Towards strong nonapproximability results in the Lovász-Schrijver hierarchy

Mikhail Alekhnovich* Sanjeev Arora† Iannis Tourlakis‡

Abstract

Lovász and Schrijver described a generic method of tightening the LP and SDP relaxation for any 0-1 optimization problem. These tightened relaxations were the basis of several celebrated approximation algorithms (such as for MAX-CUT, MAX-3SAT, and SPARSEST CUT).

We prove strong nonapproximability results in this model for well-known problems such as MAX-3SAT, Hypergraph Vertex Cover and Minimum Set Cover. We show that the relaxations produced by as many as $\Omega(n)$ rounds of the LS_+ procedure do not allow nontrivial approximation, thus ruling out the possibility that the LS_+ approach gives even slightly subexponential approximation algorithms for these problems.

We also point out why our results are somewhat incomparable to known nonapproximability results proved using PCPs, and formalize several interesting open questions.

1 Introduction

The past decade has seen a dramatic improvement of our understanding of the approximation properties of many NP-hard optimization problems. Many new approximation algorithms were designed, especially using linear programming (LP) or semidefinite programming (SDP) relaxations. For many problems it was proved using probabilistically checkable proofs (PCPs) that these algorithms are the best possible polynomial-time algorithms unless **NP** has subexponential time algorithms.

This paper is motivated by two nagging facts about the state of the art. First, for some problems there is a large gap between the approximation ratio achieved by the best algorithms known and the approximation ratio ruled out by PCP-based results. For instance, the two ratios are 1.5 and 1.02 respectively for metric TSP, and 2 and 1.36... respectively for Vertex Cover in graphs. Second, current PCP-based results do not rule out the existence of slightly subexponential-time approximation algorithms. This happens because they often use reductions that greatly increase the instance size. For example, the reduction from 3SAT to Vertex Cover by Dinur and Safra [6] reduces 3SAT instances of size n to 1.36-approximation for Vertex Cover on graphs of size n^C where C is an astronomical constant. Thus it does not rule out that 1.2-approximation to Vertex Cover is possible in say $2^{n^{0.01}}$ time —an interesting possibility.

A recent paper of Arora, Bollobás, and Lovász [2] pointed out both these issues. It also suggested a concrete approach to study such questions: rule out good approximation algorithms that use "standard methods" of writing LP and SDP relaxations. After all, even though linear programming is P-complete, and hence in principle capable of representing arbitrary polynomial-time computations, current approximation algorithms are designed by writing LP relaxations in a certain way. A lowerbound for "large" families of relaxations could thus be viewed as a lowerbound for a restricted but important computational model (analogous to lowerbounds for

*Institute for Advanced Study, Princeton, US. Supported by CCR grant $\mathcal{N}$CCR-0324906.

†Supported by David and Lucile Packard Fellowship and NSF grants CCR 0205594. `arora@cs.princeton.edu`

‡Princeton University, `itourlak@cs.princeton.edu`

restricted circuit classes or proof systems). Arora et al. proved that several families of relaxations cannot achieve an approximation ratio better than $2 - o(1)$ for Vertex Cover. Related work of this nature is described below.

In this paper, we concentrate on lowerbounds for the approximation ratio of relaxations obtained by a general technique of Lovász and Schrijver [12]. (A related "lift and project" technique was also proposed by Sherali and Adams [14].) Given an arbitrary relaxation of a 0-1 optimization problem, this gives two procedures LS and LS_+ for obtaining tighter and tighter relaxations for the integral polytope (formal definitions appear in Section 2). The relaxation obtained from r rounds is solvable in $n^{O(r)}$ time. Thus though $r = O(1)$ is the most interesting case, if we are also interested in subexponential algorithms then any value of r less than $n/\log n$ is also interesting. In general n rounds suffice to obtain the integral polytope, which achieves an approximation ratio 1. But it is conceivable that much fewer rounds suffice to get a very nontrivial approximation ratio. The well-known SDP relaxation in the Goemans-Williamson [10] 0.878-approximation to MAX-CUT is obtained by using one round of LS_+ on the standard LP relaxation, and the SDP relaxation with triangle inequalities used in the recent Arora-Rao-Vazirani [3] approximation for SPARSEST CUT is implied by two rounds of LS_+.

We show that $\Omega(n)$ rounds of LS_+ do not suffice to achieve the following approximations for any $\epsilon > 0$: (i) approximating MAX-3SAT within a factor better than $7/8 - \epsilon$ (ii) approximating Vertex Cover in rank-k hypergraphs within a factor better than $k - 1 - \epsilon$ (iii) approximating SET COVER within a factor better than $(1 - \epsilon)\ln n$.

Note that there are nonapproximability results in the PCP setting where all the above factors appear [5, 11, 8, 13]. However, as mentioned already, those results use reductions that greatly blow up the instance size, and hence imply the above integrality gaps—under any complexity assumption at all—for n^{δ} rounds instead of for $\Omega(n)$ rounds. (An integrality gap of $(1 - \epsilon)\ln n$ for SET COVER would be implied only for $n^{o(1)}$ rounds [8].)

We further note a curious difference between the above lowerbounds for LS_+ and PCP-based nonapproximability results. In the PCP world, once we have proved an inapproximability results for "canonical" problems (such MAX-3SAT) we can use reductions to prove inapproximability results for many other problems. Proving integrality gaps via reductions in the LS world seems much harder if not impossible. In general this should not be surprising, since reductions use arbitrary polynomial-time computations, which may be outside the purview of the limited "reasoning" available in the LS_+ system (note that LS_+ is technically a proof system). What is more surprising to us is that even the simple gadget-based reductions typically encountered in NP-hardness proofs seem outside the purview of LS_+ reasoning. To give an example, approximating MAX-3SAT within a factor better than 7/8 is reducible via a textbook reduction (carried out entirely with local gadgets) to approximating Vertex Cover in graphs within a factor better than 17/16. Nevertheless, we are unable to rule out $17/16 - \epsilon$ (or even weaker) approximations to Vertex Cover in graphs, even though we have ruled out $7/8 - \epsilon$ approximations to MAX-3SAT. We describe the difficulties in Section 7.

This raises the tantalizing possibility that the familiar interrelationships among approximation problems that emerged in the past decade breaks down when one considers subexponential time approximation algorithms. Only further work can resolve such issues, and we list some interesting open problems later. (We should mention that our lowerbound results were motivated by a failure to prove upperbounds.)

Comparison with related results. Goemans and Tuncel show that that LS_+ procedure cannot derive some simple inequalities in $\Omega(n)$ rounds, showing its limitations. Results relating to integrality gaps appear in more recent papers. Arora et al. [2] show that the integrality for Vertex Cover remains $2 - \delta$ after $\Omega(\sqrt{\log n})$ rounds of LS. Feige and Krauthgamer [7] show a large gap remains for the maximum independent set problem after $\Omega(\log n)$ rounds of LS liftings.

Buresh-Oppenheim et al. [4] considered the problem of proving integrality gaps from the angle of propositional proof complexity. In the proof complexity setting, LS-type procedures can be viewed as deduction systems with a prescribed set of derivation rules. The axioms are the polytope constraints and the derivation rules give the inequalities implied by one round of LS_+. (See also Section 7.) Their paper [4] shows a linear lower bound on the number of LS_+ rounds needed to refute an unsatisfiable linear system for MAX-kSAT and MAX-kXOR-SAT when $k \geq 5$.

Our work has been partially inspired by this paper, namely we tried to apply their methods for other optimization problems. It is not difficult to show how ideas similar to [4] allow one to prove (albeit loose) integrality gaps for Vertex Cover on rank-k hypergraphs for big values of k. However their techniques seem to break down for $k = 3$ and $k = 4$—the most interesting cases after the celebrated vertex cover problem. We overcome the arising difficulties by introducing two new elements: a more subtle *expansion correction* strategy as well as a new family of positive semidefinite matrices specifically designed to protect noninteger points in the vertex cover polytopes. The latter concept plays a key role in proving rank lower bounds for LS-lifting procedures, so we would like to elaborate more on it.

With a couple of exceptions, lowerbounds in all prior papers use a certain "*protection lemma*" due to Lovász and Schrijver (see Section 2): this gives a sufficient condition for a point x outside the integral hull to survive one round of lifting. In particular, the protection lemma reduces the problem of showing that such a point survives r rounds, to showing that some specific set T of points survives $r - 1$ rounds. The simple protection lemma observes that it suffices to let T be the $2n$ points that differ from x in exactly one coordinate. (The lone exceptions are two proofs in [12] and [2] where the set T is obtained by LP duality and not explicitly described.) The simple protection lemma fails to prove the integrality gaps that interest us, and we introduce a new protection lemma that may be of interest in subsequent work. One curious feature is that in order for this protection lemma to work for even one round, we need the underlying problem instance to have some *expansion* properties. In fact, expansion plays a key role in our lowerbounds, analogous to the role it played in Buresh-Oppenheim et al. [4].

2 Relaxations, Tightenings and our Methodology

Using the Vertex Cover problem (for graphs) we explain relaxations and how to tighten them using LS_+ lifting. The integer program (IP) characterization for a graph $G = (V, E)$ is: minimize $\sum_{i \in V} v_i$ such that $v_i + v_j \geq 1$ for all $\{i, j\} \in G$ in the graph, where $v_i \in \{0, 1\}$. The *integer hull* denoted I, is the convex hull of all solutions to this problem. The standard LP relaxation is to allow $0 \leq v_i \leq 1$. The value of the LP is no more than that of the IP. Let P be the convex hull of all solutions vectors in $[0, 1]^n$ to the LP. A linear relaxation is *tightened* by adding more and more constraints that also hold for the integral hull; in general this gives some polytope P' such that $P \subseteq P' \subseteq I$.

The quality of a linear relaxation is measured by $\frac{\text{optimum value over } I}{\text{optimum value over } P}$, usually called its *integrality gap*. For the vertex cover relation, this ratio is 2. (Note: When designing approximation algorithms, one also needs some kind of *rounding algorithm* to convert fractional solutions to integer ones in polynomial time, but we ignore this aspect.)

Lovász and Schrijver [12] present a so-called "lift-and-project" technique for deriving tighter and tighter relaxations of a 0-1 integer program. The n dimensional relaxed polytope is lifted to n^2 dimensions, new constraints are introduced, and then projected back into the original space.

The notation for the LS-liftings uses homogenized inequalities. In particular, given a polytope $P \subseteq \mathbb{R}^n$ for a linear relaxation, let $Q \subseteq \mathbb{R}^{n+1}$ be the cone $\{\binom{a}{\vec{x}} : \vec{x}/a \in P\}$. For example, if P is the Vertex Cover polytope, then $(x_0, x_1, \ldots, x_n) \in Q$ iff the edge constraints $x_i + x_j \geq x_0$ hold for all edges $\{x_i, x_j\}$ in the graph. Denote by $N^r(Q)$ and $N_+^r(Q)$ the feasible cone of all inequalities derivable in r rounds of the LS and LS_+ lifting procedure, respectively. Let $N^0(Q) = N_+^0(Q) = Q$. The rth round polytope is then obtained by projecting along the hyperplane $x_0 = 1$. We will often abuse notation and write $N_+^r(P)$ to indicate the polytope obtained after r rounds of LS_+ lifting. Lovász and Schrijver prove that if there exists a polynomial time separation oracle for P then one can optimize a linear function over $N^r(P)$ and $N_+^r(P)$ in time $n^{O(r)}$.

The following lemma from [12] characterizes the N and N_+ operators.

Lemma 2.1. *Let Q be a cone as defined above. Then $y = \binom{1}{x}$ is in $N(Q)$ iff there is a symmetric matrix $Y \in \mathbb{R}^{(n+1)\times(n+1)}$ such that:*

1. *$Ye_0 = diag(Y) = y$,*
2. *For each i such that $x_i = 0$, $Ye_i = 0$; for each i such that $x_i = 1$, $Ye_i = y$; Otherwise Ye_i/x_i and $Y(e_0 - e_i)/(1 - x_i)$ are both in Q.*

Moreover, $y = \binom{1}{x}$ is in $N_+(Q)$ iff in addition Y is positive semidefinite.

Note that $N_+^{r+1}(Q) \subseteq N_+^r(Q)$. Note that to prove that $y \in N_+^{r+1}(Q)$, we have to construct a specific matrix Y and prove that the $2n$ vectors defined in Lemma 2.1 are in $N_+^r(Q)$. Choosing such a Y gives what Buresh-Oppenheim et al. call a "protection lemma": a point survives one round of LS if the $2n$ vectors given by Y are in the previous polytope.

The simplest Y one could conceive of is $Y_{ij} = y_i y_j$ (this is trivially positive semidefinite); however, this matrix satisfies $diag(Y) = y$ only if y is a 0-1 vector. Indeed, this proves that all polytopes resulting from LS-type liftings contain the integral hull. The next simplest Y one could conceive is $Y = yy^T + \mathrm{Diag}(y - y^2)$, that is, the matrix that has $Y_{ij} = y_i y_j$ except along the diagonal where $Y_{ii} = y_i$. For $y \in [0, 1]^n$ this is clearly positive semidefinite, and indeed, this matrix was used in early results by Lovasz-Schrijver [12] and Goemans-Tuncel [9], and more recently, Buresh-Oppenheim et al. [4].

With this choice of Y, the vectors Ye_i/x_i and $Y(e_0 - e_i)/(1 - x_i)$ are obtained from changing *one* coordinate in y to a 0 or a 1. However, for MAX-3SAT and hypergraph Vertex Cover, these vectors are not guaranteed to be in the polytope. Thus other than for the Set Cover problem, this simple protection lemma does not suffice for us.

Our response is to use a more complicated Y, such that most entries satisfy $Y_{ij} = y_i y_j$, but some don't. Then the $2n$ vectors generated above correspond to modifying Y in a small number of entries. (A similar idea occurred in [2], except the Y there was not explicit.) This is at the heart of our new protection lemmas for MAX-3SAT and hypergraph Vertex Cover. To make this choice of Y work out, we need certain expansion requirements to be met.

With our "protection lemma" in hand, the lowerbound strategy will be as follows: Given our relaxed polytope P, we identify a point $w \in P$ for which the ratio between the integral optimum and the value of the objective function at w is large. We will then prove the lowerbound by showing that w survives many rounds of LS_+. We do this via a Prover-Adversary game where the Prover is trying to prove that $w \in N_+^r(P)$ and the Adversary's goal is to show the opposite. For the Adversary to win, it will suffice for him to exhibit a vector amongst the $2n$ vectors given by our "protection lemma" that is not in $N_+^{r-1}(P)$. He picks such a vector x and "challenges" the Prover to show it is in $N_+^{r-1}(P)$. Things continue this way, and the Prover loses if she cannot keep the game going for r steps. To keep the argument clean, we need to maintain the vector x in a nice form throughout the game. To this end, we borrow an idea from [4]: during each round, to prove that a particular point x is in a certain polytope, the Prover can also choose to express the point as a convex combination $\sum_j \rho_j z_j$ and claims that every $z_j \in N_+^{r-1}(P)$ (and consequently so is x). To counter this claim, the Adversary picks some z_j which he thinks is not in $N_+^{r-1}(P)$, and the game continues for that vector. We will show that if the equations defining P satisfy certain expansion requirements, then for appropriate w, the Prover has a linear round strategy against any Adversary.

3 Incidence Graphs of Equations and their Properties

Given a hypergraph $G = (V, E)$, let H_G be the bipartite incidence graph on $E \times V$ where each each hyperedge is connected to the vertices it contains. We will often use the notion of *expansion* in a bipartite graph.

Definition 3.1. A bipartite graph $G = (V_1, V_2, E)$ is an (r, c)-*expander* if every subset $S \subseteq V_1$, $|S| \leq r$, satisfies $|\Gamma(S)| \geq c|S|$, where $\Gamma(S)$ is the set of neighbours of S in V_2.

Throughout this paper we will deal with constraints of the form $\sum_i v_i^{\epsilon_i} \geq 1$ where $v_i^{\epsilon_i}$ represents v_i if $\epsilon_i = 1$ and $v_i^{\epsilon_i}$ represents $1 - v_i$ if $\epsilon_i = 0$. Say that a variable $v_i^{\epsilon_i}$ occurs *negated* in a constraint if $\epsilon_i = 0$. Let C be a set of such constraints on a set V of n variables. Given an assignment vector $x \in [0,1]^n$ for V, we define $C(x)$ to be the set of constraints obtained from C as follows: (a) If $x_i = 0$, remove all constraints containing v_i negated; (b) if $x_i = 1$, remove all constraints containing v_i unnegated; and (c) remove all variables set to 0-1 by x from the remaining constraints. Intuitively, $C(x)$ is the set of simplified constraints in C not trivially satisfied by x. In particular, if x satisfies $C(x)$, then x satisfies C.

Let $V(x)$ be the set of those variables in V not set to 0-1 by x and let $H(x)$ be the bipartite incidence graph on $C(x) \times V(x)$; that is, for each equation in $C(x)$ there is an edge to every variable it contains. Let H be the incidence graph on $C \times V$. We will say that $C(x)$ is an (r,c)-expander if $H(x)$ is an (r,c)-expander. We will say that the *arity* of a constraint is t if it has t neighbours in $H(x)$. For a subset $S \subseteq C(x)$ of equations, denote the variables in S (i.e., the neighbours of S in $H(x)$) by $\Gamma(S)$.

Usually $C(x)$ will have some expansion property, and in particular will be at least a $(2, k-1-\epsilon)$-expander. Then all equations in $C(x)$ will have arity at least $k-1$. Moreover, whenever $C(x)$ is an expander, equations of arity $k-1$ will enjoy some special properties of which we will take advantage. For a vector $x \in \mathbb{R}^n$, let $R(x)$ to denote the set of all indices to non-integral coordinates of x.

Definition 3.2. Let $0 < \epsilon < 1/2$ and $x \in \{0, \frac{1}{k-1}, 1\}^n$ and suppose $C(x)$ is a $(2, k-1-\epsilon)$-expander. Two indices $i, j \in R(x)$ are $C(x)$-*equivalent* (written $i \sim_{C(x)} j$) if there is a constraint in $C(x)$ of arity $k-1$ containing v_i and v_j. Let $E(x) \subseteq R(x)$ contain all indices $i \in R(x)$ for which there exists $j \in R(x), j \neq i$ such that $i \sim_{C(x)} j$.

The following proposition will be used repeatedly in our lower bound proofs and follows easily from expansion.

Proposition 3.3 (Facts about $C(x)$-equivalence classes). *Let $0 < \epsilon < 1/2$ and $x \in \{0, \frac{1}{k-1}, 1\}^n$ and suppose $C(x)$ is $(2, k-1-\epsilon)$-expanding.*

1. *A given variable can only occur in one arity $k-1$ equation in $C(x)$. Hence, each $C(x)$-equivalence class has exactly $k-1$ elements.*
2. *Any given equation in $C(x)$ (other than the arity $k-1$ equation defining the equivalence) can contain at most one variable from any given $C(x)$-equivalence class.*

4 Lowerbounds for Hypergraph Vertex Cover

Let $G = (V, E)$, $E \subseteq V^k$, be a k-uniform hypergraph. The vertex cover problem for G is to find the smallest subset $S \subseteq V$ such that all hyperedges in G contain at least one element from S. The problem is expressed by the following integer program where variable $v_i \in \{0,1\}$ corresponds to vertex i in the graph:

$$\min \sum_{v_i \in V} v_i : \qquad \sum_{j=1}^{k} v_j \geq 1 \qquad \forall (v_1, \ldots, v_k) \in E. \tag{4.1}$$

Let us relax to $0 \leq v_i \leq 1$. Let $\mathrm{VC}(G)$ then be the polytope consisting of all points $w \in \mathbb{R}^n$ satisfying the relaxed constraints, that is, $w \in \mathbb{R}^n$ is in $\mathrm{VC}(G)$ if setting $v_i = w_i$ for all i results in all relaxed constraints being satisfied. It is easy to see for the complete k-uniform hypergraph on n vertices the optimal value of the integer program is $n - k + 1$ while the optimum value of the relaxed linear program is n/k. Therefore, the integrality gap between the integer and linear programs is at least $k - o(1)$.

We prove that even after a linear number of rounds of LS_+ tightenings of $\mathrm{VC}(G)$ there still exists some graph for which the integrality gap is $k - 1 - o(1)$.

Theorem 4.1. *Let $k \geq 3$. For all $\alpha > 0$ there exists $\gamma > 0$ and a k-uniform hypergraph G such that the integrality gap of any γn round LS_+ relaxation of* $\mathrm{VC}(G)$ *is at least* $(k-1)(1-\alpha)$.

Given $G = (V, E)$, let C_G be the set of hyperedge constraints defining $\mathrm{VC}(G)$. Since the underlying graph G will usually be clear, we omit the subscript unless extra precision is needed. In this section we will always have $x \in \{0, \frac{1}{k-1}, 1\}^n$ and $C(x)$ will be at least a $(2, k-1-\epsilon)$-expander. Then all equations in $C(x)$ will have arity at least $k-1$ and the following will hold:

Proposition 4.2. *Let $0 < \epsilon < 1/2$, and $x \in \{0, \frac{1}{k-1}, 1\}^n$, and suppose that $C(x)$ is $(2, k-1-\epsilon)$-expanding. Then $x \in \mathrm{VC}(G)$.*

We now define the vectors that will appear in our "protection lemma" for vertex cover. For the remainder of this section we will always assume $0 < \epsilon < 1/2$.

Definition 4.3. Given $x \in [0,1]^n$, for all $i \in R(x)$ and all $a \in \{0,1\}$ define $x^{(i,a)}$ to be identical to x except that $x_i^{(i,a)} = a$.

Definition 4.4. Let $x \in \{0, \frac{1}{k-1}, 1\}^n$, and suppose $C(x)$ is $(2, k-1-\epsilon)$-expanding. For all $i \in E(x)$ define $x^{[i]}$ to be identical to x except that $x_i^{[i]} = 1$ and $x_j^{[i]} = 0$ for all $j \sim_{C(x)} i$. Let the set $T_x \subseteq \{0, \frac{1}{k-1}, 1\}^n$ equal the union $\{x^{[i]} : i \in E(x)\} \cup \{x^{(i,a)} : i \in R(x) \backslash E(x), a \in \{0,1\}\}$.

Lemma 4.5. *Let $x \in \{0, \frac{1}{k-1}, 1\}^n$, and suppose $C(x)$ is $(2, k-1-\epsilon)$-expanding. Then $R(x) \subseteq \mathrm{VC}(G)$. Moreover, for all $y \in T_x$, each equation in $C(y)$ has arity at least $k-1$.*

Proof. There are two types of points in T_x: (1) $x^{(i,a)}$ for $i \in R(x) \backslash E(x)$ and (2) $x^{[i]}$ for $i \in E(x)$. Consider a point $x^{(i,a)}$ in T_x where $i \in R(x) \backslash E(x)$. In this case, v_i does not belong to any arity $k-1$ equation in $C(x)$. Hence, every equation in $C(x^{(i,a)})$ has arity at least $k-1$ in $C(x^{(i,a)})$, and is therefore satisfied by $x^{(i,a)}$.

Now consider a point $x^{[i]}$ in T_x such that $i \in E(x)$. By Fact 2 on equivalences and the definition of $x^{[i]}$, every equation in $C(x)$ that had arity k in $C(x)$ has arity at least $k-1$ in $C(x^{(i,a)})$, and hence is satisfied by $x^{(i,a)}$. By Fact 1 on equivalences, the only arity $k-1$ equation in $C(x)$ for which the values of any of its variables changes under $x^{[i]}$ is the unique arity $k-1$ equation containing v_i. But such an equation is satisfied by $x^{[i]}$ since v_i is set to 1 in $x^{[i]}$. ◻

Lemma 4.6 ("Protection Lemma"). *Let $x \in \{0, \frac{1}{k-1}, 1\}^n$, and suppose $C(x)$ is $(2, k-1-\epsilon)$-expanding. Suppose moreover that $T_x \subseteq N_+^m(\mathrm{VC}(G))$. Then $x \in N_+^{m+1}(\mathrm{VC}(G))$.*

Proof. Let $y = \binom{1}{x}$. The proof uses Lemma 2.1 and the following choice of an $(n+1) \times (n+1)$ positive semi-definite symmetric matrix Y that is $yy^T + \mathrm{Diag}(y - y^2)$ except that $Y_{ij} = 0$ whenever $i \sim_{C(x)} j$. Note that Y is symmetric and that $Ye_0 = diag(Y) = y$. Moreover, by Proposition 9.1 in the appendix, Y is positive semi-definite. So by Lemma 2.1, to show that $x \in N_+^{m+1}(\mathrm{VC}(G))$ it remains only to show that for all $i \in R(x)$, Ye_i/x_i and $Y(e_0 - e_i)/(1-x_i)$ are in $N_+^m(\mathrm{VC}(G))$.

For $i \in R(x) \backslash E(x)$, $Ye_i/x_i = \binom{1}{x^{(i,1)}}$ and $Y(e_0 - e_i)/(1-x_i) = \binom{1}{x^{(i,0)}}$ and hence are both in $T_x \subseteq N_+^m(\mathrm{VC}(G))$. For $i \in E(x)$, $Ye_i/x_i = \binom{1}{x^{[i]}}$ which is in $T_x \subseteq N_+^m(\mathrm{VC}(G))$. Finally, for $i \in E(x)$, $Y(e_0 - e_i)/(1-x_i) = \binom{1}{z}$ where $z = \frac{1}{k-2} \sum_{j \sim_{C(x)} i, j \neq i} x^{[j]}$. In particular, $Y(e_0 - e_i)/(1-x_i)$ is in the convex hull of $T_x \subseteq N_+^m(\mathrm{VC}(G))$, and hence is also in $N_+^m(\mathrm{VC}(G))$. ◻

4.1 Proof of Theorem 4.1

Let $\alpha, \epsilon > 0$ be arbitrarily small. By Lemma 9.2 in the Appendix, there are constants $\beta, \delta > 0$ such that a rank k hypergraph G exists with n vertices and βn edges such that the bipartite graph H_G is a $(\delta n, k-1-\epsilon)$-expander, and every vertex cover of G has size at least $(1-\alpha)n$. We show that the vector $w = (\frac{1}{k-1}, \ldots, \frac{1}{k-1})$, corresponding to a fractional vertex cover of "size"

$n/(k-1)$, is in $N_+^r(\mathrm{VC}(G))$ where $r = \frac{\epsilon\delta n}{k-1}$. It follows that this many rounds of LS_+ cannot reduce the integrality gap below $(k-1)(1-\alpha)$, and Theorem 4.1 then follows for $\gamma = \frac{\epsilon\delta}{k-1}$.

Note that H_G is isomorphic to $H(w)$, and hence, $C(w)$ is $(\delta n, k-1-\epsilon)$-expanding. This will be crucial for the lower bound.

The lowerbound follows from a Prover-Adversary game of the type discussed in Section 2. We describe the game more formally. In round i there is a parameter $\ell_i \geq 2$ and a current point $x \in \{0, \frac{1}{k-1}, 1\}^n$. For $i = 0$, x is w. At the beginning of round i, $C(x)$ will be an $(\ell_i, k-1-2\epsilon)$-expander. In round i the following two moves are made.

1. **Adversary Move:** The Adversary selects z from T_x.
2. **Expansion Correction:** The Prover constructs a set $Y \subseteq \{0, \frac{1}{k-1}, 1\}^n$ such that (1) z is in the convex hull of Y, and (2) for all $y \in Y$, $C(y)$ is an $(\ell_{i+1}, k-1-2\epsilon)$-expander where $\ell_{i+1} \leq \ell_i$. The Adversary selects one point $y \in Y$ to be the new x.

The game ends when $\ell_{i+1} \leq 1$.

Intuitively, the Adversary fixes more and more coordinates in the initial point w to 0-1 values by replacing the current point x with a point z from T_x (note that once a coordinate is set to 0-1 it remains fixed). The Prover wants this to continue for as long as possible but may run into trouble if $C(z)$ is no longer a good expander. The Prover therefore does Expansion Correction to obtain a new x for which $C(x)$ is a good expander. The next lemma shows that a good Prover strategy implies w has high rank.

Lemma 4.7. *If the Prover has an m round strategy against any adversary, then $w \in N_+^m(\mathrm{VC}(G))$.*

Proof. By induction on m. By Proposition 4.2, $w \in \mathrm{VC}(G) = N_+^0(\mathrm{VC}(G))$, and case $m = 0$ follows. So suppose the claim is true for m and that the Prover has an $m+1$ round strategy against any adversary. Consider the first round of the game and suppose the Adversary picks $z \in T_x$. Let Y be the set subsequently constructed by the Prover in the Expansion Correction move. Since the game runs for m more rounds regardless of which $y \in Y$ the Adversary chooses, $Y \subseteq N_+^m(\mathrm{VC}(G))$ by induction, and $z \in N_+^m(\mathrm{VC}(G))$ by convexity. This holds no matter which $z \in T_x$ the Adversary chooses, and so $T_x \subseteq N_+^m(\mathrm{VC}(G))$. Lemma 4.6 then implies $w \in N_+^{m+1}(\mathrm{VC}(G))$. ∎

So to prove $w \in N_+^r(\mathrm{VC}(G))$ and complete the proof of Theorem 4.1, it suffices to describe an r round strategy for the Prover.

Lemma 4.8. *If $C(w)$ is a $(\delta n, k-1-\epsilon)$-expander, then the Prover has an r round strategy against any Adversary, where $r = \frac{\epsilon\delta n}{k-1}$.*

Proof. We start the game with $x = w$. Proposition 4.2 implies $w \in \mathrm{VC}(G)$. In round i of the strategy the parameter ℓ_i will be defined such that for the current point x the Prover can ensure $C(x)$ is an $(\ell_i, k-1-2\epsilon)$-expander. At the start, $\ell_1 = \delta n$.

The strategy will work as follows: The two moves made in each round of the game remove more and more variable vertices from the incidence graph $H(w)$ on $C(w) \times V(w)$. In each round at most $k-1$ variable vertices are removed from $H(w)$ by the Adversary choosing $z \in T_x$. As for the Expansion Correction move, the Prover will "correct" expansion in round i by identifying a maximal non-expanding set S_i of equations of size at most ℓ_i and removing it and its neighbours from $H(x)$. Letting $\ell_{i+1} = \ell_i - |S_i|$, the resulting graph would then be an $(\ell_{i+1}, k-1-2\epsilon)$-expander. The Prover removes these equations in S_i by having the assignments Y be 0-1 on $\Gamma(S_i)$ and equal to x outside $\Gamma(S_i)$. If $\ell_{i+1} \leq 1$, the game ends; otherwise, the game continues. The claim is that such a strategy results in at least r rounds: Suppose the strategy lasts m rounds and consider $S = \cup S_i$. Then $|S| = \sum_{i=1}^m |S_i| = \sum_{i=1}^m \ell_i - \ell_{i+1} = \delta n - \ell_{m+1}$. By expansion, S had at least $(k-1-\epsilon)|S|$ neighbours in $H(w)$. However, at the end of the game, S has no neighbours. Expansion Correction removes at most $(k-1-2\epsilon)|S|$ neighbours. Since the Adversary Move removes at most $k-1$ neighbours per round, there must be at least $\epsilon\delta n/(k-1)$ rounds.

It remains to describe the Prover's strategy in round i in detail: If $\ell_i \leq 1$ the game ends. Otherwise, Proposition 4.2 implies $x \in \mathrm{VC}(G)$ and the Adversary selects $z \in T_x$. Note that Lemma 4.5 implies $z \in \mathrm{VC}(G)$ and that every equation in $C(z)$ has arity at least $k-1$. We now describe how the Prover constructs the set Y for Expansion Correction:

1. If $C(z)$ is an $(\ell_i, k-1-2\epsilon)$-expander, the Prover takes $Y = \{z\}$ and sets $S_i = \emptyset$.
2. Otherwise, let $S_i \subseteq C(z)$, $|S_i| \leq \ell_i$, be a maximal subset of equations with expansion less than $k-1-2\epsilon$ in $C(z)$. If $|S_i| \geq \ell_i - 1$, i.e., $\ell_{i+1} \leq 1$, the game ends, and we let the final x be the same as z except it is 0 on $\Gamma(S_i)$.
3. Otherwise we claim that for all subsets $S' \subseteq S_i$ of equations in $C(z)$, $|\Gamma(S')| > (k-2)|S'|$: Either the Adversary chose some $x^{(j,a)} \in T_x$ where j is not in any $C(x)$-equivalence class (in which case S' has expansion greater than $k-2$ in $C(z)$), or it chose $x^{[j]}$ where v_j occurs in some arity $k-1$ equation $\phi \in C(x)$. Suppose ϕ shares t variables with $\Gamma(S')$. By expansion of $C(x)$,

$$|\Gamma(S')| = |\Gamma(S' \cup \{\phi\})| - k + 1 + t \geq (k-1-2\epsilon)|S'| + t - 2\epsilon.$$

   Since S' has exactly t fewer neighbours in $C(z)$ than in $C(x)$, the claim follows.
4. Let $S_i = (e_1, \ldots, e_t)$. By Lemma 9.4 in the appendix there exists a mapping $\eta : S \to \mathcal{P}(\Gamma(S))$ such that (1) for all $i \in [t]$, $|\eta(e_i)| = k-1$, and (2) for all $i \in [t]$, $|\eta(e_i) \setminus \bigcup_{j<i} \eta(e_j)| \geq k-2$. We construct $k-1$ assignments $y^1, \ldots, y^{k-1}$ inductively according to the ordering $e_1, \ldots, e_t$. At the beginning all the y^j equal x outside $C(z)$ and are undefined on $\Gamma(S_i)$. Assume that at step t the values y_i^j for all $j \in [k-1]$ and for all i such that $v_i \in \bigcup_{i'<t} \eta(e_{i'})$ have been defined so that the constructed partial assignments satisfy all $e_{i'}$, $i' < t$, and the assigned values $y_i^1, \ldots, y_i^{k-1}$ contain exactly one 1 for each i. Consider e_t. Choose $k-2$ vertices $v_{i_1}, \ldots, v_{i_{k-2}} \in \eta(e_t)$ such that the values $y_{i_1}^j, \ldots, y_{i_{k-2}}^j$ are undefined for all $j \in [k-1]$ (these vertices exist by definition of η). Let $v_{i_{k-1}}$ be the other vertex in $\eta(e_{t+1})$. If the corresponding variables $y_{i_{k-1}}^1, \ldots, y_{i_{k-1}}^{k-1}$ are undefined then set the last of these variables to one and the rest to zeros. Assume without loss of generality that $y_{i_{k-1}}^{k-1} = 1$. For all other vertices in $\eta(e_{t+1})$ we set $y_{i_j}^j = 1$ and the rest to zeros. We have extended our partial assignments for $\eta(e_t)$ in a way that satisfies the induction hypothesis. At the the end, $y^1, \ldots, y^{k-1}$ each satisfy S_i and z is their average. Let $Y = \{y^1, \ldots, y^{k-1}\}$.

□

5 Lowerbounds for MAX-3SAT

The arguments used to prove Theorem 4.1 can be adapted to prove integrality gaps for 3-SAT. Given a 3-CNF formula ϕ, we convert its clauses to inequalities in the obvious way, i.e., $x_1 \vee x_2 \vee \neg x_3$ becomes $x_1 + x_2 + (1 - x_3) \geq 1$. Let C_ϕ be the set of these inequalities. Note that the 0-1 solutions to the inequalities correspond exactly to the satisfying assignments for ϕ. Relaxing to $x_i \in [0,1]$ yields a polytope $\mathrm{SAT}(\phi)$ whose integral points are solutions for ϕ. Let $w = (\frac{1}{2}, \ldots, \frac{1}{2})$ and note that $w \in \mathrm{SAT}(\phi)$ for any formula ϕ.

Lemma 5.1. *Suppose $C_\phi(w)$ is a $(\delta n, 2-\epsilon)$-expander, $0 < \epsilon < \frac{1}{2}$. Then $w \in N_+^{\epsilon\delta n/2}(\mathrm{SAT}(\phi))$.*

Proof. The proof is similar to the proof of Lemma 4.8. The only changes are: (1) The "protection lemma" (in particular Lemmas 4.5 and 4.6) must be altered to take negated variables into account; and (2) in the game, the Prover's Expansion Correction strategy has to accommodate negated variables. We sketch proofs of these differences.

Definition 5.2. Suppose $x \in \frac{1}{2}\mathbb{Z}^n$ and let $i \in R(x), a \in \{0,1\}$. Let $x^{[i,a]} \in \frac{1}{2}\mathbb{Z}^n$ be identical to x except (1) $x_i^{[i,a]} = a$, and (2) if there exists an arity 2 equation $v_i^{\epsilon_i} + v_j^{\epsilon_j} \geq 1$ in $C(x)$, then $x_j^{[i,a]} = 1 - a$ if $\epsilon_i = \epsilon_j$ and $x_j^{[i,a]} = a$ if $\epsilon_i \neq \epsilon_j$.

Arguing as in Lemma 4.5, it follows that if $C(x)$ is $(2, 2-\epsilon)$-expanding, then for all $i \in R(x), a \in \{0,1\}$, each equation in $C(x^{[i,a]})$ has arity at least 2 and hence, $x^{[i,a]} \in \mathrm{SAT}(\phi)$. If we let $y = \binom{1}{x}$, then the "Protection Lemma" for $\mathrm{SAT}(\phi)$ will use the matrix Y that is the matrix $yy^T + \mathrm{Diag}(y - y^2)$ except that if $x_i^{\epsilon_i} + x_j^{\epsilon_j} \geq 1$ is an equation in $C(x)$, then $Y_{ij} = 0$ if $\epsilon_i = \epsilon_j$ and $Y_{ij} = \frac{1}{2}$ if $\epsilon_i \neq \epsilon_j$. It can be verified that Y is positive semidefinite, and moreover, for all $i \in R(x)$, $Ye_i/x_i = \binom{1}{x^{[i,1]}}$ and $Y(e_0 - x_i)/(1-x_i) = \binom{1}{x^{[i,0]}}$. Hence, as the set of vectors for the protection lemma we can take $T_x = \{x^{[i,a]} : i \in R(x), a \in \{0,1\}\}$.

We now sketch how the Expansion Correction Strategy is altered. The overall argument goes the same way using Lemma 9.4 with the only difference being that for $v_i \in \eta(e_t)$, the y_i^j, $j = \{1,2\}$, are set according to the signs the variables have in clause e_t so as to satisfy e_t. $\square$

It is well-known that for all $\alpha, \epsilon > 0$, there exist constants $\beta, \delta > 0$ such that if we pick a random 3-CNF ϕ with βn clauses, then with high probability (1) no boolean assignment satisfies more than a $\frac{7}{8} + \alpha$ fraction of the clauses in ϕ and (2) C_ϕ is a $(\delta n, 2-\epsilon)$-expander. On the other hand, Lemma 5.1 says that w, which satisfies all clauses in ϕ, is in $N_+^r(\mathrm{SAT}(\phi))$ where $r = \epsilon\delta n/2$. We have proved the following:

Theorem 5.3. *For any constant $\alpha > 0$, there exist constants $\beta, \gamma > 0$ such that if ϕ is a random βn clause 3-CNF formula on n variables, then the integrality gap of any γn round LS_+ relaxation of $\mathrm{SAT}(\phi)$ is at least $\frac{7}{8} - \alpha$ with high probability.*

6 Lowerbounds for SET COVER

An instance of SET COVER consists of a tuple (S, C) where C is a collection of n subsets of a finite set S of size m. The objective is to find a minimum size subset $C' \subseteq C$ such that each element of S is in some set in C'. If for each set $S_i \in C$ we have a variable x_i indicating whether or not set S_i is included in the set cover, then the SET COVER problem is expressed by the following integer program:

$$\min \sum_{i=1}^{n} x_i : \qquad \forall j \in [m], \quad \sum_{i: j \in S_i} x_i \geq 1.$$

The relaxed SET COVER polytope $\mathrm{MSC}(S, C)$ is the polytope defined by the above constraints but where we allow $0 \leq x_i \leq 1$. Note now that if $G = (V, E)$ is a k-uniform hypergraph, and we let $S = E$ and $C = \{S_v\}_{v \in V}$ where $S_v = \{e \in E : v \in e\}$, then $\mathrm{MSC}(S, C)$ is identical to $\mathrm{VC}(G)$. Hence, integrality gaps for the hypergraph VC polytope yield integrality gaps for MSC.

Fix $\epsilon, \delta, \gamma > 0$ such that $\epsilon - \delta > 0$. By Lemma 9.3 in the appendix, there exists an $(\epsilon - \delta)n$-uniform hypergraph $G = (V, E)$ on n vertices with n edges such that the minimum vertex cover is at least $\log_{1+\epsilon} n$. Consider the hyperedge constraints C_G defining $\mathrm{VC}(G)$. Let w be the all-$(\frac{1+\gamma}{(\epsilon-\delta)n})$ point and note that w is in $\mathrm{VC}(G)$. Moreover, at least $\lfloor \frac{\gamma(\epsilon-\delta)n}{1+\gamma} \rfloor$ coordinates of w can be switched to 0 or 1 with the resulting point still satisfying all the constraints C_G.

Recall the simple protection lemma used in Buresh-Oppenheim et al. [4] and described in section 2: For a relaxed polytope P, a point x is in $N_+(P)$ if for all $i \in R(x)$ and all $a \in \{0,1\}$, $x^{(i,a)}$ is in P. That is, x is in $N_+(P)$ if whenever we change exactly one coordinate of x to 0 or 1, the resulting point is in P. So by induction, this simple protection lemma together with the observation about w in the previous paragraph prove the following:

Lemma 6.1. *The point w is in $N^r(\mathrm{VC}(G))$ where $r = \lfloor \frac{\gamma(\epsilon-\delta)n}{1+\gamma} \rfloor$.*

Finally note that since the minimum vertex cover for G has size $\log_{1+\epsilon} n$, the integrality gap for w is $\frac{(\epsilon-\delta)\ln n}{(1+\gamma)\ln(1+\epsilon)}$ which approaches $\ln n$ from below as $\epsilon, \delta, \gamma \to 0$. Thus we have proved the following theorem:

Theorem 6.2. *For all $\epsilon > 0$, there exists $\delta > 0$ and an instance (S, C), $|S| = n$, of SET COVER for which the integrality gap of the δn round LS_+ relaxation of $\mathrm{MSC}(S, C)$ is $(1-\epsilon)\ln n$.*

7 Discussion

Technically speaking, proof complexity studies non-deterministic provers showing that a given problem does not have a solution, e.g., a formula is unsatisfiable. For some proof systems, whenever a proof exists it may also be found efficiently in which case the proof system is called *automatizable*. In this case, the non-deterministic prover becomes a deterministic algorithm, and propositional complexity becomes computational complexity. This is exactly what is going on when one considers LS round lowerbounds. The classical propositional complexity measure is however the number of lines in LS-style reasoning, and no lower bounds are known for this measure. Since in case of LS round lowerbounds the propositional complexity and computational look similar, we only talk about the latter in this paper.

In this view, we note why reductions are problematic. The standard reduction from 3SAT to VC replaces each clause by a triangle of vertices. One may imagine that we add new *auxiliary variables*, each variable being a function of at most 3 old variables. Indeed, if a triangle corresponds to a clause in 3SAT then the variables in these triangle are functions over the values of variables of this clause.

However, in general, when one introduces such auxiliary variables, the proof complexity may change drastically. For example, weak resolution turns into powerful Extended Frege. On the other hand, in our case all auxiliary variables are local so intuitively adding them should not make a big difference. However, our arguments using protection lemmas break down and a newer lowerbound idea seems necessary.

8 Open Problems

It seems important to extend the nonapproximability results to a variety of problems. (Or to prove that actually many important optimization problems do have good slightly subexponential time approximation algorithms via the LS_+ procedure or other lift-and-project procedures.) As we noted above, reductions are problematic in this regard.

Our paper leaves much space for further research. Methods based on games over expanders do not seem to help against the notoriously difficult vertex cover problem: there are no expanders of degree 2. This question seems related to proving $k - \epsilon$ integrality gap for k-hypergraphs (a similar picture with these problems is observed in the PCP world).

Our result for SET COVER is interesting in a different respect: In [8] integrality gaps of $(1-\epsilon)\ln n$ are only ruled out under the assumption $\mathbf{NP} \neq \mathbf{DTIME}(n^{\log\log n})$. Since we rule out $(1-\epsilon)\ln n$ integrality gaps for $\Omega(n)$ rounds of LS_+, this strengthens the possibility that stronger PCP results are possible for this problem. In particular, it further supports the conjecture that $(1-\epsilon)\ln n$ integrality gaps ought to be ruled out under the weaker assumption of $\mathbf{NP} \neq \mathbf{ZPP}$ or even $\mathbf{NP} \neq \mathbf{P}$.

9 Acknowledgements

The authors would like to thank Toniann Pitassi and Joshua Buresh-Oppenheim for suggesting the problem of proving integrality gaps for MAX-3SAT as well as for many helpful discussions that aided in the positive resolution of this question.

The third author would like to thank Shlomo Hoory, Avner Magen and Toniann Pitassi for suggestions on how to simplify the exposition.

References

[1] M. Alekhnovich and E. Ben-Sasson. Linear upper bounds for random walk on small density random 3-cnf. In *Proceedings of the 44th Symposium on Foundations of, Computer Science (FOCS-03)*, pages 352–361, 2003.

[2] S. Arora, B. Bollobás, and L. Lovász. Proving integrality gaps without knowing the linear program. In *Proceedings of the 43rd Symposium on Foundations of Computer Science (FOCS-02)*, pages 313–322, Los Alamitos, Nov. 16–19 2002.

[3] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. In *Proceedings of the thirty-sixth annual ACM Symposium on Theory of Computing (STOC-04)*, pages 222–231, New York, June 13–15 2004. ACM Press.

[4] J. Buresh-Oppenheim, N. Galesi, S. Hoory, A. Magen, and T. Pitassi. Rank bounds and integrality gaps for cutting planes procedures. In *FOCS: IEEE Symposium on Foundations of Computer Science (FOCS)*, 2003.

[5] I. Dinur, V. Guruswami, S. Khot, and O. Regev. A new multilayered PCP and the hardness of hypergraph vertex cover. In ACM, editor, *Proceedings of the Thirty-Fifth ACM Symposium on Theory of Computing, San Diego, CA, USA, June 9–11, 2003*, pages 595–601, New York, NY, USA, 2003. ACM Press.

[6] I. Dinur and S. Safra. The importance of being biased. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC-02)*, pages 33–42, New York, May 19–21 2002. ACM Press.

[7] Feige and Krauthgamer. The probable value of the Lovasz-Schrijver relaxations for maximum independent set. *SICOMP: SIAM Journal on Computing*, 32, 2003.

[8] U. Feige. A threshold of ln n for approximating set cover. In *Proceedings of The Twenty-Eighth Annual ACM Symposium On The Theory Of Computing (STOC '96)*, pages 314–318, New York, USA, May 1996. ACM Press.

[9] M. X. Goemans and L. Tunçel. When does the positive semidefiniteness constraint help in lifting procedures. *Mathematics of Operations Research*, 26:796–815, 2001.

[10] M. X. Goemans and D. P. Williamson. .878-approximation algorithms for MAX CUT and MAX 2SAT. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on the Theory of Computing*, pages 422–431, Montréal, Québec, Canada, 23–25 May 1994.

[11] J. Håstad. Some optimal inapproximability results. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 1–10, El Paso, Texas, 4–6 May 1997.

[12] 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.

[13] R. Raz and S. Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 475–484, El Paso, Texas, 4–6 May 1997.

[14] 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.

Appendix

Proof that protection lemma matrices are positive semidefinite

Proposition 9.1. *The matrix Y constructed in Lemma 4.6 is positive semi-definite (PSD).*

Proof. By definition, Y equals the matrix $yy^T + \mathrm{Diag}(y - y^2)$ except $Y_{ij} = 0$ if $i \sim_C (x) j$. By Fact 1 on equivalences, there exist disjoint sets $I_1, \ldots, I_t$ of indices such that (a) $|I_j| = k - 1$ for all $j \in [t]$, (b) all indices belonging to an equivalence are in one of the I_j, and (c) for each $j \in [t]$ all indices in I_j are mutually equivalent. Then,

$$Y = yy^T + \mathrm{Diag}(y - y^2) + \sum_{j \in [t]} \left(\mathrm{Diag}(y_{I_j}^2) - y_{I_j} y_{I_j}^T \right),$$

where y_I equals y but is zero outside I.

To show Y is PSD, we show that $z^T Y z \geq 0$ for all $z \in \mathbb{R}^{n+1}$. Note that $z^T(yy^T)z = (y^T z)^2 \geq 0$ for all $z \in \mathbb{R}^{n+1}$. Moreover, $\mathrm{Diag}(w)$ is PSD for any w such that $w_j \geq 0$ for all j. Hence, since the sum of PSD matrices is PSD, to show that Y is PSD it suffices to show for each I_j that the following quantity is non-negative:

$$z^T(\mathrm{Diag}(y_{I_j} - y_{I_j}^2) + \mathrm{Diag}(y_{I_j}^2) - y_{I_j} y_{I_j}^T)z = z^T(\mathrm{Diag}(y_{I_j}) - y_{I_j} y_{I_j}^T)z.$$

Since the argument is identical for all I_j we drop the subscript j and assume $I = [k-1]$. The above then simplifies to $\sum_{i \in [k-1]}(z_i^2 x_i) - (\sum_{i \in [k-1]} z_i x_i)^2$. Since $x_i = \frac{1}{k-1}$ for all indices in an equivalence, this further simplifies to $\frac{1}{k-1}\sum_{i \in [k-1]} z_i^2 - \frac{1}{(k-1)^2}(\sum_{i \in [k-1]} z_i)^2$. The latter is non-negative by the fact that $\sum_{i \in [\ell]} a_i^2 \geq \frac{1}{\ell}(\sum_{i \in [\ell]} a_i)^2$. □

Graph Theory Lemmas

Lemma 9.2. *Let $\Delta(\epsilon, k, \beta) = \left(\frac{e^{\epsilon-k}}{5\beta(k-1-\epsilon)^{1+\epsilon}}\right)^{1/\epsilon}$. Then for all $\alpha, 0 < \alpha < 1$, and all $\epsilon > 0$, there exists $\mu(\alpha)$ such that for all $\beta \geq \mu(\alpha)\alpha^{-k}$ and all δ, $0 < \delta < \Delta(\epsilon, k, \beta)$, the probability that a random k-uniform hypergraph $G = (V, E)$ on n vertices with βn hyperedges (1) has no vertex cover of size smaller than $(1-\alpha)n$ and (2) H_G is a $(\delta n, k-1-\epsilon)$ expander is at least $1/2$.*

Proof. Let $\beta = \mu(\alpha)\alpha^{-k}$ and suppose the hypergraph has βn randomly and uniformly chosen hyperedges where $\mu(\alpha)$ is chosen below. The probability that there exists a vertex cover of size $(1-\alpha)n$ equals the probability that there exists a set $S \subseteq V$, $|S| = \alpha n$, such that no edge contains only elements from S. This probability is bounded by

$$\binom{n}{\alpha n}(1-\alpha^k)^{\beta n} \leq \left(\frac{e}{\alpha}\right)^{\alpha n}(1-\alpha^k)^{\beta n} = \left(\frac{e}{\alpha}\right)^{\alpha n}\left(\frac{1}{e}\right)^{\mu(\alpha)n}. \tag{9.1}$$

Let $\mu(\alpha) > 0$ be such that the above is less than $1/4$.

Now consider the bipartite graph H_G mapping E to V. Note that $|E| = \beta n$. The probability that a subset of $s = \delta n$ constraints of F does not have expansion more than $c = k-1-\epsilon$ is

$$\binom{\beta n}{s}\binom{n}{cs}\left(\frac{cs}{n}\right)^{ks} \leq \left(\frac{e\beta n}{s}\right)^{s}\left(\frac{en}{cs}\right)^{cs}\left(\frac{cs}{n}\right)^{ks} = \left[\delta^{\epsilon}\beta e^{k-\epsilon}c^{1+\epsilon}\right]^{s}.$$

Let $r = \delta^{\epsilon}\beta e^{k-\epsilon}c^{1+\epsilon}$. Then $r < 1/5$ when $\delta < \left(\frac{e^{\epsilon-k}}{5\beta c^{1+\epsilon}}\right)^{1/\epsilon}$. Hence, the probability that some subset of E of size at most δn fails to have expansion greater than $k-1-\epsilon$ is bounded by

$$\sum_{s=1}^{\delta n} r^s \leq \sum_{s \geq 1} r^s = \frac{r}{1-r} < \frac{1}{4}.$$

So with probability at least $1/2$, both G has no vertex cover of size less than $(1-\alpha)n$ and H_G is a $(\delta n, k-1-\epsilon)$ expander. □

Lemma 9.3. *For any constant $\epsilon, \delta \in (0,1)$ for all n there exists an $(\epsilon-\delta)n$-regular hypergraph with n vertices and n edges that has vertex cover greater than $\log_{(1+\epsilon)} n$.*

Proof. Let $\epsilon' = \epsilon - \delta/2$. Consider a random hypergraph G with n edges over n vertices in which every vertex belongs to an edge independently with probability ϵ'. Let $k = \log_{1+\epsilon} n$. The probability that G contains a vertex cover of size k is less than or equal

$$\binom{n}{k} \cdot \left[1-(1-\epsilon')^k\right]^m \leq n^k e^{-m\cdot(1-\epsilon')^k} = o(1).$$

Finally, with high probability every edge in G contains at least $(\epsilon' - \delta/2)n = (\epsilon-\delta)n$ elements, we may throw away any vertices from each edge so that it contains exactly $(\epsilon-\delta)n$ of them. □

Lemma 9.4. *Let $H = (V_1, V_2, E)$ be a bipartite graph and let $S \subseteq V_1$ be such that for for all $S' \subseteq S$, $|\Gamma(S')| > k|S'|$. Assume $S = \{e_1, e_2, ..., e_\ell\}$. Then there exists a mapping $\eta : S \to \mathcal{P}(\Gamma(S))$ such that (1) for all $i \in [\ell]$, $|\eta(e_i)| = k + 1$, and (2) for all $i \in [\ell]$, $|\eta(e_i) \setminus \bigcup_{j<i} \eta(e_j)| \geq k$.*

Proof. By the generalization of Hall's theorem there exists a k-matching from S into $\Gamma(S)$. Fix such a k-matching ν once and for all. We construct η in the following recursive way. By assumption, $\Gamma(S)$ contains at least $\ell k+1$ elements. So by the pigeon-hole principle there exists a vertex $v \in \Gamma(S)$ which does not belong to $\bigcup_{e \in S} \nu(e)$. Consider any vertex $e \in S$ that is adjacent to v (such a point exists because $v \in \Gamma(S)$) and let $\eta(e) = \{v\} \cup \nu(e)$. Finally, denote $S' = S \setminus \{e\}$ and repeat the process recursively for S'. The vertices in S' are ordered according to the way they were ordered in S.

Clearly for all vertices e_i in S, $\eta(e_i)$ is a $k + 1$ element subset of $\Gamma(S)$. To check the second required property for η note that at each step of the inductive process, no vertex of $\nu(e)$ may be joined to any of the $\eta(e')$ from earlier steps, because $\eta(e')$ consists of $\nu(e')$ and v', $v' \notin \nu(e)$. The lemma follows. $\square$