mapsto \frac{1-\beta^2}{q(1)} \left(\frac{2}{n^t} \underbrace{\mathbf{u}'_i \otimes \ldots \otimes \mathbf{u}'_i}_{2t \text{ times}}, \frac{2}{\sqrt{n}} 2t\lambda^{2t-1} \mathbf{u}'_i \right), \tag{10}$$

is an isometry from $(\{\mathbf{y}_i : i \in V\}, \|\cdot\|^2)$ to ℓ_1. For $i \in V$, we have

$$\|f(\mathbf{y}_i)\|_1 = \frac{1-\beta^2}{q(1)} \left(\frac{2}{n^t} \times \frac{n^{2t}}{n^t} + \frac{2}{\sqrt{n}} 2t\lambda^{2t-1} \frac{1}{\sqrt{n}} + 0 \right) = \frac{1-\beta^2}{q(1)} \times (2 + 4t\lambda^{2t-1})$$

Since $\beta = \Theta(\frac{1}{t})$, recalling that $\lambda = 1 - \frac{1}{2t}$, it is easy to see that for every $i \in V$, $\lim_{t\to\infty} \|f(\mathbf{y}_i)\|_1 = 2$. On the other hand for every $i \in V$

$$\lim_{t\to\infty} \|\mathbf{y}_i - \mathbf{y}_0\|^2 = \lim_{t\to\infty} 2 - 2(\mathbf{y}_i \cdot \mathbf{y}_0) = \lim_{t\to\infty} 2 - 2\beta = 2.$$

So if we extend f to $\{\mathbf{y}_i : i \in V \cup \{0\}\}$ by defining $f(\mathbf{y}_0) = \mathbf{0}$, we obtain a mapping from $(\{\mathbf{y}_i : i \in V \cup \{0\}\}, \|\cdot\|^2)$ to ℓ_1 whose distortion tends to 1 as t goes to infinity.

4.3 Karakostas' and pentagonal SDP formulations

Karakostas suggests the following SDP relaxation, that is the result of adding to SDP (3) the triangle inequalities applied to the set $\{\mathbf{v}_i : i \in V \cup \{0\}\} \cup \{-\mathbf{v}_i : i \in V \cup \{0\}\}$.

$$\begin{array}{lll} \textbf{Min} & \sum_{i\in V}(1+\mathbf{v}_0\mathbf{v}_i)/2 & \\ \text{s.t.} & (\mathbf{v}_i-\mathbf{v}_0)\cdot(\mathbf{v}_j-\mathbf{v}_0)=0 & \forall\, ij\in E \\ & (\mathbf{v}_i-\mathbf{v}_k)\cdot(\mathbf{v}_j-\mathbf{v}_k)\geq 0 & \forall\, i,j,k\in V \\ & (\mathbf{v}_i+\mathbf{v}_k)\cdot(\mathbf{v}_j-\mathbf{v}_k)\geq 0 & \forall\, i,j,k\in V \\ & (\mathbf{v}_i+\mathbf{v}_k)\cdot(\mathbf{v}_j+\mathbf{v}_k)\geq 0 & \forall\, i,j,k\in V \\ & \|\mathbf{v}_i\|=1 & \forall\, i\in\{0\}\cup V. \end{array} \tag{11}$$

Theorem 4 *The integrality gap of SDP (11) is* $2 - O(\sqrt{\log\log|V|/\log|V|})$.

Proof. We show that Charikar's construction satisfies formulation (11). By [6] and from the discussion in Section 4.1, it follows that all edge constraints and triangle inequalities of the original points hold. Hence we need only consider triangle inequalities with at least one nonoriginal point. By homogeneity, we may assume that there is exactly one such point.

Since all coordinates of $\mathbf{y}_i$ for $i > 0$ assume only two values with the same absolute value, it is clear that not only does the metric they induce is ℓ_1 but also taking $\pm\mathbf{y}_i$ for $i > 0$ gives an ℓ_1 metric; in particular all triangle inequalities that involve these vectors are satisfied. In fact, we may fix our attention to triangles in which $\pm\mathbf{y}_0$ is the middle point. This is since

$$(\pm\mathbf{y}_i - \pm\mathbf{y}_j)\cdot(\mathbf{y}_0 - \pm\mathbf{y}_j) = (\pm\mathbf{y}_j - \mathbf{y}_0)\cdot(\mp\mathbf{y}_i - \mathbf{y}_0).$$

Consequently, and using symmetry, we are left with checking the nonnegativity of $(\mathbf{y}_i + \mathbf{y}_0)\cdot(\mathbf{y}_j + \mathbf{y}_0)$ and $(-\mathbf{y}_i - \mathbf{y}_0)\cdot(\mathbf{y}_j - \mathbf{y}_0)$.

$$(\mathbf{y}_i+\mathbf{y}_0)\cdot(\mathbf{y}_j+\mathbf{y}_0) = 1+\mathbf{y}_0\cdot(\mathbf{y}_i+\mathbf{y}_j)+\mathbf{y}_i\cdot\mathbf{y}_j \geq 1+2\beta+\beta^2-(1-\beta^2) = 2\beta(1+\beta)\geq 0.$$

Finally, $(-\mathbf{y}_i-\mathbf{y}_0)\cdot(\mathbf{y}_j-\mathbf{y}_0) = 1+\mathbf{y}_0\cdot(\mathbf{y}_i-\mathbf{y}_j)-\mathbf{y}_i\cdot\mathbf{y}_j = 1-\mathbf{y}_i\cdot\mathbf{y}_j\geq 0$ as $\mathbf{y}_i,\mathbf{y}_j$ are of norm 1. ■

By now we know that taking all the ℓ_1 constraints leads to an exact relaxation, but not a tractable one. Our goal here is to explore the possibility that stepping towards ℓ_1 embeddability while still maintaining computational feasibility would considerably reduce the integrality gap. A canonical subset of valid inequalities for ℓ_1 metrics is the so-called *Hypermetric inequalities*. Metrics that satisfy all these inequalities are called *hypermetrics*. Again, taking all these constraints is not feasible, and yet we do not know whether this may lead to a better integrality gap (notice that we do not know that Theorem 2 remains true if we replace the ℓ_1 embeddability constraints with a hypermetricity constraint). See [9] for a related discussion about hypermetrics. We instead consider the effect of adding a small number of such constraints. The simplest hypermetric inequalities beside triangle inequalities are the *pentagonal* inequalities. These constraints consider two sets of points of size 2 and 3, and require that the sum of the distances between points in different sets is at least the sum of the distances within sets. Formally, let $S, T \subset X$, $|S| = 2, |T| = 3$, then we have the inequality $\sum_{i\in S, j\in T} d(i,j) \geq \sum_{i,j\in S} d(i,j) + \sum_{i,j\in T} d(i,j)$. To appreciate this inequality it is useful to describe where it fails. Consider the graph metric of $K_{2,3}$. Here, the LHS of the inequality is 6 and the RHS is 8, hence $K_{2,3}$ violates the pentagonal inequality. In the following theorem we show that this strengthening past the triangle inequalities fails to reduce the integrality gap significantly.

Theorem 5 *The integrality gap of SDP (4) is* $2 - O(\sqrt{\log\log|V|/\log|V|})$.

Proof. We note that in order to satisfy the triangle inequalities, the conditions that should be satisfied by the "tensoring-polynomial" used in the construction ("q" in the notation of the previous subsection) are rather modest. Essentially we needed that $q'(-\lambda) = 0$, $q(-\lambda)/q(1)$ approaches -1, and that $q''(-\lambda) \geq 0$. For the pentagonal inequalities we need to require more properties from q, namely that it is convex on its entire domain and that its derivative satisfies certain linear conditions, all of which turn out to be true.

We show that the metric space used in Charikar's construction is a feasible solution. By ignoring $\mathbf{y}_0$ the space defined by $d(i,j) = \|\mathbf{y}_i - \mathbf{y}_j\|^2$ is ℓ_1 embeddable. Therefore, the only ℓ_1-valid inequalities that may be violated are ones containing $\mathbf{y}_0$. Hence, we wish to consider a pentagonal inequality containing $\mathbf{y}_0$ and four other vectors, denoted by $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4$. Assume first that the partition of the five points in the inequality puts $\mathbf{y}_0$ together with two other points; then, using the fact that $d(0,1) = d(0,2) = d(0,3) = d(0,4)$ and triangle inequality we get that such an inequality must hold. It remains to consider a partition of the form $(\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}, \{\mathbf{y}_4, \mathbf{y}_0\})$, and show that:

$$d(1,2) + d(1,3) + d(2,3) + d(0,4) \leq d(1,4) + d(2,4) + d(3,4) + d(0,1) + d(0,2) + d(0,3)$$

As the vectors are of unit norm, it is clear that $d(0,i) = 2 - 2\beta$ for all $i > 0$ and that $d(i,j) = 2 - 2\mathbf{y}_i\mathbf{y}_j$. Recall that every $\mathbf{y}_i$ is associated with a $\{-1,1\}$ vector $\mathbf{u}_i$ and with its normalized multiple $\mathbf{u}'_i$. Also, it is simple to check that $\mathbf{y}_i \cdot \mathbf{y}_j = \beta^2 + (1-\beta^2)q(\mathbf{u}'_i \cdot \mathbf{u}'_j)/q(1)$ where $q(x) = x^{2t} + 2\lambda^{2t-1}x$. After substituting the distances as functions of the normalized vectors, our goal will then be to show:

$$E = q(\mathbf{u}'_1 \cdot \mathbf{u}'_2) + q(\mathbf{u}'_1 \cdot \mathbf{u}'_3) + q(\mathbf{u}'_2 \cdot \mathbf{u}'_3) - q(\mathbf{u}'_1 \cdot \mathbf{u}'_4) - q(\mathbf{u}'_2 \cdot \mathbf{u}'_4) - q(\mathbf{u}'_3 \cdot \mathbf{u}'_4) \geq -\frac{2q(1)}{1+\beta} \qquad (12)$$

The rest of the proof analyzes the minima of the function E and ensures that (12) is satisfied at those minima. We first partition the coordinates of the original hypercube into four sets according to the values assumed by $\mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{u}_3$. We may assume that in any coordinate at most one of these get the value 1 (otherwise multiply the values of the coordinate by -1). We get four sets, P_0 for the coordinates in which all three vectors assume value -1, and P_1, P_2, P_3 for the coordinates in which exactly $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ respectively assumes value 1.

We now consider $\mathbf{u}_4$. We argue that without loss of generality, we may assume that $\mathbf{u}_4$ is "pure" on each of the P_0, P_1, P_2, P_3 at a minimum of E; in other words it is either all 1 or all -1 on each one of P_0, P_1, P_2, P_3.

Proposition 1 *If there is a violating configuration, then there is one in which* $\mathbf{u}_4$ *is either all* 1 *or all* -1 *on each one of* P_0, P_1, P_2, P_3.

Proof. Assume for the sake of contradiction that there are w coordinates in P_0 on which $\mathbf{u}_4$ assumes value -1, and that $0 < w < |P_0|$. Let $\mathbf{u}_4^+$ (similarly $\mathbf{u}_4^-$) be identical to $\mathbf{u}_4$ except we replace one 1 in P_0 by -1 (replace one -1 in P_0 by 1). We show that replacing $\mathbf{u}_4$ by $\mathbf{u}_4^+$ or by $\mathbf{u}_4^-$ we decrease the expression E. Let $p_i = \mathbf{u}_i \cdot \mathbf{u}_4$, $p_i^+ = \mathbf{u}'_i \cdot (\mathbf{u}_4^+)'$ and $p_i^- = \mathbf{u}'_i \cdot (\mathbf{u}_4^-)'$ for $i = 1,2,3$. Notice that the above replacement only changes the negative terms in (12) so our goal now is to show that $\sum_{i=1}^3 q(p_i) < \max\{\sum_{i=1}^3 q(p_i^+), \sum_{i=1}^3 q(p_i^-)\}$. But:

$$\max\{\sum_{i=1}^3 q(p_i^+), \sum_{i=1}^3 q(p_i^-)\} \geq \sum_{i=1}^3 \frac{q(p_i^+) + q(p_i^-)}{2} > \sum_{i=1}^3 q\left(\frac{p_i^+ + p_i^-}{2}\right) = \sum_{i=1}^3 q(p_i)$$

where the last inequality is using the (strict) convexity of q. This of course applies to P_1, P_2 and P_3 in precisely the same manner. ■

For P_0, we can in fact say something stronger than we do for P_1, P_2, P_3:

Proposition 2 *If there is a violating configuration, then there is one in which* $\mathbf{u}_4$ *has all the* P_0 *coordinates set to* -1.

The above characterizations significantly limit the type of configurations we need to check. Proposition 1 was based solely on the (strict) convexity of q. Proposition 2 is more involved and uses more properties of the polynomial q. If q was a monotone increasing function it would be obvious, but of course the whole point behind q is that it brings to minimum some intermediate value $(-\lambda)$ and hence can not be increasing. We postpone the proof of Proposition 2 till the end of the Section and we will now continue our analysis assuming the proposition.

The cases that are left are characterized by whether $\mathbf{u}_4$ is 1 or -1 on each of P_1, P_2, P_3. By symmetry all we really need to know is $\xi(\mathbf{u}_4) = |\{i : \mathbf{u}_4 \text{ is 1 on } P_i\}|$. If $\xi(\mathbf{u}_4) = 1$ it means that $\mathbf{u}_4$ is the same as one of $\mathbf{u}_1, \mathbf{u}_2$ or $\mathbf{u}_4$ hence the pentagonal inequality reduces to the triangle inequality, which we already know is valid. If $\xi(\mathbf{u}_4) = 3$, it is easy to see that $\mathbf{u}_1'\mathbf{u}_4' = \mathbf{u}_2'\mathbf{u}_3'$, and likewise $\mathbf{u}_2'\mathbf{u}_4' = \mathbf{u}_1'\mathbf{u}_3'$ and $\mathbf{u}_3'\mathbf{u}_4' = \mathbf{u}_1'\mathbf{u}_2'$ hence E is 0 for these cases, which means that (12) is satisfied.

We are left with the cases $\xi(\mathbf{u}_4) \in \{0, 2\}$.

Case 1: $\xi(\mathbf{u}_4) = 0$

Let $x = \frac{2}{n}|P_1|, y = \frac{2}{n}|P_2|, z = \frac{2}{n}|P_3|$. Notice that $x + y + z = \frac{2}{n}(|P_1| + |P_2| + |P_3|) \leq 2$, as these sets are disjoint. Now, think of

$$E = q(1 - (x + y)) + q(1 - (x + z)) + q(1 - (y + z)) - q(1 - x) - q(1 - y) - q(1 - z)$$

as a function from $\mathbb{R}^3$ to $\mathbb{R}$. We will show that E achieves its minimum at points where either x, y or z are zero. Assume that $0 \leq x \leq y \leq z$.

Consider the function $g(\delta) = E(x - \delta, y + \delta, z)$. It is easy to see that $g'(0) = q'(1 - (x + z)) - q'(1 - (y + z)) - q'(1 - x) + q'(1 - y)$. We will prove that $g'(\delta) \leq 0$ for every $\delta \in [0, x]$. This, by the Mean Value Theorem implies that $E(0, x + y, z) \leq E(x, y, z)$, and hence we may assume that $x = 0$. This means that $\mathbf{y}_1 = \mathbf{y}_4$ which reduces to the triangle inequality on $\mathbf{y}_0, \mathbf{y}_2, \mathbf{y}_3$.

Note that in $g'(0)$, the two arguments in the terms with positive sign have the same average as the arguments in the terms with negative sign, namely $\mu = 1 - (x + y + z)/2$. We now have $g'(0) = q'(\mu + b) - q'(\mu + s) - q'(\mu - s) + q'(\mu - b)$, where $b = (x - y + z)/2, s = (-x + y + z)/2$. After calculations:

$$\begin{aligned} g'(0) &= 2t[(\mu + b)^{2t-1} + (\mu - b)^{2t-1} - (\mu + s)^{2t-1} - (\mu - s)^{2t-1}] \\ &= 4t \sum_{i \text{ even}} \binom{2t-1}{i} \mu^{2t-1-i}(b^i - s^i) \end{aligned}$$

Observe that $\mu \geq 0$. Since $x \leq y$, we get that $s \geq b \geq 0$. This means that $g'(0) \leq 0$. It can be easily checked that the same argument holds if we replace x, y by $x - \delta$ and $y + \delta$. Hence $g'(\delta) \leq 0$ for every $\delta \in [0, x]$, and we are done.

Case 2: $\xi(\mathbf{u}_4) = 2$ The expression for E is now:

$$E = q(1 - (x + y)) + q(1 - (x + z)) + q(1 - (y + z)) - q(1 - x) - q(1 - y) - q(1 - (x + y + z))$$

Although $E(x, y, z)$ is different than in Case 1, the important observation is that if we consider again the function $g(\delta) = E(x - \delta, y + \delta, z)$ then the derivative $g'(\delta)$ is the same as in Case 1 and hence the same analysis shows that $E(0, x + y, z) \leq E(x, y, z)$. But if $x = 0$, then $\mathbf{y}_2$ identifies with $\mathbf{y}_4$ and the inequality reduces to the triangle inequality on $\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_3$. ■

To complete the proof, it remains to prove Proposition 2.

Proof of Proposition 2 : Fix a configuration for $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ and as before let $x = \frac{2}{n}|P_1|$, $y = \frac{2}{n}|P_2|$, $z = \frac{2}{n}|P_3|$, and $w = \frac{2}{n}|P_0|$, where $w > 0$. Consider a vector $\mathbf{u}_4$ that has all -1's in P_0. Let $H_i = \frac{2}{n}H(\mathbf{u}_i, \mathbf{u}_4)$, where $H(\mathbf{u}_i, \mathbf{u}_4)$ is the Hamming distance from $\mathbf{u}_4$ to $\mathbf{u}_i$, $i = 1, 2, 3$. It suffices to show that replacing the P_0-part of $\mathbf{u}_4$ with 1's (which means adding w to each H_i) does not decrease the LHS of (12), i.e., that:

$$q(1 - H_1) + q(1 - H_2) + q(1 - H_3) \geq q(1 - (H_1 + w)) + q(1 - (H_2 + w)) + q(1 - (H_3 + w)) \quad (13)$$

Because of the convexity of q, the cases that we need to consider are characterized by whether $\mathbf{u}_4$ is 1 or -1 on each of P_1, P_2, P_3. By symmetry there are 4 cases to check, corresponding to the different values of $\xi(\mathbf{u}_4)$. In most of these cases, we use the following argument: consider the function $g(\delta) = q(1 - (H_1 + \delta)) + q(1 - (H_2 + \delta)) + q(1 - (H_3 + \delta))$, where $\delta \in [0, w]$. Let $a_i = 1 - (H_i + \delta)$. The derivative $g'(\delta)$ is:

$$g'(\delta) = -(q'(a_1) + q'(a_2) + q'(a_3)) = -2t(a_1^{2t-1} + a_2^{2t-1} + a_3^{2t-1} + 3\lambda^{2t-1})$$

If we show that the derivative is negative for any $\delta \in [0, w]$, that would imply that $g(0) \geq g(w)$ and hence we are done since we have a more violating configuration if we do not add w to the Hamming distances.

Case 1: $\xi(\mathbf{u}_4) = 0$

In this case $H_1 = x$, $H_2 = y$, $H_3 = z$. Note that $x + y + z + w = 2$. Hence, if $H_i \geq 1$ for some i, say for H_1, then $H_2 + \delta \leq 1$ and $H_3 + \delta \leq 1$. This implies that $a_2 \geq 0$ and $a_3 \geq 0$. Thus

$$g'(\delta) \leq -(-1 + 3\lambda^{2t-1}) \leq 1 - 3/e < 0$$

since $\lambda^{2t-1} = (1 - \frac{1}{2t})^{2t-1} \geq 1/e$. Hence we are done.

Therefore, we can assume that $H_i < 1$ for all i, i.e., $1 - H_i \geq 0$. We now compare the LHS and RHS of (13). In particular we claim that each term $q(1 - H_i)$ is at least as big as the corresponding term $q(1 - (H_i + w))$. This is because of the form of the function q. Note that q is increasing in $[0, 1]$ and also that the value of q at any point $x \in [0, 1]$ is greater than the value of q at any point $y \in [-1, 0)$. Therefore since $1 - H_i > 0$ and since we only subtract w from each point, it follows that (13) holds.

Case 2: $\xi(\mathbf{u}_4) = 1$

Assume without loss of generality that $\mathbf{u}_4$ is 1 on P_1 only. In this case, $H_1 = 0$, $H_2 = x + y$ and $H_3 = x + z$. The LHS of inequality (13) is now:

$$LHS = q(1) + q(1 - (x + y)) + q(1 - (x + z))$$

whereas the RHS is:

$$RHS = q(1 - w) + q(1 - (x + y + w)) + q(1 - (x + z + w)) = q(1 - w) + q(-1 + z) + q(-1 + y)$$

by using the fact that $x + y + w = 2 - z$.

Let $\alpha_1 = 1$, $\alpha_2 = 1 - (x + y)$, $\alpha_3 = 1 - (x + z)$. The LHS is the sum of the values of q at these points whereas the RHS is the sum of the values of q after shifting each point α_i to the left by w. Let $\alpha_i' = \alpha_i - w$. The difference $\Delta = q(1) - q(1 - w)$ will always be positive since $q(1)$ is the highest value that q achieves in $[-1, 1]$. Therefore to show that (13) holds it is enough to show that the potential gain in q from shifting α_2 and α_3 is at most Δ. Suppose not and consider such a configuration. This means that either $q(\alpha_2') > q(\alpha_2)$ or $q(\alpha_3') > q(\alpha_3)$ or both. We consider the case that both points achieve a higher value after being shifted. The same arguments apply if we have only one point that improves its value. Hence we assume that $q(\alpha_2') > q(\alpha_2)$ and $q(\alpha_3') > q(\alpha_3)$. Before we proceed, we state some properties of q, which can be easily verified.

Claim 1 *The function q is decreasing in $[-1, -\lambda]$ and increasing in $[-\lambda, 1]$. Furthermore, for any 2 points x, y such that $x \in [-1, 2 - 3\lambda]$ and $y \geq 2 - 3\lambda$, $q(y) \geq q(x)$.*

Using the above claim, we argue about the location of α_2 and α_3. If $\alpha_2 \geq 2 - 3\lambda \geq -\lambda$, then $q(\alpha_2) \geq q(\alpha_2')$. Thus both α_2 and α_3 must belong to $[-1, 2 - 3\lambda] = [-1, -1 + \frac{3}{2t}]$. We will restrict further the location of α_2 and α_3 by making some more observations about q. The interval $[-1, 2 - 3\lambda]$ is the union of $A_1 = [-1, -\lambda]$ and $A_2 = [-\lambda, 2 - 3\lambda]$ and we know q is decreasing in A_1 and increasing in A_2. We claim that α_2, α_3 should belong to A_1 in the worst possible violation of (13). To see this, suppose $\alpha_2 \in A_2$ and $\alpha_3 \in A_2$ (the case with $\alpha_2 \in A_2$, $\alpha_3 \in A_1$ can be handled similarly). We know that q is the sum of a linear function and the function x^{2t}. Hence when we shift the 3 points to the left, the difference $q(1) - q(1 - w)$ is at least as big as a positive term that is linear in w. This difference has to be counterbalanced by the differences $q(\alpha_2') - q(\alpha_2)$ and $q(\alpha_3') - q(\alpha_3)$. However the form of q ensures that there is a point $\zeta_2 \in A_1$ such that $q(\alpha_2) = q(\zeta_2)$ and ditto for α_3. By considering the configuration where $\alpha_2 \equiv \zeta_2$ and $\alpha_3 \equiv \zeta_3$ we will have the same contribution from the terms $q(\alpha_2') - q(\alpha_2)$ and $q(\alpha_3') - q(\alpha_3)$ and at the same time a smaller w.

Therefore we may assume that $w \leq |A_1| = \frac{1}{2t}$. By substituting the value of q, (13) is equivalent to showing that:

$$1 - (1 - w)^{2t} + 6t\lambda^{2t-1}w \geq (\alpha_2 - w)^{2t} - \alpha_2^{2t} + (\alpha_3 - w)^{2t} - \alpha_3^{2t}$$

It is easy to see that the difference $1 - (1 - w)^{2t}$ is greater than or equal to the difference $(\alpha_2 - w)^{2t} - \alpha_2^{2t}$ by convexity. Hence it suffices to show:

$$6t\lambda^{2t-1}w \geq (\alpha_3 - w)^{2t} - \alpha_3^{2t}$$

We know that the LHS is at least $(6t/e)w$. The difference $(\alpha_3 - w)^{2t} - \alpha_3^{2t}$ can be estimated using the derivatives of x^{2t} and turns out to be at most $(6t/e)w$. Therefore no configuration in this case can violate (13).

Case 3: $\xi(\mathbf{u}_4) = 2$

Assume that $\mathbf{u}_4$ is 1 on P_1 and P_2. Now $H_1 = y$, $H_2 = x$, $H_3 = x + y + z$. The LHS and RHS of (13) are:

$$\begin{aligned} LHS &= q(1 - y) + q(1 - x) + q(1 - (x + y + z)) \\ RHS &= q(1 - (y + w)) + q(1 - (x + w)) + q(-1) \end{aligned}$$

As in case 2, let $\alpha_1 = 1 - y$, $\alpha_2 = 1 - x$ and $\alpha_3 = 1 - (x + y + z)$ be the 3 points before shifting by w. First note that either $\alpha_1 > 0$ or $\alpha_2 > 0$. This comes from the constraint that $x + y + z + w = 2$.

Assume that $\alpha_1 > 0$. Hence $q(\alpha_1) - q(\alpha_1 - w) > 0$. If $\alpha_2 \notin [-1, 2-3\lambda]$ then we would be done because by the above claim, $q(\alpha_2) - q(\alpha_2 - w) > 0$. Therefore the only way that (13) can be violated is if the nonlinear term $(\alpha_3 - w)^{2t} - \alpha_3^{2t}$ can compensate for the loss for the other terms. It can be easily checked that this cannot happen. Hence we may assume that both $\alpha_2, \alpha_3 \in [-1, 2-3\lambda]$ and that $q(\alpha_2 - w) > q(\alpha_2)$, $q(\alpha_3 - w) > q(\alpha_3)$. The rest of the analysis is based on arguments similar to case 2 and we omit it.
Case 4: $\xi(\mathbf{u}_4) = 3$ This case can also be done using similar arguments with case 2 and 3. ∎

5 Lower bound for embedding negative type metrics into ℓ_1

While, in view of Theorem 3, Charikar's metric does not supply an example that is far from ℓ_1, we may still (partly motivated by Theorem 2) utilize the idea of "tensoring the cube" and then adding some more points in order to achieve negative type metrics that are not ℓ_1 embeddable. Our starting point is an isoperimetric inequality on the cube that generalizes the standard one. Such a setting is also relevant in [20, 22] where harmonic analysis tools are used to bound expansion; these tools are unlikely to be applicable to our case where the interest and improvements lie in the constants.

Theorem 1 *(Generalized Isoperimetric inequality) For every set $S \subseteq Q_n$,*

$$|E(S, S^c)| \geq |S|(n - \log_2 |S|) + p(S).$$

where $p(S)$ denotes the number of vertices $\mathbf{u} \in S$ such that $-\mathbf{u} \in S$.

Proof. We use induction on n. Divide Q_n into two sets $V_1 = \{\mathbf{u} : \mathbf{u}_1 = 1\}$ and $V_{-1} = \{\mathbf{u} : \mathbf{u}_1 = -1\}$. Let $S_1 = S \cap V_1$ and $S_{-1} = S \cap V_{-1}$. Now, $E(S, S^c)$ is the disjoint union of $E(S_1, V_1 \setminus S_1)$, $E(S_{-1}, V_{-1} \setminus S_{-1})$, and $E(S_1, V_{-1} \setminus S_{-1}) \cup E(S_{-1}, V_1 \setminus S_1)$. Define the operator $\widehat{\cdot}$ on Q_n to be the projection onto the last $n-1$ coordinates, so for example $\widehat{S_1} = \{\mathbf{u} \in Q_{n-1} : (1, \mathbf{u}) \in S_1\}$. It is easy to observe that $|E(S_1, V_{-1} \setminus S_{-1}) \cup E(S_{-1}, V_1 \setminus S_1)| = |\widehat{S_1} \Delta \widehat{S_{-1}}|$. We argue that

$$p(S) + |S_1| - |S_{-1}| \leq p(\widehat{S_1}) + p(\widehat{S_{-1}}) + |\widehat{S_1} \Delta \widehat{S_{-1}}|. \tag{14}$$

To prove (14), for every $\mathbf{u} \in \{-1, 1\}^{n-1}$, we show that the contribution of $(1, \mathbf{u})$, $(1, -\mathbf{u})$, $(-1, \mathbf{u})$, and $(-1, -\mathbf{u})$ to the right hand side of (14) is at least as large as their contribution to the left hand side: This is trivial if the contribution of these four vectors to $p(S)$ is not more than their contribution to $p(\widehat{S_1})$, and $p(\widehat{S_{-1}})$. We therefore assume that the contribution of the four vectors to $p(S)$, $p(\widehat{S_1})$, and $p(\widehat{S_{-1}})$ are 2, 0, and 0, respectively. Then without loss of generality we may assume that $(1, \mathbf{u}), (-1, -\mathbf{u}) \in S$ and $(1, -\mathbf{u}), (-1, \mathbf{u}) \notin S$, and in this case the contribution to both sides is 2. By induction hypothesis and (14) we get

$$\begin{aligned}
|E(S, S^c)| &= |E(\widehat{S_1}, Q_{n-1} \setminus \widehat{S_1}| + |E(\widehat{S_{-1}}, Q_{n-1} \setminus \widehat{S_{-1}}| + |\widehat{S_1} \Delta \widehat{S_{-1}}| \\
&\geq |S_1|(n - 1 - \log_2 |S_1|) + p(\widehat{S_1}) + |S_{-1}|(n - 1 - \log_2 |S_{-1}|) + p(\widehat{S_{-1}}) + |\widehat{S_1} \Delta \widehat{S_{-1}}| \\
&\geq |S|n - |S| - (|S_1| \log_2 |S_1| + |S_{-1}| \log_2 |S_{-1}|) + p(\widehat{S_1}) + p(\widehat{S_{-1}}) + |\widehat{S_1} \Delta \widehat{S_{-1}}| \\
&\geq |S|n - (2|S_{-1}| + |S_1| \log_2 |S_1| + |S_{-1}| \log_2 |S_{-1}|) + p(S).
\end{aligned}$$

Now the lemma follows from the fact that $2|S_{-1}| + |S_1| \log_2 |S_1| + |S_{-1}| \log_2 |S_{-1}| \leq |S| \log_2 |S|$, which can be obtained using easy calculus. ∎

We call a set $S \subseteq Q_n$ *symmetric* if $-\mathbf{u} \in S$ whenever $\mathbf{u} \in S$. Note that $p(S) = |S|$ for symmetric sets S.

Corollary 1 *For every symmetric set $S \subseteq Q_n$*

$$|E(S, S^c)| \geq |S|(n - \log_2 |S| + 1).$$

The corollary above implies the following Poincaré inequality.

Proposition 3 *(Poincaré inequality for the cube and an additional point) Let $f : Q_n \cup \{\mathbf{0}\} \to \mathbb{R}^m$ satisfy that $f(\mathbf{u}) = f(-\mathbf{u})$ for every $\mathbf{u} \in Q_n$, and let $\alpha = \frac{\ln 2}{14 - 8\ln 2}$.*

Then the following Poincaré inequality holds.

$$\frac{1}{2^n} \cdot \frac{8}{7}(4\alpha + 1/2) \sum_{\mathbf{u},\mathbf{v} \in Q_n} \|f(\mathbf{u}) - f(\mathbf{v})\|_1 \leq \alpha \sum_{\mathbf{uv} \in E} \|f(\mathbf{u}) - f(\mathbf{v})\|_1 + \frac{1}{2} \sum_{\mathbf{u} \in Q_n} \|f(\mathbf{u}) - f(\mathbf{0})\|_1$$

Proof. It is enough to prove the above inequality for $f : V \to \{0, 1\}$. We may assume without loss of generality that $f(\mathbf{0}) = 0$. Associating S with $\{\mathbf{u} : f(\mathbf{u}) = 1\}$, the inequality of the proposition reduces to

$$\frac{1}{2^n}\frac{8}{7}(4\alpha + 1/2)|S||S^c| \leq \alpha|E(S, S^c)| + |S|/2, \tag{15}$$

where S is a symmetric set, owing to the condition $f(\mathbf{u}) = f(-\mathbf{u})$. From the isoperimetric inequality of Theorem 1 we have that $|E(S, S^c)| \geq |S|(x + 1)$ for $x = n - \log_2 |S|$ and so

$$\left(\frac{\alpha(x+1) + 1/2}{1 - 2^{-x}}\right) \frac{1}{2^n}|S||S^c)| \leq \alpha|E(S, S^c)| + |S|/2.$$

Lemma 1 below shows that $\frac{\alpha(x+1)+1/2}{1-2^{-x}}$ attains its minimum in $[1, \infty)$ at $x = 3$ whence $\frac{\alpha(x+1)+1/2}{1-2^{-x}} \geq \frac{4\alpha+1/2}{7/8}$, and Inequality (15) is proven. ■

Lemma 1 *The function $f(x) = \frac{\alpha(x+1)+1/2}{1-2^{-x}}$ for $\alpha = \frac{\ln 2}{14-8\ln 2}$ attains its minimum in $[1, \infty]$ at $x = 3$.*

Proof. The derivative of f is

$$\frac{1 - 2^{-x} - (\alpha(x + 1) + 1/2)\ln(2)2^{-x}}{(1 - 2^{-x})^2}.$$

It is easy to see that $f'(3) = 0$, $f(1) = 4\alpha + 1 > 8/7$, and $\lim_{x\to\infty} f(x) = \infty$. So it is sufficient to show that

$$g(x) = 1 - 2^{-x} - (\alpha(x + 1) + 1/2)\ln(2)2^{-x},$$

is an increasing function in the interval $[1, \infty)$. To show this note that

$$g'(x) = 2^{-x}\ln(2)\,(1 - \alpha + \alpha x \ln(2) + \alpha \ln(2)) > 0,$$

for $x \geq 1$. ■

Theorem 6 *Let $V = \{\tilde{\mathbf{u}} : \mathbf{u} \in Q_n\} \cup \{\mathbf{0}\}$, where $\tilde{\mathbf{u}} = \mathbf{u} \otimes \mathbf{u}$. Then for the semi-metric space $X = (V, \|\cdot\|^2)$ we have $c_1(X) \geq \frac{8}{7} - \epsilon$, for every $\epsilon > 0$ and sufficiently large n.*

Proof. We start with an informal description of the proof. The heart of the argument is showing that the cuts that participate in a supposedly good ℓ_1 embedding of X cannot be balanced on one hand, and cannot be imbalanced on the other. First notice that the average distance in X is almost double that of the distance between $\mathbf{0}$ and any other point (achieving this in a cube structure without violating the triangle inequality was where the tensor operation came in handy). For a cut metric on the points of X, such a relation only occurs for very imbalanced cuts; hence the representation of balanced cuts in a low distortion embedding cannot be large. On the other hand, comparing the (overall) average distance to the average distance between neighbouring points in the cube shows that any good embedding must use cuts with very small edge expansion, and such cuts in the cube must be balanced (the same argument says that one must use the dimension cuts when embedding the hamming cube into ℓ_1 with low distortion). The fact that only *symmetric cuts* participate in the ℓ_1 embedding (or else the distortion becomes infinite due to the tensor operation) enables us to use the stronger isoperimetric inequality which leads to the current lower bound. We now proceed to the proof.

We may view X as a distance function with points in $\mathbf{u} \in Q_n \cup \{\mathbf{0}\}$, and $d(\mathbf{u},\mathbf{v}) = \|\tilde{\mathbf{u}} - \tilde{\mathbf{v}}\|^2$. We first notice that X is indeed a metric space, i.e., that triangle inequalities are satisfied: notice that $X \setminus \{\mathbf{0}\}$ is a subset of $\{-1,1\}^{n^2}$. Therefore, the square Euclidean distances is the same (upto a constant) as their ℓ_1 distance. Hence, the only triangle inequality we need to check is $\|\tilde{\mathbf{u}} - \tilde{\mathbf{v}}\|^2 \le \|\tilde{\mathbf{u}} - \mathbf{0}\|^2 + \|\tilde{\mathbf{v}} - \mathbf{0}\|^2$, which is implied by the fact that $\tilde{\mathbf{u}} \cdot \tilde{\mathbf{v}} = (\mathbf{u}\cdot\mathbf{v})^2$ is always nonnegative.

For every $\mathbf{u},\mathbf{v} \in Q_n$, we have $d(\mathbf{u},\mathbf{0}) = \|\tilde{\mathbf{u}}\|^2 = \tilde{\mathbf{u}}\cdot\tilde{\mathbf{u}} = (\mathbf{u}\cdot\mathbf{u})^2 = n^2$, and $d(\mathbf{u},\mathbf{v}) = \|\tilde{\mathbf{u}} - \tilde{\mathbf{v}}\|^2 = \|\tilde{\mathbf{u}}\|^2 + \|\tilde{\mathbf{v}}\|^2 - 2(\tilde{\mathbf{u}}\cdot\tilde{\mathbf{v}}) = 2n^2 - 2(\mathbf{u}\cdot\mathbf{v})^2$. In particular, if $\mathbf{uv} \in E$ we have $d(\mathbf{u},\mathbf{v}) = 2n^2 - 2(n-2)^2 = 8(n-1)$. We next notice that

$$\sum_{\mathbf{u},\mathbf{v}\in Q_n} d(\mathbf{u},\mathbf{v}) = 2^{2n} \times 2n^2 - 2\sum_{\mathbf{u},\mathbf{v}} (\mathbf{u}\cdot\mathbf{v})^2 = 2^{2n} \times 2n^2 - 2\sum_{\mathbf{u},\mathbf{v}} (\sum_i \mathbf{u}_i\mathbf{v}_i)^2 = 2^{2n}(2n^2 - 2n),$$

as $\sum_{\mathbf{u},\mathbf{v}} \mathbf{u}_i\mathbf{v}_i\mathbf{u}_j\mathbf{v}_j$ is 2^{2n} when $i = j$, and 0 otherwise.

Let f be a nonexpanding embedding of X into ℓ_1. Notice that

$$d(\mathbf{u},-\mathbf{u}) = 2n^2 - 2(\mathbf{u}\cdot\mathbf{v})^2 = 0,$$

and so any embedding with finite distortion must satisfy $f(\mathbf{u}) = f(-\mathbf{u})$. Therefore Inequality (3) can be used and we get that

$$\frac{\alpha \sum_{\mathbf{uv}\in E} \|f(\tilde{\mathbf{u}}) - f(\tilde{\mathbf{v}})\|_1 + \frac{1}{2}\sum_{\mathbf{u}\in Q_n} \|f(\tilde{\mathbf{u}}) - f(\mathbf{0})\|_1}{\frac{1}{2^n}\sum_{\mathbf{u},\mathbf{v}\in Q_n} \|f(\tilde{\mathbf{u}}) - f(\tilde{\mathbf{v}})\|_1} \ge \frac{8}{7}(4\alpha + 1/2). \tag{16}$$

On the other hand,

$$\frac{\alpha \sum_{\mathbf{uv}\in E} d(\mathbf{u},\mathbf{v}) + \frac{1}{2}\sum_{\mathbf{u}\in Q_n} d(\mathbf{u},\mathbf{0})}{\frac{1}{2^n}\sum_{\mathbf{u},\mathbf{v}\in Q_n} d(\mathbf{u},\mathbf{v})} \frac{8\alpha(n^2 - n) + n^2}{2n^2 - 2n} = 4\alpha + 1/2 + o(1). \quad (16)$$

The discrepancy between (16) and (16) shows that for every $\epsilon > 0$ and for sufficiently large n, the required distortion of V into ℓ_1 is at least $8/7 - \epsilon$. ■

6 Discussion

We have considered the metric characterization of SDP relaxations of Vertex Cover and specifically related the amount of "ℓ_1 information" that is enforced with the resulting integrality gap. We showed that no integrality gap exists in the most powerful extreme, i.e., when ℓ_1 embeddability of the solution is enforced. We further demonstrated that integrality gap is not a continuous function of the possible distortion that is allowed, as it jumps from 1 to $2 - o(1)$ when the allowed distortion changes from 1 to $1 + \delta$. The natural extensions of these results are to (i) check whether the addition of more k-gonal inequalities (something that can be done efficiently for any finite number of such inequalities) can reduce the integrality gap or prove otherwise. It is interesting to note that related questions are discussed in the context of LP relaxations of Vertex Cover and Max Cut in [3, 11] (ii) use the nonembeddability construction and technique in Section 5 to find negative type metrics that incur more significant distortion when embedded into ℓ_1.

It is important to understand our results in the context of the Lift and Project system defined by Lovász and Schrijver [23], specifically the one that uses positive semidefinite constraints, called LS_+ (see [2] for relevant discussion). As was mentioned in the introduction, a new result of Georgiou, Magen, Pitassi and Tourlakis [13] shows that after a super-constant number of rounds of LS_+, the integrality gap is still $2-o(1)$. To relate LS_+ to SDPs one needs to use the conversion $\mathbf{y}_i = 2\mathbf{z}_i - \mathbf{z}_0$, where $\mathbf{y}_i$ is as usual the vectors of the SDP solution and the $\mathbf{z}_i$ are the Cholesky decomposition of the matrix of the lifted variables in the LS_+ system. With this relation in mind, it can be shown that the triangle inequalities with respect to $\mathbf{v}_0$ are implied after as little as one round of LS_+ and so [13] extends Charikar's result on the SDP with these types of triangle inequalities. However, at least for some graphs, triangle inequalities not involving $\mathbf{v}_0$ as well as pentagonal inequalities are *not* implied by any number of rounds of LS_+. To see this, consider the application of LS_+ system to Vertex Cover when the instance is the empty graph. Since for this instance the Linear Program relaxation is tight, lifted inequalities must appear in the first round or not at all. But it is easy to see that even the general triangle inequalities do not appear after one round and thus will never appear. It is important to note that for the graphs used in [13] (which are the same as the ones we use here) we do not know whether the general triangle inequality and whether pentagonal inequalities are implied after a few rounds of LS_+.

Last, we suggest looking at connections of ℓ_1-embeddability and integrality gaps for other NP-hard problems. Under certain circumstances, such connections may be used to convert hardness results of combinatorial problems into hardness results of approximating ℓ_1 distortion.

Acknowledgment

Special thanks to George Karakostas for very valuable discussions.

References

[1] A. Agarwal, M. Charikar, K. Makarychev, and Y. Makarychev. $O(\sqrt{\log n})$ approximation algorithms for min UnCut, min 2CNF deletion, and directed cut problems. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 573–581, New York, NY, USA, 2005. ACM Press.

[2] S. Arora, M. Alekhnovich, and I. Tourlakis. Towards strong nonapproximability results in the lovász-schrijver hierarchy. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, New York, NY, USA, 2005. ACM Press.

[3] S. Arora, B. Bollobas, L. Lovász, and I. Tourlakis. Proving integrality gaps without knowing the linear program. *Theory of Computing*, 2:19–51, 2006.

[4] S. Arora, J. Lee, and A. Naor. Euclidean distortion and the sparsest cut [extended abstract]. In *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 553–562, New York, 2005. ACM.

[5] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 222–231 (electronic), New York, 2004. ACM.

[6] M. Charikar. On semidefinite programming relaxations for graph coloring and vertex cover. In *SODA '02: Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 616–620, Philadelphia, PA, USA, 2002. Society for Industrial and Applied Mathematics.

[7] S. Chawla, A. Gupta, and H. Räcke. Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut. In *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 102–111, Vancouer, BC, Canada, 2005.

[8] N. Devanur, S. Khot, R. Saket, and N. Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, 2006.

[9] M. Deza and M. Laurent. *Geometry of cuts and metrics.* Springer-Verlag, Berlin, 1997.

[10] I. Dinur and S. Safra. On the hardness of approximating minimum vertex-cover. *Annals of Mathematics*, 162(1):439–486, 2005.

[11] W. Fernandez de la Vega and C. Kenyon-Mathieu. Linear programming relaxations of maxcut. In *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms*, 2007.

[12] P. Frankl and V. Rödl. Forbidden intersections. *Trans. Amer. Math. Soc.*, 300(1):259–286, 1987.

[13] K. Georgiou, A. Magen, T. Pitassi, and I. Tourlakis. Integrality gaps of $2-o(1)$ for vertex cover sdps in the lovász-schrijver hierarchy. In *ECCCTR: Electronic Colloquium on Computational Complexity, technical reports, TR06-152*, 2006.

[14] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995.

[15] E. Halperin. Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. *SIAM J. Comput.*, 31(5):1608–1623 (electronic), 2002.

[16] H. Hatami, A. Magen, and E. Markakis. Integrality gaps of semidefinite programs for vertex cover and relations to ℓ_1 embeddability of negative type metrics. In *RANDOM-APPROX*, 2007.

[17] G. Karakostas. A better approximation ratio for the vertex cover problem. In *Proceedings of the Thirty-Second International Colloquium on Automata, Languages and Programming*, 2005.

[18] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 767–775 (electronic), New York, 2002. ACM.

[19] S. Khot and O. Regev. Vertex cover might be hard to approximate to within $2 - \epsilon$. In *Proceedings of the 18th IEEE Conference on Computational Complexity*, pages 379–386, 2003.

[20] S. Khot and N. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proceedings of The 46-th Annual Symposium on Foundations of Computer Science*, 2005.

[21] J. Kleinberg and M. X. Goemans. The Lovász theta function and a semidefinite programming relaxation of vertex cover. *SIAM J. Discrete Math.*, 11(2):196–204 (electronic), 1998.

[22] R. Krauthgamer and Y. Rabani. Improved lower bounds for embeddings into l_1. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, 2006.

[23] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, May 1991.