Integrality gaps of $2 - o(1)$ for Vertex Cover SDPs in the Lovász-Schrijver hierarchy

Konstantinos Georgiou   Avner Magen   Toniann Pitassi   Iannis Tourlakis

Department of Computer Science
University of Toronto
10 Kings College Road, Toronto, Ontario M5S 3G4
{cgeorg, avner, toni, iannis}@cs.toronto.edu

Abstract

Linear and semidefinite programming are highly successful approaches for obtaining good approximations for NP-hard optimization problems. For example, breakthrough approximation algorithms for MAX CUT *and* SPARSEST CUT *use semidefinite programming.*

Perhaps the most prominent NP-hard problem whose exact approximation factor is still unresolved is VERTEX COVER. *PCP-based techniques of Dinur and Safra [7] show that it is not possible to achieve a factor better than 1.36; on the other hand no known algorithm does better than the factor of 2 achieved by the simple greedy algorithm. Furthermore, there is a widespread belief that SDP techniques are the most promising methods available for improving upon this factor of 2.*

Following a line of study initiated by Arora et al. [3], our aim is to show that a large family of LP and SDP based algorithms fail to produce an approximation for VERTEX COVER *better than 2. Lovász and Schrijver [21] introduced the systems LS and LS_+ for systematically tightening LP and SDP relaxations, respectively, over many rounds. These systems naturally capture large classes of LP and SDP relaxations; indeed, LS_+ captures the celebrated SDP-based algorithms for* MAX CUT *and* SPARSEST CUT *mentioned above.*

We rule out polynomial-time $2 - \Omega(1)$ approximations for VERTEX COVER *using LS_+. In particular, we prove an integrality gap of $2-o(1)$ for* VERTEX COVER *SDPs obtained by tightening the standard LP relaxation with $\Omega(\sqrt{\log n/\log\log n})$ rounds of LS_+. While tight integrality gaps were known for* VERTEX COVER *in the weaker LS system [23], previous results did not rule out a $2-\Omega(1)$ approximation after even two rounds of LS_+.*

Funded in part by NSERC

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. The minimum VERTEX COVER problem asks what size the minimum vertex cover in G is. Determining how well we can approximate VERTEX COVER is one of the outstanding open problems in the complexity of approximation: while VERTEX COVER has a trivial 2-approximation algorithm, no better approximation algorithms are known.

This contrasts with the situation for another famous problem, MAX CUT: for many years, no approximation algorithm was known that could yield better than a $(0.5 + o(1))$-approximation (the trivial randomized algorithm gives a 0.5-approximation) until the seminal paper of Goemans and Williamson [12] which used semidefinite programming (SDP) to obtain a 0.878-approximation algorithm. Since then semidefinite programming has yielded breakthrough approximation algorithms for various NP-hard optimization problems and has arguably become our most powerful tool for designing approximation algorithms. Consequently, semidefinite programming is believed (see Lovász [20] for instance) to be the most promising technique for attacking the VERTEX COVER problem.

However, Kleinberg and Goemans [19] showed in '95 that the standard SDP for VERTEX COVER has an integrality gap of $2 - o(1)$. Subsequently, Charikar [6] showed that the integrality gap remains $2 - o(1)$ even if we add additional triangle inequality constraints.

Hatami, Magen and Markakis [14] strengthened this further, showing that this state of affairs remains even when we add the so-called pentagonal inequality constraints.

Indeed, the state of the art is such that SDP-based algorithms for VERTEX COVER must settle for competing in "how big" the "little oh" term is in the $2 - o(1)$ factor. Halperin [13] gives a $(2 - \log\log\Delta/\log\Delta)$-approximation, where Δ is the maximal degree of the graph. The best approximation algorithm currently known for arbitrary graphs is due to Karakostas [16] who obtains a $(2 - \Omega(1/\sqrt{\log n}))$-approximation algorithm using a stronger SDP relaxation.

Nevertheless, it is consistent with the known hardness results for VERTEX COVER that there could be some other SDP with integrality gap, say, 1.4. In particular, the best PCP-based hardness result known (Dinur and Safra [7]) only shows that 1.36-approximation of VERTEX COVER is NP-hard. Only by assuming Khot's Unique Games Conjecture [17] do we get a tight $2-o(1)$ inapproximability result [18]. However, determining the validity of the Unique Games Conjecture (or directly improving on [7]) remains a difficult open problem.

To get a better picture of the approximability of VERTEX COVER (especially in light of the inability to resolve the issue with PCP-based methods), Arora et al. [3] suggested the following approach: rule out good approximations by large families of algorithms. One such family is the class of relaxations for VERTEX COVER in the Lovász-Schrijver hierarchies. Lovász and Schrijver [21] define procedures LS and LS_+ for systematically tightening linear and semidefinite relaxations, respectively, over many rounds. Important algorithmic properties of LS and LS_+ are: (a) n rounds of even the weaker LS procedure suffice to obtain exact solutions and (b) we can optimize a linear function over the rth round LS and LS_+ relaxations in $n^{O(r)}$ time (provided the original relaxation had a polynomial-time separation oracle).

Many celebrated SDP-based algorithms, including the seminal MAX CUT algorithm of Goemans-Williamson [12] and the Arora-Rao-Vazirani algorithm [4] for SPARSEST CUT, can be derived using a constant number of rounds of LS_+. Thus proving inapproximability results for LS_+ based algorithms rules out one of the most promising classes of algorithms that we currently have for obtaining $2 - \Omega(1)$ approximations for VERTEX COVER. Furthermore, unlike PCP-based results we emphasize that such results do not rely on any complexity theoretic assumptions.

Arora et al. [3] obtained the first result along these lines for VERTEX COVER showing that $\Omega(\log n)$ rounds of the weaker LS procedure has an integrality gap of $2 - o(1)$. Tourlakis [25] subsequently proved an integrality gap of $1.5 - o(1)$ for VERTEX COVER for $\Omega(\log^2 n)$ rounds of LS. Very recently, a beautiful result by Schoenebeck, Trevisan and Tulsiani [23] showed that the integrality gap is $2-o(1)$ even after $\Omega(n)$ rounds of LS. Unfortunately, the hard examples used in these papers cannot be used to prove a $2-o(1)$ integrality gap for even one round of LS_+.

The only known integrality gaps for VERTEX COVER LS_+ relaxations prior to the current paper were proved by Schoenebeck, Trevisan and Tulsiani [22] who showed that the integrality gap remains $7/6$ for $\Omega(n)$ rounds of LS_+. The graphs they use are obtained using the standard FGLSS [8] reduction from MAX-3XOR to VERTEX COVER. Such instances cannot prove stronger integrality gaps for LS_+ since their integrality gaps are *at most* $7/6$ after one round of LS_+.

To summarize, previously known results do not preclude a polynomial time $2 - \Omega(1)$ approximation algorithm for VERTEX COVER using LS_+ tightenings. In particular, showing a $2 - o(1)$ integrality gap for even two rounds of LS_+ remained a challenging open problem (Charikar's construction [6] does imply a $2 - o(1)$ gap for one round).

In this paper we rule out such approximations. Our starting point is the graph families used to show $2 - o(1)$ integrality gaps for various VERTEX COVER SDPs in [19, 6, 14] (similar graphs were used by Alon and Kahale [2] in independent work contemporaneous with [19] studying the Lovász theta function). We briefly describe these graphs. The vertex set is $\{-1, 1\}^m$ and two vertices are adjacent if their Hamming distance is exactly $(1 - \gamma)m$. A result of Frankl and Rödl [10] bounds from above the size of any independent set in such graphs by $m(2 - \Omega(\gamma^2))^m$. Hence, for constant $\gamma > 0$ (or even γ a slowly vanishing function of m) any vertex cover has size $(1 - o(1))|V|$. Of course for $\gamma = 0$ these graphs are just perfect matchings on 2^m vertices. The cleverness of the construction lies in how a minuscule increase in γ dramatically changes the independent set size while not appreciably altering the "geometry" of the graph (and hence not appreciably increasing the SDP value from the perfect matching case).

We use this graph family to show that $\Omega(\sqrt{\log n/\log\log n})$ rounds of LS_+ has an integrality gap of $2 - o(1)$ for VERTEX COVER. Our main theorem also implies that the integrality gap remains at least $2 - O(\sqrt{\log\log n/\log n})$ after $O(1)$ rounds of LS_+. Hence, the approximation ratio achieved by Karakostas' [16] algorithm is essentially tight for "polynomial" time LS_+ relaxations. Our main technical tool is the construction of a sequence of tensoring operations on vectors. These operations have the property that inner products on the set of tensored vectors are a polynomial function of the inner products of the original vectors. These extend similar tensoring operations used by Charikar [6] (and implicit in earlier work by Kahn and Kalai [15]). However, our application calls for more complicated polynomials, and moreover the polynomials (and hence the tensored vectors) change as the induction unwinds in our lower bound argument (details in Section 3).

Section 2 contains all necessary definitions including a description of LS_+. Section 3 outlines our approach while Section 4 contains the proof of our main result. Section 5 discusses limitations of our approach and poses some open problems.

2 Definitions, Notation and Tools

2.1 Standard SDPs for Vertex Cover

The standard way to formulate Vertex Cover for a graph $G = (V, E)$ as a quadratic integer program is:

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

The set of vertices i for which $x_i = x_0$ corresponds to the minimal vertex cover. This quadratic program leads to the following semidefinite programming relaxation:

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

We can strengthen this relaxation by adding the vector analogues of constraints valid in the integral case. Examples are the triangle and "extended" triangle inequalities (respectively),

$$(\mathbf{v}_i - \mathbf{v}_j) \cdot (\mathbf{v}_i - \mathbf{v}_k) \geq 0 \quad \forall i, j, k \in \{0\} \cup V, \tag{2}$$

$$(\mathbf{v}_i \pm \mathbf{v}_j) \cdot (\mathbf{v}_i \pm \mathbf{v}_k) \geq 0 \quad \forall i, j, k \in \{0\} \cup V. \tag{3}$$

Relaxation (1) was studied in [19]. The SDP tightened using (2) was studied in [6] while the SDP tightened using (2) *and* (3) (as well as the so-called pentagonal inequalities) was studied in [14].

2.2 Lovász-Schrijver Lift-and-Project

A convex cone is a set $K \subseteq \mathcal{R}^{n+1}$ such that for every $\mathbf{y}, \mathbf{z} \in K$, and for every $\alpha, \beta \geq 0$, $\alpha\mathbf{y} + \beta\mathbf{z} \in K$. Let $\mathbf{e}_i$ denote the vector with 1 in coordinate i and 0 everywhere else. Hence, $Y\mathbf{e}_i$ denotes the ith column of a matrix Y. If $K \subseteq \mathbb{R}^{n+1}$ is a convex cone, $M_+(K) \subseteq \mathbb{R}^{(n+1)\times(n+1)}$ consists of all symmetric $(n+1)\times(n+1)$ matrices Y such that,

1. For all $i = 0, 1, \ldots, n$, $Y_{0i} = Y_{ii}$.
2. For all $i = 0, 1, \ldots, n$, $Y\mathbf{e}_i, Y\mathbf{e}_0 - Y\mathbf{e}_i \in K$.
3. Y is positive semidefinite (PSD).

We then define $N_+(K) = \{Y\mathbf{e}_0 : Y \in M_+(K)\} \subseteq \mathbb{R}^{n+1}$. That is, a vector $\mathbf{y} = (y_0, \ldots, y_n)$ is in $N_+(K)$ if there exists $Y \in M_+(K)$ such that $Y\mathbf{e}_0 = \mathbf{y}$ in which case Y is called a *protection matrix for* $\mathbf{y}$. Define $N_+^k(K)$ inductively by setting $N_+^0(K) = K$ and $N_+^k(K) = N_+(N_+^{k-1}(K))$.

Let $G = (V, E)$ be a graph and assume that $V = \{1, \ldots, n\}$. The Vertex Cover convex cone for G, $VC(G)$, is the set of vectors $\mathbf{y} \in \mathbb{R}^{n+1}$ such that:

$$y_i + y_j \geq y_0 \quad \text{for all } ij \in E \tag{4}$$

$$y_0 \geq y_i \geq 0 \quad \text{for all } i \in V \tag{5}$$

$$y_0 \geq 0$$

Constraints (4) are called the *edge constraints* and constraints (5) are called the *box constraints*.

The value of the Vertex Cover relaxation arising from k rounds of LS_+ is the solution of

$$\begin{array}{ll} \min & \sum_{i=1}^{n} y_i \\ \text{s.t.} & (y_0, \ldots, y_n) \in N_+^k(VC(G)) \text{ and } y_0 = 1 \end{array}$$

The *integrality gap* of this relaxation (for n-vertex graphs) is the largest ratio between the minimum vertex cover size of G and the optimum in the above program, over all n-vertex graphs G.

To get an idea of the power of LS_+, we note first that the relaxation $N_+(VC(G))$ is at least as strong as the standard SDP relaxation for Vertex Cover since the Cholesky decomposition of any matrix $Y \in M_+(VC(G))$ satisfies (under an affine transformation) SDP (1). In fact, it even satisfies the triangle inequalities (2) for the case $i = 0$. On the other hand, one can show that adding both the standard and "extended" triangle inequalities (constraints (2) and (3), respectively) to the standard Vertex Cover SDP results in a relaxation at least as strong as $N_+(VC(G))$. Indeed, we will (implicitly) exploit the latter fact when constructing SDP solutions for our lower bound.

2.3 Vectors and Tensoring

We will use $\mathbf{0}$ to denote the all-0 vector. Given two vectors $\mathbf{x}, \mathbf{y} \in \{-1, 1\}^n$ their *Hamming distance*

$d_H(\mathbf{x},\mathbf{y})$ is $|\{i \in [n] : x_i \neq y_i\}|$. 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 on the first n coordinates is $\mathbf{u}$ and on the last m coordinates is $\mathbf{v}$.

Recall that the tensor product $\mathbf{u} \otimes \mathbf{v}$ of vectors $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$is the vector in $\mathbb{R}^{nm}$ indexed by ordered pairs from $n \times m$ and that assumes the value $\mathbf{u}_i\mathbf{v}_j$ at coordinate (i,j). Define $\mathbf{u}^{\otimes d}$ to be the vector in $\mathbb{R}^{n^d}$ obtained by tensoring $\mathbf{u}$ with itself d times.

Definition 1 *Let $P(x) = c_1x^{t_1} + \ldots + c_qx^{t_q}$ be a polynomial with nonnegative coefficients. Then we define T_P to be the function that maps a vector $\mathbf{u}$ to the vector $T_P(\mathbf{u}) = (\sqrt{c_1}u^{\otimes t_1}, \ldots, \sqrt{c_q}u^{\otimes t_q})$.*

Fact 1 *For all $\mathbf{u},\mathbf{v} \in \mathbb{R}^d$, $T_P(\mathbf{u}) \cdot T_P(\mathbf{v}) = P(\mathbf{u} \cdot \mathbf{v})$.*

2.4 Frankl-Rödl Graphs

Definition 2 *Fix γ, $0 \leq \gamma \leq 1$ and an integer $m \geq 1$. The* Frankl-Rödl graph *G_m^γ is the graph with vertices $\{-1,1\}^m$ and where two vertices $i,j \in \{-1,1\}^m$ are adjacent if $d_H(i,j) = (1-\gamma)m$.*

Relatives of the following lemma appear in [10] in various guises, but it seems as if the exact statement that we will use requires a further small step which we sketch in Appendix A. The key difference with variants in [10] is that we explicitly allow γ to be a function of m.

Lemma 1 *Let m be an integer and let $\gamma = \gamma(m) > 0$ be a sufficiently small number so that $\gamma \cdot m$ is an even integer. Then there are no independent sets in G_m^γ of size larger than $m2^m(1-\gamma^2/64)^m$.*

2.5 Saturated Vectors

In general, our lower bounds will be proved by arguing about vectors whose coordinates are either 0/1 or take on at most one other fixed value. The following definition formalizes this.

Definition 3 *A vector $\mathbf{y} \in [0,1]^{n+1}$ is an* ϵ-vector *if $y_0 = 1$ and $y_i \in \left\{0, \frac{1}{2}+\epsilon, 1\right\}$ for all $1 \leq i \leq n$.*

Note that ϵ-vectors have the property that the sum of any two non-0/1 coordinates is $1+2\epsilon$. A weaker condition on vectors in $[0,1]^{n+1}$ would be to only require that the sum of any two non-0/1 coordinates is at least $1+2\epsilon$. Such vectors were used in [23] and the following definition is adapted from their paper:

Definition 4 ([23]) *Let $G = (V,E)$ be a graph. A vector $\mathbf{y} \in VC(G)$ is* ϵ-saturated *if for every edge $ij \in E$ such that y_i and y_j are both not integral, $y_i + y_j \geq 1+2\epsilon$.*

Saturated vectors have the following important property proved in [23] (we include a proof in Appendix B for completeness):

Lemma 2 ([23]) *Let $G = (V,E)$ be any graph and suppose $\mathbf{x} \in VC(G)$ is ϵ-saturated. Then $\mathbf{x}$ is a convex combination of ϵ-vectors in $VC(G)$.*

The lemma essentially says that proving lower bounds for ϵ-saturated vectors reduces to proving lower bounds for ϵ-vectors. This will be crucial for our arguments since we only know how to find protection matrices for ϵ-vectors. We remark that our definition for saturation is slightly different than the one in [23] as there they only require that *one* of y_i or y_j in Definition 4 be non-integral. Consequently, Lemma 2 becomes somewhat stronger to accommodate this difference, but the additional argument for this strengthening is trivial (see Appendix B).

3 Overview of the Proof

We start with a Frankl-Rödl graph $G = G_m^\gamma$ and denote by $n = 2^m$ the size of G. We will show that the point $\mathbf{x} = (1, 1/2+\epsilon, \ldots, 1/2+\epsilon)$ is contained in the polytope defined after $\Omega(\sqrt{\log n / \log\log n})$ rounds of LS_+. This clearly gives us our desired $2 - o(1)$ integrality gap.

The standard way to prove that a certain point $\mathbf{x}$ is in the polytope resulting from r rounds of LS_+ (hereafter, the "rth polytope") is as follows: (1) Exhibit a symmetric PSD "protection" matrix Y for $\mathbf{x}$ such that the diagonal and first column of Y equal $\mathbf{x}$. (2) Show inductively that the vectors Ye_i and $Y(\mathbf{e}_0 - \mathbf{e}_i)$ are in the $(r-1)$st polytope. By definition of LS_+ it will then follow that $\mathbf{x}$ is in the rth polytope.

To define a protection matrix for $\mathbf{x}$ we will start with the canonical set of vectors associated with the vertices of G, namely the normalized versions of the vectors $\{-1,1\}^m$ (these vectors were also the starting point for [19, 6, 14]). These vectors have the appealing property that the inner product of vectors associated with two vertices i and j is solely a function of their Hamming distance $d_H(i,j)$. Observe that this property will not be compromised by applying the T_P tensoring transformation to the vectors. Indeed, we will use this tensoring transformation with a specific polynomial P to obtain a new set of tensored vectors and then define our candidate protection matrix to be essentially the Gram matrix of these vectors. (Note that Charikar [6] also uses a tensor transformation to prove his integrality gap for the SDP with triangle inequalities.)

A consequence of the observation above is that the values on the diagonal of the Gram matrix are all identical. So this protection matrix recipe only works for vectors like $\mathbf{x}$ where all fractional values are the same. In fact, for technical reasons which we do not get into in this outline, this recipe produces valid protection matrices only when $\mathbf{x}$ is a ρ-vector for some $0 < \rho < 1/2$.

To continue our inductive argument we would in turn like to use the same recipe to find candidate protection matrices for each of the $2n$ vectors $Y\mathbf{e}_i$ and $Y(\mathbf{e}_0 - \mathbf{e}_i)$ (or, more accurately, for the projections of these vectors onto the hyperplane $x_0 = 1$). The problem is that while these $2n$ vectors may indeed be in the $(r-1)$st polytope, they may not be ρ-vectors. (This is because the entries Y_{ij} of $Y\mathbf{e}_i$ are a polynomial function of $d_H(i,j)$ and the latter is distributed like a binomial distribution when i is fixed.) So the recipe cannot be used without extra work.

To remedy the situation, we will apply a "correction" phase as follows. (Note that "correction" phases of some sort or another can be found in many previous works [3, 1, 5, 25, 22, 23].) We will construct the tensored vectors so that the vectors $Y\mathbf{e}_i$, $Y(\mathbf{e}_0 - \mathbf{e}_i)$ have high saturation. We will then use Lemma 2 to express these vectors as convex combinations of ρ'-vectors from $VC(G)$ for some $\rho' > 0$ (this is the "correction" part). We then carry on the induction with these ρ'-vectors to show that they lie in the $(r-1)$st polytope. Convexity then implies that the vectors $Y\mathbf{e}_i$, $Y(\mathbf{e}_0 - \mathbf{e}_i)$ are also in the $(r-1)$st polytope.

To summarize, we start with a vector $\mathbf{x} = (1, 1/2 + \epsilon_0, \ldots, 1/2 + \epsilon_0)$, $\epsilon_0 = \epsilon$, and after one round we need to show that the $2n$ vectors $Y\mathbf{e}_i$, $Y(\mathbf{e}_0 - \mathbf{e}_i)$ corresponding to $\mathbf{x}$'s protection matrix Y have large saturation ϵ_1; and then we continue with vectors with fractional values $1/2 + \epsilon_1$, and so on. In this process, the obvious objective is to make the sequence $\epsilon_0, \epsilon_1, \epsilon_2, \ldots$ as slowly decreasing as possible, thereby making it last for many rounds before it becomes negative (which amounts to negative saturation, and hence that the corresponding vectors are not in $VC(G)$ at all). We will show that for each round i, we can ensure that $\epsilon_i = \epsilon_{i-1} - O(\gamma)$. Thus for arbitrarily small initial ϵ_0, we get an induction chain of length $\Omega(\epsilon_0/\gamma)$.

The engine of this process and our main technical tool are the tensor-inducing polynomials. Along with the sequence of decreasing saturation values we shall have a sequence of polynomials with positive coefficients, $P_0, P_1, P_2, \ldots$ where P_i depends on ϵ_i and determines ϵ_{i+1}. The choice of this sequence is at the heart of the matter. The non-negativity requirement on the coefficients is what makes this a challenging task. Charikar [6] used a polynomial designed to produce vectors that satisfy the triangle inequality. This polynomial is the sum of a linear term and a degree $O(1/\gamma)$ monomial that unfortunately produces a poor saturation, and hence cannot be used to proceed beyond one round of LS_+. In particular, the saturation it provides is about $1/m \ll \gamma$. The problem is intrinsic: let us suppose that we are dealing with $Y(\mathbf{e}_0 - \mathbf{e}_i)$ for some fixed i. It is easy to see that no matter which polynomial we use, edges incident to vertex i will have no slack at all in $Y(\mathbf{e}_0 - \mathbf{e}_i)$. Such an edge ij will not in itself affect the saturation as its vertices will have integral values; however, the continuous nature of the construction means that nearby edges $i'j'$ will not have integral values since their values will correspond to evaluating the polynomial at points only slightly different than those for ij. But then, to ensure that $i'j'$ has good saturation, our polynomial must vary a lot between the cases corresponding to ij and $i'j'$. This calls for a polynomial with a very large derivative, and hence one with very high degree $d \gg m$; in contrast, the polynomial that Charikar uses has degree independent of m.

4 Main Theorem

Lemma 3 *Let m be a sufficiently large integer and $\gamma > 0$. Let $n = 2^m$ and let ϵ be a sufficiently small constant such that $\epsilon > 5\gamma$. Suppose in addition that $\mathbf{y} \in \mathbb{R}^{n+1}$ is an ϵ-vector in $VC(G_m^\gamma)$. Then there exists a protection matrix Y for $\mathbf{y}$ such that for all i with $0 < y_i < 1$, Ye_i/y_i and $Y(e_0 - e_i)/(1-y_i)$ are convex combinations of $(\epsilon - 6\gamma)$-vectors. In particular, $y \in N_+(VC(G_m^\gamma))$.*

Given Lemma 3, we can prove our main theorem from which the integrality gaps for LS_+ stated in the introduction immediately follow.

Theorem 5 *Let m be sufficiently large, and fix $\gamma \geq 12\sqrt{\frac{\log m}{m}}$ such that γm are all even. Let ϵ be a sufficiently small constant such that $\epsilon > 5\gamma$. Let $n = 2^m$ and let $r = \lfloor \frac{\epsilon}{6\gamma} \rfloor - 1$. Then the integrality gap of $N_+^r(VC(G_m^\gamma))$ is at least $2 - 4\epsilon - 2/m$.*

Proof: Let $\mathbf{y} = (1, \frac{1}{2} + \epsilon, \ldots, \frac{1}{2} + \epsilon) \in \mathbb{R}^{n+1}$. Clearly $\mathbf{y} \in VC(G_m^\gamma)$. A simple inductive argument using Lemma 3 then implies that $\mathbf{y} \in N_+^r(VC(G_m^\gamma))$.

On the other hand, Lemma 1 implies that the largest independent set in G_m^γ has size at most

$$2^m m\left(1-\frac{\gamma^2}{64}\right)^m \le \frac{m2^m}{e^{\frac{\gamma^2 m}{64}}} \le \frac{m2^m}{e^{\frac{144}{64}\log m}} \le \frac{2^m}{m}.$$

Hence, the integrality gap for $N_+^r(VC(G_m^\gamma))$ is at least, $\frac{2^m-2^m/m}{n(\frac{1}{2}+\epsilon)} = \frac{2(1-1/m)}{1+2\epsilon} \ge 2-4\epsilon-\frac{2}{m}$. $\square$

4.1 Proof of Lemma 3

Fix m and γ and consider $G = G_m^\gamma$. Denote the vertices V of G as vectors $\mathbf{w}_i \in \{-1,1\}^m$, $1 \le i \le 2^m$, and for each vector $\mathbf{w}_i \in V$ define $\mathbf{u}_i = \frac{1}{\sqrt{m}}\mathbf{w}_i$. Note that $\|\mathbf{u}_i\| = 1$ for all $i \in V$ and $\mathbf{u}_i \cdot \mathbf{u}_j = 2\gamma - 1$ for all $ij \in E$. Moreover, $-1 \le \mathbf{u}_i \cdot \mathbf{u}_j \le 1 - \frac{2}{m}$ for all $1 \le i < j \le 2^m$.

Given a polynomial P with nonnegative coefficients we will now define a procedure that takes the vectors $\{\mathbf{u}_i\}$, applies the tensoring operation T_P from Section 2.3 to obtain a new set of vectors, and then applies a linear transformation to the resulting vectors. The Gram matrix of the vectors resulting from this procedure will be called $Y(P,\mathbf{y})$. Our goal will be to pick P so that $Y(P,\mathbf{y})$ is a protection matrix for $\mathbf{y}$.

First, define $\mathbf{v}_0 = (1,0,\ldots,0)$. For each vertex $1 \le i \le 2^m$ define,

$$\mathbf{v}_i = \begin{cases} \mathbf{v}_0, & \text{if } y_i = 1 \\ \mathbf{0}, & \text{if } y_i = 0 \\ (\frac{1}{2}+\epsilon, \frac{\sqrt{1-4\epsilon^2}}{2}\cdot T_P(\mathbf{u}_i)), & \text{if } y_i = \frac{1}{2}+\epsilon \end{cases}$$

Let $Y(P,\mathbf{y}) \in \mathbb{R}^{(n+1)\times(n+1)}$ be the PSD matrix defined by $Y(P,\mathbf{y})_{ij} = \mathbf{v}_i \cdot \mathbf{v}_j$. We define a class of polynomials and show that for any polynomial P in this class, $Y(P,\mathbf{y})$ is a protection matrix for $\mathbf{y}$.

Definition 6 *A polynomial $P(x)$ is called (γ,ϵ,m)-*useful *if it satisfies the following conditions:*

1. *P has only nonnegative coefficients.*
2. *$P(1) = 1$,*
3. *$P(x) \ge P(2\gamma-1) = -\frac{1-2\epsilon}{1+2\epsilon}$ for all $x \in [-1,1]$.*
4. *For all $i \in \{1,\ldots,2^m\}$ and all $jk \in E$,*

$$-\frac{4\epsilon}{1-2\epsilon} \le P(\mathbf{u}_i\cdot\mathbf{u}_j) + P(\mathbf{u}_i\cdot\mathbf{u}_k) \le \frac{4\epsilon}{1+2\epsilon}. \quad (6)$$

Claim 1 *If P is (γ,ϵ,m)-useful, then $Y = Y(P,\mathbf{y}) \in M_+(VC(G))$. In particular, Y is a protection matrix for $\mathbf{y}$ and hence, $\mathbf{y} \in N_+(VC(G))$.*

Proof: Since Y is PSD by definition, to show that Y is a protection matrix for $\mathbf{y}$ it suffices to show that: **A)** for all $0 \le i \le n$, $Y_{i0} = Y_{ii} = y_i$, and **B)** for all $1 \le i \le n$, $Y\mathbf{e}_i, Y(\mathbf{e}_0 - \mathbf{e}_i) \in VC(G)$.

Consider **A** first. Clearly $Y_{i0} = Y_{ii} = y_i$ whenever $y_i \in \{0,1\}$. In particular, note that $Y_{00} = 1$. So assume that $y_i = 1/2+\epsilon$. Clearly $Y_{i0} = \frac{1}{2}+\epsilon$, so consider Y_{ii}. We have

$$\begin{aligned} Y_{ii} = \mathbf{v}_i\cdot\mathbf{v}_i &= \left(\frac{1}{2}+\epsilon\right)^2 + \frac{1-4\epsilon^2}{4}T_P(\mathbf{u}_i)\cdot T_P(\mathbf{u}_i) \\ &= \frac{1}{4}+\epsilon+\epsilon^2+\frac{1-4\epsilon^2}{4}P(\mathbf{u}_i\cdot\mathbf{u}_i) = \frac{1}{2}+\epsilon, \end{aligned}$$

where the last equality follows from the fact that the $\mathbf{u}_i$ are unit vectors and $P(1) = 1$.

Now consider **B**. We must show that for $1 \le i \le n$, $Y\mathbf{e}_i$ and $Y(\mathbf{e}_0 - \mathbf{e}_i)$ both satisfy the edge constraints (4) and the box constraints (5). Note that if $y_i \in \{0,1\}$, then $\{Y\mathbf{e}_i, Y(\mathbf{e}_0-\mathbf{e}_i)\} = \{\mathbf{0}, Y\mathbf{e}_0\} \subseteq VC(G)$ and these constraints are trivially satisfied. So assume $y_i = \frac{1}{2}+\epsilon$.

The box constraints require for all $1 \le j \le n$ that $0 \le Y_{ij} \le Y_{i0}$ and $0 \le Y_{0j} - Y_{ij} \le Y_{00} - Y_{i0}$. Equivalently, for all $1 \le j \le n$,

$$Y_{i0} + Y_{j0} - Y_{00} \le Y_{ij} \le Y_{i0}. \quad (7)$$

On the other hand, the edge constraints require for all $1 \le i \le n$ and all $jk \in E$ that

$$Y_{ij} + Y_{ik} \ge Y_{i0}, \quad (8)$$

$$(Y_{0j} - Y_{ij}) + (Y_{0k} - Y_{ik}) \ge Y_{00} - Y_{i0}. \quad (9)$$

Since (7) holds when $y_i \in \{0,1\}$, by symmetry it also holds if $y_j \in \{0,1\}$. So assume $y_j = \frac{1}{2}+\epsilon$. We first show that the right inequality in (7) holds. Fix $j \in \{1,\ldots,n\}$. Note that since $P(1) = 1$, it follows that $\|\mathbf{v}_i\| = \|\mathbf{v}_j\|$. So, $Y_{ij} = \mathbf{v}_i\cdot\mathbf{v}_j \le \|\mathbf{v}_i\|^2 = Y_{ii} = Y_{i0}$.

Now consider the left inequality in (7). We have that,

$$\begin{aligned} Y_{ij} + Y_{00} - Y_{i0} - Y_{j0} &= Y_{ij} - 2\epsilon \\ &= \left[\frac{1}{4}+\epsilon+\epsilon^2+\frac{1-4\epsilon^2}{4}T_P(\mathbf{u}_i)\cdot T_P(\mathbf{u}_j)\right] - 2\epsilon \\ &= \frac{1}{4}-\epsilon+\epsilon^2+\frac{1-4\epsilon^2}{4}P(\mathbf{u}_i\cdot\mathbf{u}_j) \ge 0, \end{aligned}$$

where the last inequality follows by Property 3 of a (γ,ϵ,m)-useful polynomial and the fact that the $\mathbf{u}_i$ are unit vectors. So (7) holds.

Now consider the remaining constraints. Fix $j,k \in \{0,1,\ldots,2^m\}$. Using constraints (7), the fact that $Y_{ii} = Y_{i0}$ for all i, and the fact that $\mathbf{y}$ is an ϵ-vector in $VC(G)$, it is easy to verify that constraints (8) and (9) hold whenever one of y_j or y_k are integral. So assume $y_j = y_k = \frac{1}{2} + \epsilon$.

Constraint (8) then holds if the following is at least 1:

$$\begin{aligned}
&\frac{Y_{ij} + Y_{ik}}{Y_{i0}} \\
&= 2\left(\frac{1}{2} + \epsilon\right) + \frac{1-2\epsilon}{2} T_P(\mathbf{u}_i) \cdot (T_P(\mathbf{u}_j) + T_P(\mathbf{u}_k)) \\
&= 1 + 2\epsilon + \frac{1-2\epsilon}{2}(P(\mathbf{u}_i \cdot \mathbf{u}_j) + P(\mathbf{u}_i \cdot \mathbf{u}_k)). \quad (10)
\end{aligned}$$

Similarly, (9) holds if the following is at least 1:

$$\begin{aligned}
&\frac{(Y_{0j} - Y_{ij}) + (Y_{0k} - Y_{ik})}{Y_{00} - Y_{i0}} \\
&= 1 + 2\epsilon - \frac{1+2\epsilon}{2}(P(\mathbf{u}_i \cdot \mathbf{u}_j) + P(\mathbf{u}_i \cdot \mathbf{u}_k)). \quad (11)
\end{aligned}$$

But by Property 4 of a (γ, ϵ, m)-useful polynomial, for all $1 \le i \le n$ and all $jk \in E$, equations (10) and (11) are indeed both at least 1. The claim follows. □

By Lemma 2, to complete the proof of Lemma 3 it suffices to show that there exists a (γ, ϵ, m)-useful polynomial P such that if $Y = Y(P, \mathbf{y})$, then for all i such that $y_i = \frac{1}{2} + \epsilon$ the vectors $Y\mathbf{e}_i/y_i$ and $Y(\mathbf{e}_0 - \mathbf{e}_i)/(1-y_i)$ are $(\epsilon - 6\gamma)$-saturated. (The vectors $Y\mathbf{e}_i/y_i$ and $Y(\mathbf{e}_0 - \mathbf{e}_i)/(1-y_i)$ are the "normalized" versions of Ye_i and $Y(\mathbf{e}_0 - \mathbf{e}_i)$, i.e., their projections onto the hyperplane $x_0 = 1$.)

To that end, let us first compute the saturation of these vectors for an arbitrary but fixed (γ, ϵ, m)-useful polynomial P. Fix i such that $y_i = \frac{1}{2} + \epsilon$ and consider $Y\mathbf{e}_i/y_i$. Let $I = \{i\} \cup \{j : y_j \in \{0, 1\}\}$. Then the saturation of $Y\mathbf{e}_i/y_i$ is at least

$$\begin{aligned}
&\min_{j,k \notin I, jk \in E} \frac{1}{2}((Y_{ij} + Y_{ik})/y_i - 1) \\
&= \min_{j,k \notin I, jk \in E} \left[\epsilon + \frac{1-2\epsilon}{4}(P(\mathbf{u}_i \cdot \mathbf{u}_j) + P(\mathbf{u}_i \cdot \mathbf{u}_k))\right] \\
&\ge \min_{j,k \ne i, jk \in E} \left[\epsilon + \frac{1-2\epsilon}{4}(P(\mathbf{u}_i \cdot \mathbf{u}_j) + P(\mathbf{u}_i \cdot \mathbf{u}_k))\right],
\end{aligned}$$

where the equality follows by (10) and the fact that $y_j, y_k \notin \{0, 1\}$. Similarly, the saturation of $Y(\mathbf{e}_0 - \mathbf{e}_i)/(1-y_i)$ is at least

$$\begin{aligned}
&\min_{j,k \notin I, jk \in E} \frac{1}{2}\left(\frac{(Y_{0j} - Y_{ij}) + (Y_{0k} - Y_{ik})}{1 - y_i} - 1\right) \\
&= \min_{j,k \notin I, jk \in E} \left[\epsilon - \frac{1+2\epsilon}{4}(P(\mathbf{u}_i \cdot \mathbf{u}_j) + P(\mathbf{u}_i \cdot \mathbf{u}_k))\right] \\
&\ge \min_{j,k \ne i, jk \in E} \left[\epsilon - \frac{1+2\epsilon}{4}(P(\mathbf{u}_i \cdot \mathbf{u}_j) + P(\mathbf{u}_i \cdot \mathbf{u}_k))\right],
\end{aligned}$$

where the equality follows by (11) and the fact that $y_j, y_k \notin \{0, 1\}$.

Lemma 3 now follows from the following lemma proved in Section 4.2 which shows that (γ, ϵ, m)-useful polynomials of the type we require do in fact exist:

Lemma 4 *Let m be an integer and γ a sufficiently small positive real such that $\frac{m}{2\gamma}$ and $\frac{1}{2\gamma}$ are even integers and m is significantly larger than $\frac{1}{\gamma}$. Suppose $\epsilon > 5\gamma$. Then there exists a (γ, ϵ, m)-useful polynomial P such that for all $i, j, k \in \{-1, 1\}^m$ where $j, k \ne i$ and $jk \in E$,*

$$|P(\mathbf{u}_i \cdot \mathbf{u}_j) + P(\mathbf{u}_i \cdot \mathbf{u}_k)| \le 20\gamma. \quad (12)$$

4.2 Proof of Lemma 4: Constructing (γ, ϵ, m)-useful polynomials

In this section we prove Lemma 4. Fix ϵ and γ as in the statement of the lemma. Let R be the subset of $\mathbb{R}^2$ that consists of all $(x, y) \in [-1, 1]^2$ for which $|x + y| \le 2\gamma$, $|x - y| \le 2(1-\gamma)$, $x < 1 - \frac{1}{m}$, and $y < 1 - \frac{1}{m}$ (see Figure 1).

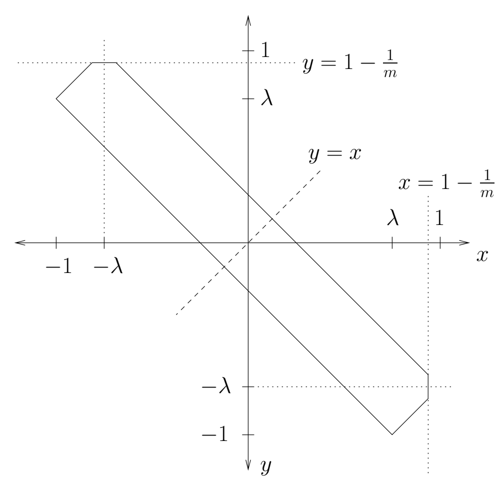

Figure 1. The domain R.

Claim 2 *To prove the lemma it suffices to find a polynomial P with nonnegative coefficients such that $P(1) = 1$, $\forall x \in [-1, 1]$ $P(x) \ge P(2\gamma - 1) = (2\epsilon - 1)/(2\epsilon + 1)$, and such that,*

$$|P(x) + P(y)| \le 20\gamma \quad \forall (x, y) \in R. \quad (13)$$

Proof: By definition, P satisfies the first three properties of a (γ, ϵ, m)-useful polynomial.

Next recall that the vectors $\mathbf{u}_i$ satisfy the property $-1 \leq \mathbf{u}_i \cdot \mathbf{u}_j \leq 1 - \frac{2}{m}$ for all $1 \leq i \neq j \leq 2^m$. Further, if $jk \in E$ and $i \neq j, k$, then since $\mathbf{u}_j + \mathbf{u}_k$ is supported on γm coordinates on which it assumes values $\pm 2/\sqrt{m}$ we get that

$$|\mathbf{u}_i \cdot \mathbf{u}_j + \mathbf{u}_i \cdot \mathbf{u}_k| = |\mathbf{u}_i \cdot (\mathbf{u}_j + \mathbf{u}_k)| \leq 2\gamma.$$

Similarly, $|\mathbf{u}_i \cdot \mathbf{u}_j - \mathbf{u}_i \cdot \mathbf{u}_k| \leq 2(1-\gamma)$. Hence, $\{(\mathbf{u}_i \cdot \mathbf{u}_j, \mathbf{u}_i \cdot \mathbf{u}_k) : j, k \neq i \text{ and } jk \in E\} \subseteq R$. So (13) implies (12). Moreover, since $5\gamma < \epsilon$, it implies Property 4 of a (γ, ϵ, m)-useful polynomial in all cases except when $i = k$. However, in that case we have

$$\begin{aligned} P(\mathbf{u}_i \cdot \mathbf{u}_i) + P(\mathbf{u}_i \cdot \mathbf{u}_j) &= P(1) + P(2\gamma - 1) \\ &= 1 + \frac{2\epsilon - 1}{2\epsilon + 1} = \frac{4\epsilon}{1 + 2\epsilon}, \end{aligned}$$

and hence Property 4 holds in that case too. □

Lemma 4 now follows from the following technical lemma:

Lemma 5 *Let m be an integer and γ a sufficiently small positive real such that $\frac{1}{\gamma}$ is an even integer and m is significantly larger than $\frac{1}{\gamma}$. Let $\epsilon > 3\gamma$ be sufficiently small. Then there exists a polynomial P satisfying the conditions in Claim 2.*

Proof: Let $P(x) = \Delta(x+1)x^{\frac{2m}{\gamma}} + cx^{\frac{1}{\gamma}} + (1-c-2\Delta)x$ where c, Δ are positive constants we will define below so that P satisfies the conditions of the lemma. Note that P has a "high" degree component (i.e., $\Delta(x+1)x^{\frac{2m}{\gamma}}$) which vanishes at -1, as well as a "medium" degree and a linear component (see Figure 2). Note that $P(1) = 1$.

A necessary condition for ensuring that $P(x) \geq P(2\gamma - 1)$ for $x \in [-1, 1]$ is that $P'(2\gamma - 1) = 0$. This condition together with the condition $P(2\gamma - 1) = (2\epsilon - 1)/(2\epsilon + 1)$ immediately determine the values of c and Δ. The following rough bounds will suffice for our analysis:

$$\begin{aligned} \frac{2\epsilon}{1 + 2\epsilon} - 5\gamma < \Delta < 3\epsilon, \\ 7\gamma < c < 8.5\gamma. \end{aligned}$$

Note that since $\epsilon > 3\gamma$ these bounds ensure that P has positive coefficients.

Next we verify that these bounds ensure that $P(x) \geq P(2\gamma - 1)$ for $x \in [-1, 1]$. Since $\frac{1}{\gamma}$ is even, $P''(x)$ is at least

$$\Delta\left(\frac{2m}{\gamma} + 1\right)\frac{2m}{\gamma}x^{\frac{2m}{\gamma} - 1} + \Delta\left(\frac{2m}{\gamma} - 1\right)\frac{2m}{\gamma}x^{\frac{2m}{\gamma} - 2}.$$

It is not hard to see then that $P''(x) \geq 0$ whenever $x \geq -1 + \frac{2\gamma}{2m+\gamma}$. So since $P'(2\gamma - 1) = 0$, it follows that $P(x) \geq P(2\gamma - 1)$ whenever $x \geq -1 + \frac{2\gamma}{2m+\gamma}$. It is more difficult to estimate P'' when $x < -1 + \frac{2\gamma}{2m+\gamma}$; instead, we will bound $P(x)$ directly for such x: our lower bounds for c and Δ and the fact that m is sufficiently large imply that for $x < -1 + \frac{2\gamma}{2m+\gamma}$,

$$\begin{aligned} P(x) &> c\left(1 - \frac{\gamma}{m}\right)^{\frac{1}{\gamma}} - (1 - c - 2\Delta) \\ &> -1 + 2\Delta + c + ce^{-\frac{1}{m}} \\ &> -\frac{1 - 2\epsilon}{1 + 2\epsilon} - 4\gamma + 0.9c \\ &> P(2\gamma - 1). \end{aligned}$$

Hence, $P(x) \geq P(2\gamma - 1)$ for every x in $[-1, 1]$.

It remains to prove that $|P(x) + P(y)| \leq 20\gamma$ on R. Firstly, since $m \gg 1/\gamma$, we (very generously) have that $(x+1)x^{\frac{2m}{\gamma}} < \frac{\gamma}{6\epsilon}$ when $x \in [-1, 1 - \frac{1}{m}]$. Secondly, $|x^{\frac{1}{\gamma}} + y^{\frac{1}{\gamma}}| \leq 2$ over R. Finally, by definition of R, we have that $|x + y| \leq 2\gamma$ for all $(x, y) \in R$. Hence, for all $(x, y) \in R$, the expression $|P(x) + P(y)|$ is bounded from above by

$$\begin{aligned} &\Delta\left|(x+1)x^{\frac{2m}{\gamma}} + (y+1)y^{\frac{2m}{\gamma}}\right| \\ + \quad &c\left|x^{\frac{1}{\gamma}} + y^{\frac{1}{\gamma}}\right| \\ + \quad &(1 - c - 2\Delta)|x + y|. \end{aligned}$$

These three terms are at most γ, 17γ and 2γ, respectively, implying that $|P(x) + P(y)| \leq 20\gamma$. □

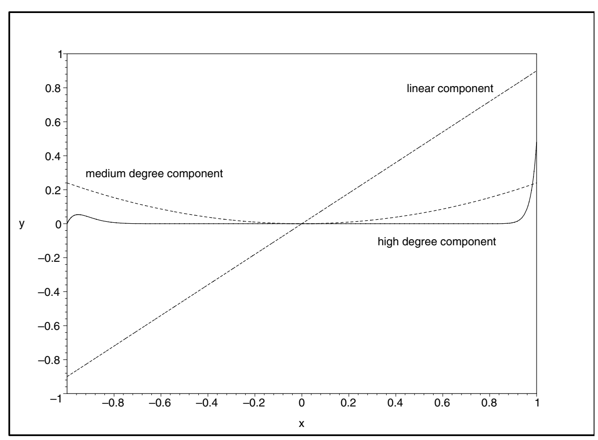

Figure 2. Relative behaviour of the three components of P.

5 Discussion

An obvious open problem is to investigate how the integrality gap for VERTEX COVER evolves beyond $\omega(\sqrt{\log n})$ rounds of LS_+. Note that our graph instances have girth essentially $O(\frac{1}{\gamma}) \approx \sqrt{\log n}$ and that proving integrality gaps for VERTEX COVER for more rounds than the girth proved quite challenging in the LS context (see [25, 23]).

It would be very interesting to expand the family of SDPs for which our methods apply. For example, the VERTEX COVER SDPs we study are incomparable to the SDP used in Karakostas's algorithm [16] and with the SDPs considered by Hatami et al. [14]. Karakostas's SDP employs the triangle and "extended" triangle inequalities (constraints (2) and (3), respectively, from Section 2.1), while Hatami et al. also consider the the so-called pentagonal inequalities. Such inequalities constrain the *geometry* of valid SDP solutions: they are constraints on the ℓ_2^2-distances of the vector solution and do not depend on the edges present in the underlying graph. It is not hard to show for the graph G_0 with no edges that there exist matrices in $M_+^r(VC(G_0))$ (for all r) whose Cholesky decompositions do not satisfy the triangle inequality when v_0 is the middle point. The technical reason for this is as follows: while r rounds of LS_+ suffice to derive all valid inequalities for any subset of r vertices, LS_+ (without strengthening the initial relaxation) cannot also derive all valid inequalities for the "lifted" variables Y_{ij} involving those r vertices. Intuitively, to derive such inequalities we need a lift-and-project method that in subsequent rounds does lifting on the vertex variables *and* the Y_{ij} variables (i.e., applies N_+ to $M_+(VC(G))$ rather than to $N_+(VC(G))$).

Sherali and Adams [24] describe precisely such a lift-and-project method. Unfortunately, our arguments do not seem to extend to this system. Indeed, no nontrivial integrality gaps are known for the SDP version of Sherali-Adams for *any* problem. Even for the LP version of Sherali-Adams only one such result is known: Fernandez de la Vega and Kenyon-Mathieu [9] prove a 0.5-integrality gap for MAX CUT after super-constant Sherali-Adams rounds.

Triangle, pentagonal and other such geometric inequalities for the Y_{ij} variables can be derived within LS_+ if one introduces new variables (and constraints) to the initial relaxation to represent the ℓ_2^2 distances of the Cholesky vectors corresponding to Y. Since geometric constraints have proved powerful in tightening relaxations for problems such as SPARSEST CUT [4], we feel that the most interesting open problem posed by our work is to extend our results to either the Sherali-Adams system or to LS_+ relaxations augmented with distance variables and constraints.

Partial progress along this line is made in [11] where it is shown that the construction from the current paper (modulo an affine transformation) satisfies VERTEX COVER SDPs tightened by local hypermetric inequalities (hypermetric inequalities are a canonical subfamily of inequalities satisfied by all ℓ_1 metrics and include triangle, pentagonal and indeed all $(2k+1)$-gonal inequalities). More precisely, the SDP solution analyzed in [11] arises by taking the Cholesky decomposition of the first-round protection matrix from the current paper and then applying the affine transformation $\mathbf{z}_i = 2\mathbf{v}_0 - \mathbf{v}_i$ (this simply maps $\{0, 1\}$ integral solutions to $\{1, -1\}$). It is shown in [11] that the ℓ_2^2 metric induced by these vectors satisfies all k-gonal inequalities for $k = \Omega(\sqrt{\log n / \log \log n})$ (and actually satisfies all hypermetric inequalities on k points). Moreover, the vectors satisfy the extended triangle inequalities (3) employed by Karakostas's SDP. Interestingly, the asymptotic bound for the parameter k in [11] is the same as the number of LS_+ rounds for which we prove our lower bound in the current paper. This hints at a deeper relationship between the families of SDPs considered in [11] and the current paper.

References

[1] M. Alekhnovich, S. Arora, 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.

[2] N. Alon and N. Kahale. Approximating the independence number via the theta-function. *Math. Program*, 80:253–264, 1998.

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

[4] 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.

[5] J. Buresh-Oppenheim, N. Galesi, S. Hoory, A. Magen, and T. Pitassi. Rank bounds and integrality gaps for cutting planes procedures. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, pages 318–327, 2003.

[6] M. Charikar. On semidefinite programming relaxations for graph coloring and vertex cover. In *SODA '02: Proceedings of the thirteenth annual ACM-SIAM sympo-*

sium on Discrete algorithms, pages 616–620, Philadelphia, PA, USA, 2002. Society for Industrial and Applied Mathematics.

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

[8] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996.

[9] W. Fernandez de la Vega and C. Kenyon-Mathieu. Linear programming relaxations of Maxcut. In *SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 2007.

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

[11] K. Georgiou, A. Magen, T. Pitassi, and I. Tourlakis. Vertex cover resists SDPs tightened by local hypermetric inequalities. Manuscript, 2007.

[12] 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.

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

[14] 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 *APPROX*, 2007.

[15] J. Kahn and G. Kalai. A counterexample to Borsuk's conjecture. *Bulletin of the American Mathematical Society*, 29(1):60–62, 1993.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] L. Lovász. Semidefinite programs and combinatorial optimization. In B. A. Reed and C. L. Sales, editors, *Recent Advances in Algorithms and Combinatorics*, volume 11 of *CMS Books in Mathematics*, pages 137–194. Springer-Verlag, Berlin-Heidelberg-New York-Hong Kong-London-Milan-Paris-Tokyo, 2003.

[21] 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.

[22] G. Schoenebeck, L. Trevisan, and M. Tulsiani. A linear round lower bound for Lovász-Schrijver SDP relaxations of vertex cover. In *Proceedings of the 22nd IEEE Conference on Computational Complexity*, pages 205–216. IEEE, 2007.

[23] G. Schoenebeck, L. Trevisan, and M. Tulsiani. Tight integrality gaps for Lovász-Schrijver LP relaxations of vertex cover and max cut. In *Proceedings of the 39th annual ACM symposium on Theory of computing*, pages 302–310. ACM, 2007.

[24] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3):411–430, Aug. 1990.

[25] I. Tourlakis. New lower bounds for vertex cover in the Lovász-Schrijver hierarchy. In *Proceedings of the 21st IEEE Conference on Computational Complexity*, 2006.

A Proof sketch of Lemma 1

In [10] we find the following similar-looking statement to Lemma 1 about sets avoiding intersections.

Lemma 6 *(Corollary 4.2 in [10]) Let η be a sufficiently small number and m an integer. Also, let $\mathcal{F}$ and $\mathcal{G}$ be two set families over the universe $[m]$ so that $|F \cap G| \neq \lfloor m\eta \rfloor$ for every $F \in \mathcal{F}$, $G \in \mathcal{G}$. Then $4^{-m}|\mathcal{F}||\mathcal{G}| \leq (1-\eta^2/4)$.*

By taking $\mathcal{F} = \mathcal{G}$ and treating set families as points in $\{-1,1\}^m$ we get that the above lemma says that a subset of size $> 2^m(1-\eta^2/4)$ must contain two points which share exactly $\lfloor m\eta \rfloor$ ones.

Let S be a set in $\{-1,1\}^m$ avoiding distance $(1-\gamma)m$. Instead of bounding the size of S we will bound the size of the biggest set of the form $S_k = \{s \in S : |s| = k\}$, where $|\cdot|$ denotes Hamming weight (i.e., the number of coordinates set to 1). Assume S_w is this largest set; clearly it is of size at least $|S|/m$. We may and will assume that $w \leq m/2$. Having reduced to the case where all points have the same Hamming weight w we relate to Lemma 6: it is easy to see that no two points in S_w may share exactly $w - m(1-\gamma)/2$ ones.

Now, let us assume first that $w > \frac{m}{2}(1-\gamma/2)$. Then S_w is a subset that avoids intersections of size ηm where $\gamma/4 \leq \eta \leq \gamma/2$. We now apply Lemma 6 (or its corollary rather) to get that

$$|S_w| \leq 2^m(1-\eta^2/4) \geq 2^m(1-\gamma^2/64)^m,$$

and so $|S| \leq m|S_w| \leq m2^m(1-\gamma^2/64)^m$. For the other case, namely $w \leq \frac{m}{2}(1-\gamma/2)$, it is enough to use the simple upper bound $S_w \leq \binom{m}{w}$. More precisely $|S_w|$

is at most

$$\binom{m}{\frac{m}{2}(1-\gamma/2)} \sim 2^{mH(1/2-\gamma/4)} \sim 2^m(2^{-\gamma^2/16})^m$$
$$\le 2^m \exp\left(-\frac{\log 2}{16}\gamma^2\right)$$
$$\le 2^m(1-\gamma^2/64)^m,$$

and again S is at most m times this bound.

The above estimate is nearly tight: consider the (open) Hamming ball B of radius $(1-\gamma)/2$; clearly this ball is an independent set in $G_{\gamma,m}$. Now $|B| = \sum_{j<\frac{m}{2}(1-\gamma)} \binom{m}{j}$ which is at least $\frac{\gamma m}{2}\binom{m}{\frac{m}{2}(1-2\gamma)}$. The last expression can be further bounded from below by

$$\frac{\gamma m}{2} 2^{mH(1/2-\gamma)} \sim \frac{\gamma m}{2} 2^{m(1-\gamma^2/4)} = 2^m \frac{\gamma m}{2} 2^{-\gamma^2 m/4}.$$

So for $|B|$ to be $o(2^m)$ we must have that $\gamma m 2^{-\gamma^2 m/4} = o(1)$ and so $\gamma = \Omega(\sqrt{\log m/m})$.

B Proof of Lemma 2

For completeness, we include in this section a proof of the lemma by Schoenebeck, Trevisan and Tulsiani [23] (Lemma 2 here) for expressing an ϵ-saturated vector as a convex combination of ϵ-vectors.

Proof: Partition V as follows: Let $V_- = \{i \in V : x_i < 1/2+\epsilon\}, V_+ = \{i \in V : x_i > 1/2+\epsilon\}, V_0 = \{i \in V : x_i = 1/2+\epsilon\}$. Let $r(0) = 0$, and for all $i \in V$ let

$$r(i) = \begin{cases} 1 - \frac{x_i}{1/2+\epsilon}, & i \in V_- \\ 1, & i \in V_0 \\ 1 - \frac{1-x_i}{1/2-\epsilon}, & i \in V_+ \end{cases}$$

setting at the end the maximum of the $r(i)$'s equal to 1. Note that since $\mathbf{x}$ is ϵ-saturated, whenever $ij \in E$ and $i \in V_-$, we must have $j \in V_+$. Moreover, for such a pair we must have that $r(j) \ge r(i)$ because

$$\begin{aligned} r(j) - r(i) &= 1 - \frac{1-x_j}{1/2-\epsilon} - \left(1 - \frac{x_i}{1/2+\epsilon}\right) \\ &= \frac{x_i}{1/2+\epsilon} - \frac{1-x_j}{1/2-\epsilon} \\ &= \frac{x_i(1/2-\epsilon) - (1-x_j)(1/2+\epsilon)}{(1/2+\epsilon)(1/2-\epsilon)} \\ &= \frac{x_i + x_j - (1+2\epsilon)}{2(1/4-\epsilon^2)} + \frac{\epsilon(x_j - x_i)}{1/4-\epsilon^2} > 0, \end{aligned}$$

where the last inequality follows from the fact that $\mathbf{x}$ is ϵ-saturated.

Reorder the $r(i)$'s so that $0 = r(i_0) \le r(i_1) \le \ldots \le r(i_{|V|})$. For each $t = 1, \ldots, |V|$, let $\mathbf{x}^{(t)}$ be the ϵ-vector where

$$x_i^{(t)} = \begin{cases} 0, & i \in V_- \text{ and } r(i) \ge r(i_t) \\ 1, & i \in V_+ \text{ and } r(i) \ge r(i_t) \\ \frac{1}{2}+\epsilon, & \text{otherwise} \end{cases}$$

We claim these vectors are in $VC(G)$. To see why consider an edge ij. The constraint $x_i^{(t)} + x_j^{(t)} \ge 1$ is satisfied unless at least one of $x_i^{(t)}$ and $x_j^{(t)}$ is 0. However, if $x_i^{(t)} = 0$, then $i \in V_-$ and $r(i) \ge r(i_t)$. So the feasibility of $\mathbf{x}$ implies $j \in V_+$ and hence $r(j) \ge r(i_t)$. So $x_j^{(t)} = 1$ and the constraint is satisfied.

It remains to argue that $\mathbf{x}$ is in the convex hull of the $\mathbf{x}^{(t)}$'s. To that end, we define a distribution $\mathcal{D}$ over the vectors $\mathbf{x}^{(t)}$ such that $\mathbf{x}^{(t)}$ is assigned the probability $r(i_t) - r(i_{t-1})$. It is easy to verify now that $\mathrm{E}_t[x_j^{(t)}] = x_j$ for all $j \in V$. $\Box$