Tight Gaps for Vertex Cover in the Sherali-Adams SDP Hierarchy*

Siavosh Benabbas[1], Siu On Chan[2], Konstantinos Georgiou[3], and Avner Magen[1]

1 **Department of Computer Science, University of Toronto**
siavosh@cs.toronto.edu
2 **Department of Computer Science, University of California, Berkeley**
siuon@cs.berkeley.edu
3 **Department of Combinatorics and Optimization, University of Waterloo**
k2georgiou@math.uwaterloo.ca

Abstract

We give the first tight integrality gap for Vertex Cover in the Sherali-Adams SDP system. More precisely, we show that for every $\epsilon > 0$, the standard SDP for Vertex Cover that is strengthened with the level-6 Sherali-Adams system has integrality gap $2-\epsilon$. To the best of our knowledge this is the first nontrivial tight integrality gap for the Sherali-Adams SDP hierarchy for a combinatorial problem with hard constraints.

For our proof we introduce a new tool to establish Local-Global Discrepancy which uses simple facts from high-dimensional geometry. This allows us to give Sherali-Adams solutions with objective value $n(1/2 + o(1))$ for graphs with small $(2 + o(1))$ vector chromatic number. Since such graphs with no linear size independent sets exist, this immediately gives a tight integrality gap for the Sherali-Adams system for superconstant number of tightenings. In order to obtain a Sherali-Adams solution that also satisfies semidefinite conditions, we reduce semidefiniteness to a condition on the Taylor expansion of a reasonably simple function that we are able to establish up to constant-level SDP tightenings. We conjecture that this condition holds even for superconstant levels which would imply that in fact our solution is valid for superconstant level Sherali-Adams SDPs.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Vertex Cover, Integrality Gap, Lift-and-Project systems, Linear Programming, Semidefinite Programming

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2011.41

1 Introduction

A vertex cover of a graph $G = (V, E)$ is a subset S of the vertices such that for every edge $ij \in E$ at least one vertex among i, j lies in S. In the Minimum Vertex Cover problem the objective is to find the vertex cover of minimum size. While a 2-approximation algorithm is rather straightforward, considerable effort has failed to yield any polynomial time algorithm with approximation ratio $2-\Omega(1)$. Indeed the best algorithm known achieves an approximation ratio of $2-O(\sqrt{1/\log n})$ [21]. On the other hand, the strongest PCP-based hardness result [12] shows that 1.36-approximating Vertex Cover is **NP**-hard. Only by

* The full version of this work is available as [5]

31st Int'l Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2011).
Editors: Supratik Chakraborty, Amit Kumar; pp. 41–54
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

assuming Khot's Unique Game Conjecture [24], whose validity is the subject of an active area of research (see [1, 25] for example), one can show a $2 - o(1)$ hardness.

Motivation for studying Vertex Cover is two-fold. For one thing it is arguably one of the simplest **NP**-hard problems whose inapproximability remains unresolved. But more importantly, studying Vertex Cover has introduced some very important techniques both in terms of approximation algorithms and hardness of approximation with [12] being a prime example. Intuitively this is, at least partly, due to the "hard constraints" of Vertex Cover, that is the solution has to satisfy a number of inflexible constraints (the edge constraints). As many of the standard techniques for proving hardness of approximation and integrality gaps produce solutions which satisfy *most* constraints in an instance, showing tight hardness for Vertex Cover has remained unresolved.

Trying to resolve the approximability of Vertex Cover, one could study the behavior of prominent algorithmic schemes, such as Linear Programming (LP) and Semidefinite Programming (SDP) relaxations, which have yielded state-of-the-art algorithms for many combinatorial optimization problems. There, the measure of efficiency is the *Integrality Gap* which sets the approximation limitation of the algorithms based on these relaxations. In this work we show that a large family of LP and SDP relaxations for Vertex Cover have integrality gap arbitrarily close to 2. Such an integrality gap rules out a rich and important family of approximation algorithms for the problem at hand.

Furthermore, there seems to be a connection between integrality gaps for strong LP/SDP relaxations of a problem and its hardness of approximation. In one direction the reductions used to establish hardness of approximation for many problems have been used to construct integrality gaps for them, e.g. [26, 9, 29, 33]. In the other direction, and specifically for Vertex Cover, Vishwanathan [34] shows that any hard instance of the problem should have subgraphs that look like the so called "Borsuk graphs". Interestingly a specific subfamily of Borsuk graphs were previously used in many integrality gap instances for Vertex Cover, e.g. [17, 8, 20, 15, 16]. To make the picture even more complete, we show that *any* Borsuk graph is a good integrality gap instance for the (so called) Sherali-Adams LP system of relaxations for Vertex Cover.

The (tight) integrality gap of the standard LP and SDP relaxations for Vertex Cover has long been resolved [17]. Nevertheless, celebrated relaxations for a number of combinatorial problems require strengthenings (addition of extra constraints) aiming to drop the integrality gap. In that direction, a number of systematic procedures, known as Lift-and-Project systems have been proposed to systematically improve the integrality gap. These systems build strong hierarchies of either LP relaxations (as the Lovász-Schrijver and the Sherali-Adams systems) or SDP relaxations (as the Lovász-Schrijver SDP, the Sherali-Adams SDP and the Lasserre systems). Lift-and-Project systems can be thought of as being applied in rounds (also called levels). The bigger the number of rounds used, the more accurate the obtained relaxation is. In fact, if as many rounds as the number of variables are used, the final relaxation is exact and no integrality gap exists. On the other hand the size of the derived relaxation grows exponentially with the number of rounds, which implies that the time one needs to solve it also grows. It is then natural to ask whether looking at a modest number of rounds (say $O(1)$ or $\log \log n$) will result in an algorithm with approximation factor better than 2.

Identifying the limitations of relaxations derived by Lift-and-Project system has attracted much attention and showing integrality gaps for the Sherali-Adams SDP and the Lasserre systems stand as the most attractive subjects in this area of research due to a number of reasons. Firstly, the best algorithms known for many combinatorial optimization problems (and Vertex Cover in particular) are based on relaxations weaker than those derived by a constant (say four) rounds of the Sherali-Adams SDP system which we study here, e.g. [18, 23, 2, 21]. Lift-and-Project hierarchies have been also used recently in designing approximation algorithms with a runtime-approximation ratio trade off, e.g. [11, 27, 10, 4, 22, 3, 19]. Finally, for some particular constraint satisfaction problems, and modulo the Unique Games Conjecture, no approximation algorithm can perform better than the one obtained by Sherali-Adams SDP of a constant number of rounds (see [28].) One can then think of algorithms based on the Sherali-Adams SDP as an interesting model of computation.

In this work we study the limitations of strong relaxations for Vertex Cover in the powerful Sherali-Adams SDP system. The performance of the same hierarchy has been studied for other combinatorial problems (see [29, 6, 7]), but its integrality gap for Vertex Cover remained open, due to the hard constraints mentioned earlier. Our main result is as follows.

▶ **Theorem 1.1.** *For every $\epsilon > 0$, the SDP derived by the level-6 Sherali-Adams SDP system for* Vertex Cover *has integrality gap $2 - \epsilon$.*

Theorem 1.1 yields the first nontrivial Sherali-Adams SDP integrality gap for Vertex Cover and in fact any problem with hard constraints. While tight integrality gaps for weaker or incomparable systems were known, there were no good candidates for Sherali-Adams SDP integrality gap solutions.In particular, while integrality gaps for the closely related but weaker Sherali-Adams LP system for Vertex Cover were known [9], the solution there does *not* satisfy the required positive semidefiniteness condition. As we explain below, apart from the significance of our new SDP integrality gap, we also believe that our proofs are interesting in their own right. In Section 3 we give a high level description of our ideas, along with a detailed explanation of how our techniques are different from existing integrality gap results.

On our way to prove the above theorem we need to define new solutions for Sherali-Adams LP relaxations of Vertex Cover. As mentioned, one of our contributions is an intuitive and geometric explanation of why this large family of LPs are fooled by a certain family of graphs, the so-called Borsuk graphs. This yields a tight level-$\Omega(\sqrt{\log n / \log\log n})$ integrality gap for Sherali-Adams LP (see Theorem 4.4.) Other than being used in our proof of Theorem 1.1, our solution is arguably simpler and more intuitive than the integrality gap of [9] for the same system.[1]

The heart of the problem in showing integrality gaps for Sherali-Adams SDPs is that the proposed solution needs to satisfy a strong positive-semidefiniteness condition. Toward establishing Theorem 1.1, we show how to reduce this condition into a clean analytic statement about a certain function parameterized by t. We are able to show that this analytic statement holds up to $t = 6$, hence the level-6 Sherali-Adams SDP gap. We have strong evidence (both theoretical and experimental) that the aforementioned analytic statement holds for any constant value of t, which we explicitly state as a conjecture in Section 5.3. To sum up, we have the following second theorem.

▶ **Theorem 1.2.** *Assuming Conjecture 5.12, for every constant $\epsilon > 0$ and $t \in \mathbb{N}$, the SDP derived by the level-t Sherali-Adams SDP system for* Vertex Cover *has integrality gap $2 - \epsilon$.*

For a brief discussion of the validity of Conjecture 5.12 see Remark 5.3.

[1] Although it should be mentioned that their integrality gap applies to more rounds.

Known integrality gaps for Vertex Cover: Considerable effort has been invested in strong lower bounds for various hierarchies for Vertex Cover. For LP hierarchies, [31] shows an integrality gap of $2-\epsilon$ for $\Omega(n)$ rounds of the Lovász-Schrijver system and [9] shows the same integrality gap for the stronger Sherali-Adams system up to $\Omega(n^\delta)$ rounds (with δ going to 0 together with ϵ.) Both results concern LP hierarchies, which are incomparable to SDP relaxations. For SDP hierarchies, and for the Lovász-Schrijver SDP system which is stronger than both the LS system and the canonical SDP formulation (but incomparable to Sherali-Adams), [14] shows an integrality gap of $2-\epsilon$ for $\Omega(\sqrt{\log n/\log\log n})$ levels.

The integrality gap of two stronger hierarchies for Vertex Cover, on the other hand, has long been open. The first is the Sherali-Adams SDP system which is stronger than the LS system, and the subject of this paper. The second is the Lasserre system, for which no tight integrality gap for Vertex Cover is known.[2] If one is content with an integrality gap less than 2, a 1.36 integrality gap for $\Omega(n^\delta)$ levels [33] and a 7/6 integrality gap for $\Omega(n)$ levels [30] of the Lasserre system are known. We will compare our proof techniques with previous ones at the end of Section 3.

2 Preliminaries

2.1 Borsuk Graphs, Frankl-Rödl Graphs and Tensoring

Our integrality gap instances are *Frankl-Rödl graphs*. These graphs are parameterized by an integer m which is considered growing and a real parameter $0<\gamma<1$.

▶ **Definition 2.1.** (*Frankl-Rödl graphs*) The *Frankl-Rödl graph* G_γ^m is the graph with vertices $\{-1,1\}^m$ where two vertices $i,j\in\{-1,1\}^m$ are adjacent iff $d_H(i,j)=(1-\gamma)m$.

Frankl-Rödl graphs exhibit an interesting "extremal" combinatorial property. While G_0^m is a perfect matching and thus has a vertex cover of size half the number of its vertices, a beautiful theorem by Frankl and Rödl states that for slightly larger γ, any vertex cover of G_γ^m is very large. The fact that such a small geometric perturbation results in a drastic change in the vertex cover size has led to the use of Frankl-Rödl graphs as tight integrality gap instances in a series of results [17, 8, 14, 15, 16].

▶ **Theorem 2.2** ([14]; slight modification of Theorem 1.4 of [13]). *Let m be an integer and let $\gamma=\Theta(\sqrt{\log m/m})$ be a sufficiently small number so that γm is an even integer. Then any vertex cover of G_γ^m contains at least a $1-o(1)$ fraction of the vertices.*

An important tool in proving strong integrality gaps is tensoring of vectors. Recall that for $\mathbf{u}\in\mathbb{R}^n$ and $\mathbf{v}\in\mathbb{R}^m$ their *tensor product* $\mathbf{u}\otimes\mathbf{v}\in\mathbb{R}^{nm}$ is a vector indexed by ordered pairs from $[n]\times[m]$ taking value u_iv_j at coordinate (i,j). For any polynomial $P(x)=c_1x^{t_1}+\ldots+c_qx^{t_q}$ *with nonnegative coefficients* consider the function T_P mapping a vector $\mathbf{u}\in\mathbb{R}^n$ to the vector $T_P(\mathbf{u})=(\sqrt{c_1}\mathbf{u}^{\otimes t_1},\ldots,\sqrt{c_q}\mathbf{u}^{\otimes t_q})\in\mathbb{R}^{\sum_i n^{t_i}}$, where $\mathbf{u}^{\otimes d}$ is the vector obtained by tensoring $\mathbf{u}$ with itself d times. Polynomial tensoring can be used to manipulate inner products in the sense that $T_P(\mathbf{u})\cdot T_P(\mathbf{v})=P(\mathbf{u}\cdot\mathbf{v})$; it was used as an ingredient in many integrality gap results such as [17, 8, 14, 15].

We often think of the vertices of the Frankl-Rödl graphs as (scaled and) embedded on the unit sphere S^{m-1}. In this sense the Frankl-Rödl graphs are subgraphs of the infinite *Borsuk graphs*.

[2] In fact there are only a few combinatorial problems for which tight Lasserre integrality gaps are known. (see [30] and [33] for some notable exceptions.)

▶ **Definition 2.3.** (*Borsuk graphs*) The Borsuk graph B^m_δ is an infinite graph with vertex set S^{m-1}. Two vertices $\mathbf{x}, \mathbf{y}$ are adjacent if they are nearly antipodal, i.e. $\|\mathbf{x}+\mathbf{y}\| \le 2\sqrt{\delta}$.

2.2 Strong relaxations for Vertex Cover

In this subsection we give a brief high level description of the Sherali-Adams SDP system applied to the Vertex Cover problem. This high level description should be enough to understand the high level of our results. The interested reader can find a rigorous definition in the full version of the paper or [32].

The starting point of the Sherali-Adams SDP for Vertex Cover is the following simple LP relaxation of Vertex Cover. Assume that $G = (V, E)$ is the input graph.

$$\min \sum_{i\in V} x_i, \qquad \text{s.t. } \forall ij \in E \;\; x_i + x_j \ge 1, \qquad \forall i \in V \;\; x_i \in [0,1] \tag{1}$$

Here x_i is the indicator variable of vertex i being part of a vertex cover. Since in the LP relaxation (1) x_i assumes any value in $[0, 1]$, we may think of x_i as encoding a local distribution $\mathcal{D}(\{i\})$ of 0-1 assignments for the elements in $\{i\}$. The Sherali-Adams LP system strengthens this relaxation by introducing variables to encode the joint status of a subset of vertices U with respect to the vertex cover, for all subsets up to a certain size. In particular, the Sherali-Adams LP system of level t, seen below, is a Linear Program with the following variables. If $U \subset V$ is any subset of the vertices of size at most t, the program will have real-valued variables to specify a distribution $\mathcal{D}(U)$ over the subsets of U. Furthermore, the program will have two kinds of constraints. The first kind (similar to the one in (1)) ensure that any subset of U that is assigned a positive probability covers all the edges inside U, i.e. the distribution $\mathcal{D}(U)$ is over vertex covers of U. The second kind of constraints ensure that the marginals of the distributions for $U_1 \subseteq U$ are consistent on U_1, i.e. any event that only depends on the vertices of U_1 has the same probability according to $\mathcal{D}(U_1)$ and $\mathcal{D}(U)$. The objective value of the program is the sum over all vertices v, of the probability that v is in the local vertex covers (which is well defined as $\mathcal{D}(U)$'s are consistent for all $U \ni v$.) That is, fix a $U \ni v$, the contribution of v to the objective function is, $\mathbb{P}_{S\sim\mathcal{D}(U)}[v \in S]$. Summarizing we have the following relaxations,

▶ **Definition 2.4** (Level-t Sherali-Adams LP relaxation of Vertex Cover)**.** Let $\mathcal{P}(U)$ denote the powerset of U.

$$\begin{array}{lll} \min & \sum_{i\in V} \mathbb{P}_{S\sim\mathcal{D}(\{i\})}[i \in S] & \\ \text{s.t.} & \mathbb{P}_{S\sim\mathcal{D}(\{i,j\})}[i \notin S, j \notin S] = 0 & \forall ij \in E \qquad \textit{(Edge constraints)} \\ & \mathbb{P}_{S\sim\mathcal{D}(U_1)}[S = T] = \mathbb{P}_{S\sim\mathcal{D}(U)}[S \cap U_1 = T] & \forall T \subseteq U_1 \subseteq U \subseteq V, |U| \le t \\ & \mathcal{D}(U) \text{ is a distribution on } \mathcal{P}(U) & \forall U \subseteq V, |U| \le t \end{array} \tag{2}$$

▶ **Definition 2.5** (Level-t Sherali-Adams SDP relaxation of Vertex Cover)**.** The Sherali-Adams SDP relaxation is the Sherali-Adams LP relaxation plus the following semi-definiteness constraint. Define M_1 to be an $(n+1)\times(n+1)$ matrix whose rows and columns are indexed by $\emptyset, \{1\}, \ldots, \{n\}$ as follows and add the following semi-definiteness condition.

$$m_{I,J} = \underset{S\sim\mathcal{D}(I\cup J)}{\mathbb{P}}[I \cup J \subseteq S] \qquad M_1 = [m_{I,J}]_{(n+1)\times(n+1)} \succeq 0. \tag{3}$$

In other words, one makes a matrix whose first row and column and diagonal are the “singleton probabilities”, i.e., the probabilities of each vertex being in a set sampled according to the local distribution, while the rest of the matrix is filled with the "doubleton probabilities", i.e., the probabilities that pairs of vertices are in a set sampled according to the local distribution together.

It is not hard to see that any *integral* solution of (1) gives rise to a solution to the Sherali-Adams SDP relaxation of any level. It is also not hard to see that the optimum of the Sherali-Adams SDP relaxation can be found in time polynomial in n^t. While the above is not the original definition of Sherali-Adams hierarchy it is equivalent. The reader can see the original definition as well as the formal theorem stating the equivalence in the full version of the paper.

3 Outline of Our Method and Comparison to Previous Work

By Theorem 2.2, for $\gamma = \sqrt{\log m/m}$, G^m_γ has no vertex cover smaller than $2^m(1 - o(1))$. A tight integrality gap therefore calls for a solution in the system of objective value at most $2^m(1/2 + \epsilon)$, for a small constant $\epsilon > 0$.

Consider the following experiment used to define our solution. A geometric way to obtain a distribution of vertex covers would be to embed G^m_γ on the unit sphere and take a sufficiently large *spherical cap* centered at a random point on the sphere. Of course, given the Frankl-Rödl theorem mentioned above, in doing so we have not achieved much since we are defining a *global* distribution of vertex covers, and thus its expected size has to be at least $2^m(1 - o(1))$. However, it is useful to understand why these vertex covers are big from a geometric point of view: the heightof the spherical cap must be at least $1 + \sqrt{\gamma}$ (as opposed to 1 for a half-sphere.) Now concentration of measure on the sphere implies that because $\sqrt{\gamma m} = \omega(1)$ the area of such a cap is a $1 - o(1)$ fraction of the whole sphere. So the probability that any vertex of the graph is in the cap is $1 - o(1)$, which is very large. Had it been the case that $\sqrt{\gamma m} = o(1)$ concentration of measure would imply that the area of the cap is $1/2 + o(1)$ of that of the sphere and we would have had a small vertex cover.

The main idea is that one only needs to define probabilities for small sets (up to size t if the goal is to show integrality gaps for level-t Sherali-Adams LP relaxations.) So one can first embed the points in such a small set in a *small dimensional sphere* and then repeat the above experiment to define a random vertex cover. The spherical caps that are required in order to cover the edges in these sets have the same height, but now, due to the lower dimension, their area is greatly reduced! Specifically, if the original set has at most t points, the experiment can be performed in a t-dimensional sphere and if $\sqrt{\gamma t} = o(1)$, the probability of any vertex participating in the vertex cover will be no more than $1/2 + o(1)$. In particular, $t = o(\sqrt{m/\log m})$ would suffice.

It is critical, of course, that the obtained distributions are consistent. But this is "built-in" in this experiment. Indeed, due to spherical symmetry, the probability that a set of points on a t dimensional sphere belong to a random cap of a fixed radius depends only on t, the radius of the cap and the pairwise Euclidean distances of the points in the set. Interestingly this construction works for any graph with vector chromatic number $2 + o(1)$. In other words, if G is an n vertex graph that can be embedded into the unit sphere so that the end points of any edge are almost antipodes, then there is a sufficiently "low-level" (but non-trivial) Sherali-Adams solution of value $(1/2 + o(1))n$.

Unfortunately, we cannot show that the above solution satisfies the extra constraints imposed by the SA SDP system. Instead we change our solution in several ways to attain positive semidefiniteness. These changes are somewhat technical and we avoid discussing them in detail here. At a high level the changes are (i) we add a small probability of picking the *whole* graph as the vertex cover. (ii) We apply a transformation of the canonical embedding of the cube in the sphere that ensures that the farthest pairs of vertices are precisely the edges, and also that the inner products have a bias to being positive (as opposed to the canonical embedding in which the average inner product is 0.)

To get some insight into the rationale of these modifications, first note that the matrix whose positive definiteness we need to prove happens to be highly symmetric. For such symmetric matrices a necessary condition for positive semi-definiteness is that the average entry is at least as large as the square of the diagonal entries. Manipulation (i) above is precisely the tool we need to ensure this condition, and has no adverse effect otherwise. The second transformation is useful although not clearly necessary. We can, however, argue that without a transformation of this nature, a good SDP solution is possible also for a graphs in which edges connect vertices that are at least as far as $m(1-\gamma)$ (rather than exactly that distance). The existence of solutions for such dense graphs seems intuitively questionable. Last, boosting the typical inner product can be shown to considerably boost the Taylor coefficients of a certain function which we need to show only has positive Taylor coefficients. The later is a condition to which we reduce the positive-semidefiniteness of our LP solution.

Comparison to Previous Work: There are more than half a dozen different integrality gap constructions for Vertex Cover in different Lift-and-Project systems known. Among these the most relevant to our work is [9]. In [9], Charikar, et al. obtain a Sherali-Adams solution that is based on embedding the vertices of the graph in the sphere. The similarity with our work is that Charikar et al. take a special case of caps, i.e. half-spheres, in order to determine probabilities. Consistency of these distribution is, just as in our case, guaranteed by the fact that these probabilities are intrinsic to the local distances of the point-set in question. However, the reason that these distributions behave differently than a global distribution (which is essential for an integrality gap construction) is completely different than ours. It is easy to see that when the caps in the construction are half spheres, the dimension does not play a role at all. However, in [9] there is no *global embedding* of the points in the sphere but rather only a local one. In contrast, our distributions can be defined for all dimensions, however as we mentioned we must keep the dimension reasonably small in order to guarantee small objective value. Another big difference pertains to the different instances. While our construction may very well be the one (or close to the one) that will give a Lasserre integrality-gap bound, the instances of [9] have no substantial integrality gap even for the standard SDP. Thus their result cannot be extended to the stronger Lasserre or Sherali-Adams SDP hierarchies.

It is also important to put our work in context with the sequence of results dealing with SDP integrality gaps of Vertex-Cover [17, 8, 14, 15, 16]. In these works the solution can be thought of as an approximation to a very simple set: a dimension cut, that is a face of the cube. This set is not a vertex cover, but in some geometric sense is close to one. The SDP solutions are essentially averaging of such dimension-cuts with some carefully crafted perturbations. Using the same language, the solution we present is based on Hamming balls of radius $m/2$ (i.e. translations of the majority function) rather than dimension-cuts (i.e. dictatorship functions). The perturbation we apply to make such a solution valid is simply the small increase in the radius of the Hamming balls. Another distinction is that while all previous results use tensoring to *construct* their solutions we mainly use it to *certify* its positive semidefiniteness. In other words, our solutions are defined geometrically and then tensoring is used to give an alternative view which helps to show they have the required positive semidefiniteness.

4 Fooling LPs derived by the Sherali-Adams System

4.1 Local Distributions of Vertex Covers for Borsuk Graphs

In this section we study relaxation (2) for discrete subgraphs of B_γ^m on n vertices. In particular, for every set $U \subseteq [n]$ we define a distribution of vertex covers that are locally consistent.

The family of distributions we are looking for arises from the following experiments. Fix a discrete subgraph $G = (V, E)$ of B_δ^m on n vertices for which we want to construct a level t Sherali-Adams LP solution with small objective value. Given that $G = (V, E)$ is a subgraph of B_δ^m we can think of its vertices as points on S^{m-1} and in particular talk about their Euclidean distances. The following experiment defines the local distributions.

Experiment Local-Global

The input is any $I \subseteq V$, of size at most t, and some $\sqrt{\delta} > 0$.
The result of the experiment is a distribution $\mathcal{D}(I)$ of 0/1 assignments on I.

(a) Embed the I-induced subgraph of G into S^{t-1} preserving all pairwise Euclidean distances.

(b) In S^{t-1} consider the complement C of a random spherical cap of height $1 - \sqrt{\delta}$.

(c) Vertices of I are assigned 1 if they are in the cap C, otherwise they are assigned 0.

Notice that step (a) is possible because $|I| \leq t$.

▶ **Lemma 4.1.** *For every finite subgraph of B_δ^m on n vertices, the family of distributions $\mathcal{D}(I)$, $I \in \mathcal{P}_t^{[n]}$, is a valid solution of (2), i.e. a family of locally consistent distributions of vertex covers.*

Proof. The second constraint of (2), i.e. local consistency, follows from the following simple geometric fact: the probability distribution $\mathcal{D}(I)$ only depends on the pairwise Euclidean distances of vertices in I and the parameter t. Given this simple observation it is not hard to see that $\mathcal{D}(U_1)$ is just the marginal of $\mathcal{D}(U)$ when $U_1 \subseteq U$.

It therefore remains to argue that $\mathcal{D}(I)$ is a distribution of vertex covers, i.e. the first constraint of (2). To that end, we need to show that in the Experiment Local-Global, two adjacent vertices cannot be at the same time outside the random cap C. This is true simply because the cap is big enough. In particular, for any two vertices i, j outside the cap if $\mathbf{z}_i, \mathbf{z}_j$ are their vectors and $\mathbf{w}$ is the vector corresponding to the tip of the cap, $\mathbf{w} \cdot \mathbf{z}_i, \mathbf{w} \cdot \mathbf{z}_j > \sqrt{\delta}$ which implies $\|\mathbf{z}_i + \mathbf{z}_j\| = \|\mathbf{w}\| \, \|(\mathbf{z}_i + \mathbf{z}_j)\| \geq \mathbf{w} \cdot (\mathbf{z}_i + \mathbf{z}_j) > 2\sqrt{\delta}$, where the penult inequality is Cauchy-Schwarz. Since G is a subgraph of B_γ^m, we conclude that ij cannot be en edge. ◀

All that remains is to show that the objective value of (2) for our solution is indeed small. In fact, we can show a stronger statement, not only is the objective value $n/2 + o(n)$ but each vertex roughly contributes 1/2 to the objective value. In particular we can show the following lemma.

▶ **Lemma 4.2.** *For any fixed $\mathbf{z} \in S^{t-1}$, we have $\mathbb{P}_{\mathbf{w} \in S^{t-1}}[\mathbf{w} \cdot \mathbf{z} \leq \eta] \leq \frac{1}{2} + \eta \sqrt{\frac{\pi}{8}(t+1)}$, when $\mathbf{w}$ is distributed uniformly on S^{t-1}. Consequently, for any vertex $i \in I$ of the graph G (subgraph of B_δ^m), we have $\mathbb{P}_{S \sim \mathcal{D}(I)}[i \in S] \leq \frac{1}{2} + \sqrt{\delta\pi(t+1)/8}$.*

The following theorems follow from Lemma 4.2. The proofs can be found in the full version.

▶ **Theorem 4.3.** *Let G be a finite subgraph of B_δ^m on n vertices. Then the level-$\left(\frac{2\epsilon^2}{\pi}\frac{1}{\delta}-1\right)$ Sherali-Adams relaxation (2) for vertex cover has objective value at most $(1/2+\epsilon)n$ for G.*

▶ **Theorem 4.4.** *For every ϵ, there are graphs on n vertices such that the level-$\Omega(\frac{\log n}{\log\log n})$ LP derived by the Sherali-Adams system for* Vertex Cover *has integrality gap $2-\epsilon$.*

5 Fooling SDPs derived by the Sherali-Adams System

5.1 Preliminary Observations for the Sherali-Adams SDP Solution

Let $\mathbf{y}$ be a Sherali-Adams solution of the LP (2), namely $y_I = \mathbb{P}_{S\sim\mathcal{D}(I)}[I \subseteq S]$. Then $\mathbf{y}$ uniquely determines the matrix $M_1 = M_1(\mathbf{y})$ in (3). In order to establish a Sherali-Adams SDP integrality gap, we need to show that $M_1(\mathbf{y})$ is positive-semidefinite for an appropriately chosen $\mathbf{y}$.

It is convenient to denote by $M_1'(\mathbf{y})$ the principal submatrix $M_1(\mathbf{y})$ indexed by nonempty sets. Note that for the solution we introduced in the previous section, all $y_{\{i\}}$ attain the same value, say y_R. In other words, $M_1(\mathbf{y}) = \begin{pmatrix} 1 & \mathbf{1}y_R \\ \mathbf{1}^T y_R & M_1'(\mathbf{y}) \end{pmatrix}$, where $\mathbf{1}$ denotes the all 1 vector of appropriate size. We leave the proof of the following fact for the full version.

▶ Fact 5.1. Suppose that $\mathbf{1}$ is an eigenvector for $M_1'(\mathbf{y})$. Then $M_1(\mathbf{y}) \succeq 0$ iff $M_1'(\mathbf{y}) \succeq 0$ and for some $j \in V$, $\mathrm{avg}_{i\in V} y_{\{i,j\}} \geq y_R^2$.

The next Lemma establishes a sufficient condition for solutions fooling SDP relaxations for Borsuk graphs. The proof uses the standard tool of tensoring introduced in Section 2.1.

▶ **Lemma 5.2.** *Let $\mathbf{y}$ be a level-t Sherali-Adams solution for* Vertex Cover *for a Borsuk graph with vector representation $\mathbf{u}_i$ and suppose that the value $y_{\{i,j\}}$ can be expressed as $f(\mathbf{u}_i \cdot \mathbf{u}_j)$. If the Taylor expansion of $f(x)$ has no negative coefficients, then $M_1'(\mathbf{y}) \succeq 0$.*

Proof. Consider the Taylor expansion of $f(x) = \sum_{i=0}^{\infty} a_i x^i$, where $a_i \geq 0$. We map $\mathbf{u}_i \in S^{m-1}$ to an infinite dimensional space as follows $\mathbf{u}_i \mapsto T_f(\mathbf{u}_i)$. Then the vectors $T_f(\mathbf{u}_i)$ constitute the Cholesky decomposition of $M_1'(\mathbf{y})$, and therefore $M_1'(\mathbf{y}) \succeq 0$. ◀

Now we examine the Sherali-Adams solution of some special case that will be instructive for our general argument. Consider some n vertex subgraph $G = (V, E)$ of $B_{\rho^2}^m$ with vector representation $\mathbf{z}_i \in S^{m-1}$. Suppose also that edges $ij \in E$ appear exactly when $\mathbf{z}_i \cdot \mathbf{z}_j = -1 + 2\rho^2$, and that for all other pairs $i, j \in V$ we have $\mathbf{z}_i \cdot \mathbf{z}_j \geq -1 + 2\rho^2$. Run Experiment Local-Global with parameters $t = 2$ and $\delta = \rho^2$ to define the level-2 Sherali-Adams solution $\mathbf{y}$

$$y_I = \mathop{\mathbb{P}}_{\mathbf{w}\in S^1}[\mathbf{w}\cdot\mathbf{z}_i \leq \rho,\ \forall i \in I] \tag{4}$$

for all I of size at most 2, where $\mathbf{w}$ is distributed uniformly on the circle.

▶ Claim 5.3. The values $y_{\{i,j\}}$ depend on the inner product $x = \mathbf{z}_i \cdot \mathbf{z}_j$ in the following way: (a) if $2\rho^2 \geq x+1$, $y_{\{i,j\}} = 1 - \frac{2\theta_\rho}{\pi}$, (b) if $2\rho^2 \leq x+1$, $y_{\{i,j\}} = 1 - \frac{\theta_\rho}{\pi}$; where $\theta_x = \arccos(x)$. In particular, when $\mathbf{z}_i \cdot \mathbf{z}_j = 1$, $y_{\{i,j\}} = 1 - \frac{\theta_\rho}{\pi}$.

The next fact is motivated by the condition of Lemma 5.2.

▶ Fact 5.4. If $\rho \in [0,1]$, the Taylor expansion of the function $1 - \frac{\theta_\rho}{\pi} - \frac{\theta_x}{2\pi}$ has no negative coefficient.

We leave the proofs of Claim 5.3 and Fact 5.4 for the full version.

Note that if we start with a configuration of vectors $\mathbf{z}_i$ for which $\mathbf{z}_i \cdot \mathbf{z}_j \geq -1 + 2\rho^2$ for all pairs $i, j \in V$, then the value $y_{\{i,j\}}$ will be described as a function on the inner product $\mathbf{z}_i \cdot \mathbf{z}_j = x$, and this function on x will have Taylor expansion with nonnegative coefficients. Unfortunately, for our Sherali-Adams solution of the previous sections this is not the case. We establish this extra condition in Section 5.2, making sure that $M_1'(\mathbf{y})$ is positive semidefinite. Proving that the matrix $M_1(\mathbf{y})$ is positive semidefinite will require one extra simple argument, which is self evident from fact 5.1.

5.2 An Easy level-2 Sherali-Adams SDP Solution

In this section we apply the techniques developed in Section 5.1 to show a tight integrality gap for Vertex Cover in the level-2 Sherali-Adams SDP system. This serves as an instructive example for higher levels whose proof are a smooth generalization of the arguments below. We will show,

▶ **Theorem 5.5.** *For any $\epsilon > 0$, there exist $\delta > 0$ and sufficiently big m, such that the level-2 Sherali-Adams SDP system for* Vertex Cover *on G_δ^m has objective value at most $2^m(1/2 + \epsilon)$.*

As the theorem states, we start with the Frankl-Rödl graph $G_\delta^m = (V, E)$, which is a subset of B_δ^m, with vector representation $\mathbf{u}_i$. Our goal is to define $\mathbf{y}$ in the context of Theorem 4.3, so as the matrix $M_1(\mathbf{y})$ to be positive semidefinite. Our Sherali-Adams solution as it appears in Theorem 4.3 does not satisfy the constraint $M_1(\mathbf{y}) \succeq 0$, for reasons that will be clear shortly. For this, we need to apply the transformation $\mathbf{u}_i \mapsto \mathbf{z}_i := (\sqrt{\zeta}, \sqrt{1-\zeta}\ T_P(\mathbf{u}_i))$, for some appropriate tensoring polynomial $P(x)$, and some $\zeta > 0$ (that is allowed to be a function of (m, δ)). We will use the following fact, first proved by Charikar [8].

▶ Fact 5.6. There exist a polynomial $P(x)$, with nonnegative coefficients and $P(1) = 1$, such that for all $x \in [-1, 1]$, we have $P(x) \geq P(-1 + 2\delta) = -1 + 2\delta_0$, for some $\delta_0 = \Theta(\delta)$. Moreover, for every constant $c > 0$ and for every $x \in (-c/\sqrt{m}, c/\sqrt{m})$, we have $|P(x)| = O(\sqrt{1/m})$.

We use the polynomial P of Fact 5.6 to map the vectors $\mathbf{u}_i$ to the new vectors $\mathbf{z}_i$. Note that with this transformation, for an edge $ij \in E$ we have $\mathbf{z}_i \cdot \mathbf{z}_j = \zeta + (1-\zeta)P(-1+2\delta) = \zeta + (1-\zeta)(-1+2\delta_0) = -1 + 2\left(\zeta(1-\delta_0) + \delta_0\right)$. If we denote $\sqrt{\zeta(1-\delta_0) + \delta_0}$ by ρ, then the above transformation maps G_δ^m to $G_{\rho^2}^{m'}$, where m' is the degree of the polynomial P. We are therefore eligible to run Experiment Local-Global with parameters $t = 2$ and ρ^2 on the vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\ T_P(\mathbf{u}_i))$. Then Lemma 4.1 implies that $\mathbf{y}$ as defined in (4) is a level-2 Sherali-Adams solution (the parameters δ, ζ will be fixed later). Next we show that for a slightly perturbed $\mathbf{y}$ we have that $M_1(\mathbf{y})$ is positive semidefinite.

First we observe that the context of Section 5.1 is relevant to the current configuration of vectors $\mathbf{z}_i$ and to our graph instances, since $\mathbf{z}_i \cdot \mathbf{z}_j \geq -1+2\rho^2$. If $\mathbf{u}_i \cdot \mathbf{u}_j = x$, then the value of $y_{\{i,j\}}$ is exactly $g(\zeta + (1-\zeta)P(x))$, where $g(x) = 1 - \frac{\theta_\rho}{\pi} - \frac{\arccos(x)}{2\pi}$. By Fact 5.4 we know that the function $g(x)$ has Taylor expansion with nonnegative coefficients. Since $\zeta + (1-\zeta)P(x)$ is a polynomial with nonnegative coefficients, it follows that $g(\zeta + (1-\zeta)P(x))$ has Taylor Expansion with nonnegative coefficients. Hence, we can apply Lemma 5.2 to obtain that

▶ **Lemma 5.7.** *The matrix $M_1'(\mathbf{y})$ is positive semidefinite.*

In what follows we describe a way to extend the positive semidefiniteness of $M_1'(\mathbf{y})$ to that of $M_1(\mathbf{y})$. In fact what we will show is general and holds for any level t (where t is the

Sherali-Adams level which solution $\mathbf{y}$ was engineered for). Since the entries of $M_1'(\mathbf{y})$ are a function of the inner product of the corresponding vectors of the hypercube, it follows that the all 1 vector is an eigenvector for $M_1'(\mathbf{y})$. By Fact 5.1 it follows that we need to show that $\text{avg}_{i \in V} y_{\{i,j\}} - y_{\{i\}}^2 \geq 0$. It turns out that this is not the case, but we can establish a weaker condition (described here in terms of a general sphere dimension D).

▶ **Lemma 5.8.** *There exist $c > 0$ (not depending on m, ρ), such that $\text{avg}_{i \in V} y_{\{i,j\}} - y_{\{i\}}^2 \geq -cD\rho$.*

We omit the proof of this lemma from this extended abstract. A rough estimate that suffices is that whenever two points have positive inner product, the probability that both are in a random cap is at least 1/4. It can be shown that due to the affine transformation, all but exponentially small fraction of the pairs will have positive inner products, hence we get that the average of $y_{\{i,j\}}$ is at least $1/4 - o(1)$. On the other hand, from Section 4 we know that $y_{\{i\}} \leq 1/2 + O(D\rho)$.

Boosting: It remains to show how to "boost" the solution to move from the relaxed condition to the exact, and necessary one. The idea is simple. Consider a ridiculously wasteful integral solution to Vertex Cover, namely the solution that takes all vertices. Clearly, if we take a convex combination of this solution with the existing one we still get a Sherali-Adams solution. If the weight of the integral solution is some small number $\xi > 0$ then the objective value increases by no more than $\xi/2$ which can be absorbed for arguments to go through as long as $\xi \leq \epsilon$. Owing to the strict convexity of the quadratic function, however, this simple perturbation does allow to improve the bound on averages as required by Fact 5.1. This observation is made precise in the following Lemma whose proof can be found in the full version.

▶ **Lemma 5.9.** *Let y' be the matrix $y' = (1-\xi)y + \xi J$ where J represents the all 1 solution. Also let $s = y_{\{i\}}$ and $s' = y'_{\{i\}}\}$. Then $\text{avg}_{i,j} y'_{\{i,j\}} - s'^2 = \Omega(\xi)$.*

We are now ready to formally prove Theorem 5.5.

Proof. (of Theorem 5.5) We start with the n-vertex Frankl-Rödl graph G_δ^m, with $\delta = \Theta(\frac{\log n}{\log \log n})$ so as to satisfy the conditions of Theorem 2.2. We use the polynomial of Fact 5.6 to obtain the vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\; T_P(\mathbf{u}_i))$, with $\zeta = \delta_0$ (where $\delta_0 = \Theta(\delta)$ by Fact 5.6). We set $\rho = \sqrt{\zeta(1-\delta_0) + \delta_0} = \sqrt{\Theta(\zeta)}$, and we run the Experiment Local-Global on the vectors $\mathbf{z}_i$ with parameters $t = 2$ and ρ^2, to obtain the vector $\mathbf{y}$. By Lemma 4.1, we have that $\mathbf{y}$ as defined in (4) is a level-2 Sherali-Adams solution. Note that since $\delta = o(1)$ we conclude from Lemma 4.2 that $y_{\{i\}} = 1/2 + \Theta(\delta)$.

Next we define $\mathbf{y}'$ as $(1-\xi)y+\xi J$. We already argued that $M_1'(\mathbf{y}')$ is positive semidefinite. By the above discussion (and Lemma 5.9) we conclude that $\text{avg}_{i,j} {y'}_{\{i,j\}} - {y'}_{\{i\}}^2 \geq 0$. We can therefore use Fact 5.1 to conclude that $M(\mathbf{y}') \succeq 0$. The last thing to note is that the contribution of every vertex in the objective value is $1/2 + O(\delta)$ ◀

5.3 The Level-$(t+2)$ Sherali-Adams SDP Tight Integrality Gap

For the level-$(t+2)$ SDP, we start with the n-vertex Frankl-Rödl graphs G_δ^m, $n = 2^m$ with vector representation $\mathbf{u}_i$. The value of δ is chosen so as to satisfy Theorem 2.2, namely $\delta = \Theta(\sqrt{\log m/m})$. As in Section 5.2 we apply to $\mathbf{u}_i$ two transformations; one using the tensoring polynomial of Fact 5.6 and one affine transformation. Then we use the resulting vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\; T_P(\mathbf{u}_i))$ to define a level-$(t+2)$ Sherali-Adams solution that we denote by $\mathbf{y}$. Our construction of $\mathbf{y}$ will have a parameter ρ to be set later.

Our goal is to meet the conditions of Fact 5.1. Namely, the first thing to ensure is that $M_1'(\mathbf{y})$ is positive semidefinite. In this direction, from Lemma 5.2 it suffices to show that the Taylor expansion of the function that describes the value of $y_{\{i,j\}}$, when $\mathbf{u}_i \cdot \mathbf{u}_j = u$, has Taylor expansion with nonnegative coefficients. Given that this function at 0 will always represent some probability, the problem is equivalent to showing that the first derivative of this function has such a good Taylor expansion. Our transformation on the vectors $\mathbf{u}_i$ can be thought as mapping their inner product u first to $x = P(u)$, and second x to $\kappa_\zeta(x) = \zeta + (1-\zeta)x$. Under this notation, we can show the following lemma that involves a number of technical calculations. The proof can be found in the full version of the paper.

▶ **Lemma 5.10.** *The derivative of the functional description of $y_{\{i,j\}}$ is*

$$D_\zeta(x) := -(\arccos(\kappa_\zeta(x)))'(1 - \frac{2\rho^2}{1+\kappa_\zeta(x)})^{t/2}.$$

Therefore, to conclude that $M_1'(\mathbf{y}) \succeq 0$ it suffices to show the next technical lemma. The proof requires arguments along the lines of that of Claim 5.7 and will appear in the full version.

▶ **Lemma 5.11.** *Set $t = 4$ and $\rho^2 \in [\zeta, \zeta + \zeta^3]$. Then for sufficiently small ζ, the function $D_\zeta(x)$ as it reads in Lemma 5.10 has Taylor expansion with nonnegative coefficients.*

Now we are ready to prove Theorem 1.1. First we obtain a level-$(t+2)$ Sherali-Adams solution from the vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\ T_P(\mathbf{u}_i))$ (the reader may think of $t = 4$). We need to set $\zeta = \sqrt[3]{\delta_0}$, where $\delta_0 = (1 + \min(P(x))/2$. Since the rounding parameter we need is $\rho = \sqrt{\zeta(1-\delta_0) + \delta_0}$, it is easy to see that $\rho^2 = \zeta + \zeta^3 - \zeta^4$. It follows by Lemma 5.11 that the matrix $M_1'(\mathbf{y})$ is positive semidefinite.

Now call c the constant for which $\text{avg}_{i \in V} y_{\{i,j\}} - y_{\{i\}}^2 \geq -ct\rho^2$. We also know that if $t\rho^2$ is no more than a small constant $\epsilon/10$, then $y_{\{i\}} \leq 1/2 + \epsilon$. Then define $\mathbf{y}' = (1 - 4c\epsilon)\mathbf{y} + (4c\epsilon)\mathbf{1}$. As we did for the level-2 Sherali-Adams SDP solution, the vector $\mathbf{y}'$ is a level-$(t+2)$ Sherali-Adams solution. Moreover, the matrix $M_1'(\mathbf{y}')$ is positive semidefinite, and $\text{avg}_{i \in V} y'_{\{i,j\}} - {y'}_{\{i\}}^2 \geq 0$. All conditions of Fact 5.1 are satisfied implying that $M_1(\mathbf{y}')$ is positive semidefinite. Finally, note that the contribution of the singletons is no more than $1/2 + \Theta(ct\rho^2)$. Hence, if we start with $t\rho^2 = o(1)$, the contribution of the singletons remains $1/2 + o(1)$. On the other hand, choosing $\delta = \Theta(\sqrt{\log m/m})$ results in graphs G_δ^m with no vertex cover smaller than $n - o(n)$.

The maximum value of t in Lemma 5.11 dictates the limitation on the level of our integrality gap. In particular we have the following conjecture and the proof of Theorem 1.2 is straightforward.

▶ Conjecture 5.12. Set t be any even integer and $\rho^2 \in [\zeta, \zeta + \zeta^3]$. Then for sufficiently small ζ, the function $D_\zeta(x)$ as it reads in Lemma 5.10 has Taylor expansion with nonnegative coefficients.

▶ **Theorem 5.13.** *Assuming Conjecture 5.12, for every constants $\epsilon > 0$ and t, the level-t SDP derived by the Sherali-Adams SDP system for* VERTEX COVER *has integrality gap $2 - \epsilon$.*

▶ Remark. [**On the validity of Conjecture 5.12**] Evidence for the validity of Conjecture 5.12 is both experimental and theoretical. In particular, some relatively simple arguments can show the following two statements: (a) For every $N_0 > 0$ there exist small enough $\zeta > 0$, such that the first N_0 Taylor coefficients of $D_\zeta(x)$ are positive, (b) For every $\zeta > 0$, there exist $N_0 > 0$ such that all *but* the first N_0 Taylor coefficients of $D_\zeta(x)$ are positive. While these partial results are not enough to imply Sherali-Adams SDP lowerbounds, they do seem to indicate that Conjecture 5.12 is true.

Discussion

We presented tight integrality gaps for level-6 Sherali-Adams SDP for Vertex Cover and how if a certain analytical conjecture is proved they can be extended to any constant number of rounds. Along the way we also gave an intuitive and geometric proof of tight Sherali-Adams LP integrality gaps for the same problem. While these LP integrality gaps apply to less rounds than [9] they remain highly nontrivial, yet significantly simplified.

For large t, proving Conjecture 5.12 seems challenging. We leave it as an open problem. Another open problem is to extend the ideas in this paper to construct tight Lasserre gaps for Vertex Cover and Unique Games, thus giving the strongest evidence that Unique Games cannot be solved with SDP hierarchies.

Acknowledgements: The authors wish to thank Toniann Pitassi for many helpful discussions and comments on an earlier version of the paper. The authors are also grateful to the anonymous reviewers for suggestions on how to improve the presentation.

References

1. Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for Unique Games and related problems. In *FOCS'10*. IEEE Computer Society, 2010.
2. Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2):1–37, 2009.
3. Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding semidefinite programming hierarchies via global correlation. In *FOCS'11*, 2011. To appear.
4. Mohammad Hossein Bateni, Moses Charikar, and Venkatesan Guruswami. MaxMin allocation via degree lower-bounded arborescences. In *STOC'09*, pages 543–552. ACM Press, 2009.
5. Siavosh Benabbas, Siu On Chan, Konstantinos Georgiou, and Avner Magen. The Sherali-Adams system applied to Vertex Cover: Why Borsuk graphs fool strong LPs and some tight Integrality Gaps for SDPs. *ECCC*, 17:169, 2011. Revision 2.
6. Siavosh Benabbas, Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. SDP gaps from pairwise independence. http://www.cs.toronto.edu/∼siavosh/pdf/pairwise-full-1.pdf, 2010.
7. Siavosh Benabbas and Avner Magen. Extending SDP integrality gaps to Sherali-Adams with applications to Quadratic Programming and MaxCutGain. In *IPCO'10*, pages 299–312. Springer, 2010.
8. Moses Charikar. On semidefinite programming relaxations for graph coloring and Vertex Cover. In *SODA'02*, pages 616–620. ACM Press, 2002.
9. Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Integrality gaps for Sherali-Adams relaxations. In *STOC'09*, pages 283–292. ACM Press, 2009.
10. Eden Chlamtac and Gyanit Singh. Improved approximation guarantees through higher levels of SDP hierarchies. In *APPROX'08*, pages 49–62. Springer-Verlag, 2008.
11. Wenceslas Fernandez de la Vega and Claire Kenyon-Mathieu. Linear programming relaxations of Max Cut. In *SODA'07*, pages 53–61. ACM Press, 2007.
12. Irit Dinur and Shmuel Safra. On the hardness of approximating minimum Vertex Cover. *Annals of Mathematics*, 162(1):439–486, 2005.
13. Peter Frankl and Vojtech Rödl. Forbidden intersections. *Trans. Amer. Math. Soc.*, 300(1):259–286, 1987.
14. Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis. Integrality gaps of $2 - o(1)$ for Vertex Cover SDPs in the Lovász-Schrijver hierarchy. In *FOCS'07*, pages 702–712. IEEE Computer Society, 2007.

15 Konstantinos Georgiou, Avner Magen, and Iannis Tourlakis. Vertex Cover resists SDPs tightened by local hypermetric inequalities. In *IPCO'08*, pages 140–153. Springer, 2008.
16 Konstantinos Georgiou, Avner Magen, and Iannis Tourlakis. On the tightening of the standard SDP for Vertex Cover with ℓ_1 inequalities. In *FSTTCS'09*, pages 203–214. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2009.
17 Michel X. Goemans and Jon Kleinberg. The Lovász theta function and a semidefinite programming relaxation of Vertex Cover. *SIAM J. Discrete Math.*, 11(2):196–204 (electronic), 1998.
18 Michel X. Goemans and David P. Williamson. Improved approximation algorithms for Maximum Cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.
19 Venkatesan Guruswami and Ali Kemal Sinop. Lasserre hierarchy, higher eigenvalues, and approximation schemes for quadratic integer programming with PSD objectives. *CoRR*, abs/1104.4746, 2011.
20 Hamed Hatami, Avner Magen, and Evangelos Markakis. Integrality gaps of semidefinite programs for Vertex Cover and relations to ℓ_1 embeddability of negative type metrics. *SIAM J. Discret. Math.*, 23:178–194, December 2008.
21 George Karakostas. A better approximation ratio for the Vertex Cover problem. *ACM Trans. Algorithms*, 5(4):1–8, 2009.
22 Anna R. Karlin, Claire Mathieu, and C. Thach Nguyen. Integrality gaps of linear and semi-definite programming relaxations for knapsack. In *IPCO'11*, pages 301–314. Springer-Verlag, 2011.
23 Howard Karloff and Uri Zwick. A 7/8-approximation algorithm for MAX 3SAT? In *FOCS'07*, pages 406–415. IEEE Computer Society, 1997.
24 Subhash Khot. On the power of unique 2-prover 1-round games. In *STOC'02*, pages 767–775. ACM Press, 2002.
25 Subhash Khot. On the Unique Games Conjecture (Invited Survey). In *CCC'10*, pages 99–121. IEEE Computer Society, 2010.
26 Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into l_1. In *FOCS'05*, pages 53–62. IEEE Computer Society, 2005.
27 Avner Magen and Mohammad Moharrami. Robust algorithms for maximum independent set on minor-free graphs based on the Sherali-Adams hierarchy. In *APPROX'09*, pages 258–271. Springer, 2009.
28 Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *STOC'08*, pages 245–254. ACM Press, 2008.
29 Prasad Raghavendra and David Steurer. Integrality gaps for strong SDP relaxations of Unique Games. In *FOCS'09*, pages 575–585. IEEE Computer Society, 2009.
30 Grant Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. In *FOCS'08*, pages 593–602. IEEE Computer Society, 2008.
31 Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. A linear round lower bound for Lovász-Schrijver SDP relaxations of Vertex Cover. In *CCC'07*, pages 205–216. IEEE Computer Society, 2007.
32 Hanif D. Sherali and Warren P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.*, 3(3):411–430, 1990.
33 Madhur Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *STOC'09*, pages 303–312. ACM Press, 2009.
34 Sundar Vishwanathan. On hard instances of approximate Vertex Cover. *ACM Trans. Algorithms*, 5(1):Art. 7, 6, 2009.