Proving integrality gaps without knowing the linear program

Sanjeev Arora* Princeton University

Béla Bollobás †‡§ University of Memphis

László Lovász¶ Microsoft Research

Abstract

Proving integrality gaps for linear relaxations of NP optimization problems is a difficult task and usually undertaken on a case-by-case basis. We initiate a more systematic approach. We prove an integrality gap of $2-o(1)$ *for three families of linear relaxations for vertex cover, and our methods seem relevant to other problems as well.*

1 Introduction

Approximation algorithms for NP-hard problems — metric TSP, Vertex Cover, graph expansion, cut problems, etc.— often use a linear relaxation of the problem (see Vazirani [19], Hochbaum [13]). For instance, a simple 2-approximation algorithm for *vertex cover* solves the following relaxation: minimize $\sum_{i\in V} x_i$ such that $x_i + x_j \geq 1 \forall \{i,j\} \in E$. One can show that in the optimum solution, $x_i \in \{0, 1/2, 1\}$. Thus rounding the 1/2's up to 1 gives a vertex cover [15]. This also proves an upper bound of 2 on the *integrality gap* of the relaxation, which is the maximum over all graphs G, of the ratio of the size of minimum vertex cover in G and the cost of optimum fractional solution. Can we write a linear relaxation with a lower integrality gap, say 1.5? Note that the LP need not even be of polynomial size so long as it comes with a polynomial time separation oracle, which is all we need to solve it with the Ellipsoid method.

Such quests for tighter relaxations can seem neverending, since even simple modifications could conceivably tighten the relaxation.

For certain problems, though, the quest for tighter relaxations —indeed, the quest for any better approximation algorithms— has ended. Results using probabilistically checkable proofs (PCPs) show that for a variety of problems such as MAX-3SAT, Set Cover, MAX-2SAT, etc., known approximation algorithms cannot be improved if P $\neq$ NP. Thus PCP-based techniques provide an explanation for our inability to provide tighter relaxations for these problems.

However, for many other problems including all four problems mentioned in the opening paragraph, the PCP-based results are fairly weak or nonexistent, and fall well below the integrality gaps of the best relaxations. The best hardness result for Vertex Cover —due to Dinur and Safra [6], who improved upon a long line of work— only shows that 1.36-approximation is NP-hard. The best hardness result for metric TSP only shows that 1.01-approximation is NP-hard [16], yet decades of work has failed to yield a relaxation with integrality gap better than 1.5 [20] (or 4/3, if one believes a well-known conjecture [10]). For graph expansion and related graph problems essentially no hardness results exist yet we only know relaxations with integrality gap $\Omega(\log n)$ (Shmoys [18]).

When decades of work has failed to turn up tighter relaxations, one should seriously investigate the possibility that *no tighter relaxations exist*. Proving such a statement may, however, be related to P vs NP, since linear programming is complete for P[1] Thus it seems necessary to work with sub-

*Supported by David and Lucile Packard Fellowship, NSF Grant CCR-0098180, and an NSF ITR grant. Work done while visiting Microsoft Research and the CS Dept at UC Berkeley. arora@cs.princeton.edu

†Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399

‡Department of Mathematical Sciences, University of Memphis, Memphis TN 38152. Research supported by NSF grant DSM 9971788 and DARPA grant F33615-01-C-1900 bollobab@mathsci.msci.memphis.edu

§Trinity College, Cambridge CB2 1TQ, UK

¶Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399 lovasz@microsoft.com

[1]**Erratum:** The conference version of this paper erroneously stated that proving such a statement is tantamount to proving P $\neq$ NP. However, it is actually an open prob-

families of linear relaxations. An integrality gap result for a large subfamily of relaxations may then be viewed as a lowerbound for a restricted computational model, analogous say to lowerbounds for monotone circuits [17] and for proof systems [4]. An example is Yannakakis's result [21] that representing TSP (the exact version) using a symmetric linear program requires exponentially many constraints.

In this paper we prove nonexistence of tighter relaxations for Vertex Cover among three fairly general families of LPs. For all families we prove an integrality gap $2 - o(1)$. (Since the complement of a vertex cover is an independent set, and vice versa, the results may also be trivially rephrased to say that the integrality gap of the independent set problem is unbounded even in graphs whose independent set number is $\Omega(n)$.) An interesting aspect of our result —also the reason for the paper's title— is that no explicit description is known for the LPs in the three families. However, we can show that they use inequalities that have a fairly local view of the graph. This lets us construct graphs in which the minimum vertex cover must contain almost all the vertices, yet the all-1/2 solution (or something close to it) is feasible for each inequality.

In the first two families of relaxations we allow only the variables $x_1, x_2, \ldots, x_n \in [0, 1]$ for the vertices and no auxiliary variables. Such a restriction seems necessary because auxiliary variables would give the LP the power of arbitrary polynomial-time computations. The third family allows auxiliary variables implicitly, but in a very controlled way (namely, as part of the lift and project procedure of Lovász and Schrijver [14]).

The first family consists of linear programs that can include *arbitrary* inequalities on any set of ϵn variables.

The second family consists of linear programs containing equations with low *defect*. For an inequality $a^T \cdot x \leq b$, where a is an integer vector, the defect is defined to be $\sum_i a_i - 2b$. The defect of facets in the independent set polytope varies from 1 to $n-2$, and we allow inequalities of defect at most ϵn. An integrality gap of $2 - o(1)$ for this family is a simple corollary of the one for the first family.

The third family consists of linear programs obtained from $O(1)$ rounds of a "lift-and-project" construction of Lovász and Schrijver [14]. This is a method that, over many rounds, generates tighter and tighter linear relaxations for 0/1 optimization problems. It is more round-efficient than classical cutting plane procedures such as Gomory-Chvatal [5], since it generates every valid inequality in at most n rounds. Even in one round it generates nontrivial inequalities for vertex cover. Furthermore, the set of inequalities derivable in $O(1)$ rounds —this could be an exponentially large set— has a polynomial-time separation oracle, thus allowing the Ellipsoid method to optimize over this set. Characterizing the set of inequalities obtained in $O(1)$ rounds has proved difficult; even the case of 2 rounds is open. (Our integrality gap of $2 - o(1)$ actually extends to $\Omega(\sqrt{\log n})$ rounds.)

Our techniques seem applicable to problems other than vertex cover and should be the subject of future work. Extending our ideas to semidefinite relaxations as well as to the semidefinite programming analog of the Lovász-Schrijver procedure is an interesting open problem.

The recent paper [9] by Feige and Krauthgamer proves bounds on how well the iterated lifting procedure, with semidefiniteness constraints included, approximates the maximum clique on a random graph with edge probabilities 1/2. However, this result does not give directly any lower bound on the approximability of vertex cover, since in this case both the vertex cover and the relaxations considered are about n.

2 The first family

In this section we prove integrality gaps for linear programs in $\{x_1, x_2, \ldots, x_n\}$ that allow any constraint of the form $a^T \cdot x \leq b$, where the coefficient vector a is nonzero for only ϵn coordinates. In other words, each constraint only involves ϵn variables. The linear program may have exponential size, and may not have a polynomial-time separation oracle. (In fact, there are linear programs in this family for which finding such an oracle would imply P = NP.) We only require that all 0/1 vertex covers in the graph are feasible for the relaxation.

To exhibit the integrality gap we will construct a special graph by the probabilistic method (Theorem 1). This result appears to be new, although it fits in a line of results starting with Erdős [8], show-

lem [21] whether P = NP implies that the Vertex Cover polytope has a polynomial size description (where additional variables are allowed).

ing that the chromatic number of a graph cannot be deduced from "local considerations" (see also Alon and Spencer [1], p. 130).

A *fractional γ-coloring* of a graph G is a multiset $\mathcal{C} = \{U_1, \ldots, U_N\}$ of independent sets of vertices such that every vertex is in at least N/γ members of $\mathcal{C}$. The *fractional chromatic number* of G is

$$\chi_f(G) = \inf\{\gamma : \ G \text{ has a fractional } \gamma - \text{coloring}\} \ .$$

Theorem 1 *Let $0 < \alpha, \delta < 1/2$ be constants. Then there are constants $\beta = \beta(\alpha, \delta) > 0$ and $n_0 = n_0(\alpha, \beta, \delta)$ such that for every $n \geq n_0$ there is a graph with n vertices and independence number at most αn such that every subgraph induced by a subset of at most βn vertices has fractional chromatic number at most $2 + \delta$.*

Let H be the graph constructed in Theorem 1 with α, δ arbitrarily close to 0 and β be as given by the theorem.

Theorem 2 *The vector with all coordinates $\frac{1+\delta}{2+\delta}$ is feasible for any linear relaxation for H in which each constraint involves at most βn variables. (Consequently, the integrality gap is $\geq (1-\alpha) \cdot \frac{2+\delta}{1+\delta}$.)*

PROOF: Suppose the linear relaxation contains, for each $I \subseteq \{1, \ldots, n\}$ of size βn, the constraints

$$A_I \cdot x \leq b_I \tag{1}$$

where A_I is an arbitrary constraint matrix with nonzero entries only in columns corresponding to I. (We emphasize that A_I may depend upon the input graph H.) Then the polytope P_I associated with these constraints is the cross-product of some polytope Q_I on the coordinates in I and the unit cube $[0,1]^{n-|I|}$ in the other coordinates. The polytope describing the entire relaxation is $\cap_I P_I$.

We show the all-$\frac{1+\delta}{2+\delta}$ vector lies in all P_I's and hence is feasible. Let I be any subset of at most βn vertices and $\{U_1, \ldots, U_N\}$ be its fractional $(2+\delta)$-coloring. Each vertex in I is contained in at least $1/(2+\delta)$ fraction of U_i's, but we can assume without loss of generality —by deleting the vertex from a few U_i's if need be—that it is contained in exactly $1/(2+\delta)$ fraction of U_i's. Notice, each $I \setminus U_i$ is a vertex cover in the subgraph induced by I and hence can be extended to a vertex cover of the entire graph. Thus any vector in $\Re^n$ that has $1_{I \setminus U_i}$ (the characteristic vector of $I \setminus U_i$) in the coordinates corresponding to I is feasible for $A_I \cdot x \leq b_I$.

Consider the vectors $v_1, v_2, \ldots, v_N \in \Re^n$ where v_i is $1_{I \setminus U_i}$ in the coordinates corresponding to I and $(1+\delta)/(2+\delta)$ otherwise. Each such vector satisfies $A_I \cdot x \leq b_I$, so convexity implies that the same is also true for the average vector $\frac{1}{N}(v_1 + v_2 + \cdots + v_N)$. Since each vertex in I lies in exactly $1 - 1/(2+\delta) = (1+\delta)/(2+\delta)$ fraction of the vertex covers, this average is the vector with all coordinates $(1+\delta)/(2+\delta)$. Thus we have shown that the all-$\frac{1+\delta}{2+\delta}$ vector is feasible for the entire LP. □

Note that the same construction also proves that the integrality gap for independent set is unbounded: for any $\epsilon, \delta > 0$ there are graphs with independence number αn and for which every linear relaxation in our family has integrality gap at least $1/\alpha(2+\delta)$.

2.1 Proof of Theorem 1

The proof uses standard random graph theory supplemented with a couple of new ideas. Let us recall the standard part (see [1]). If we pick a random graph G using the familiar $G(n,p)$ model and choose p appropriately, then in G the largest independence set has size at most αn and yet the induced subgraph on every subset of βn vertices has an independence set of size close to $\beta n/2$. By deleting a few edges —too few to disturb anything else — we can assume that G has no small cycle (i.e., has high girth).

To prove an upperbound of $2+\delta$ on the fractional chromatic number of every induced subgraph with at most βn vertices, we use induction on the subgraph size. The main idea in the inductive step is to exhibit a long path inside every subgraph. (To do this we need a sparsity condition in addition to high girth; see Lemma 6.) Peeling away the path gives a smaller subgraph that is colored (fractionally) using the inductive assumption. Extending this fractional coloring to the path vertices is easy (see Lemma 5), and this completes the induction (see Lemma 7).

Now we give details. The next Lemma concerns the "standard random graph theory" mentioned above, together with a sparsity condition, which appears to be new.

Lemma 3 *Given an integer $g \geq 3$ and real numbers α, η with $0 < \alpha < 1/2$ and $0 < \eta < 1/2$, let $\lambda > e^2$ and $\beta > 0$ be such that*

$$2\frac{\log \lambda}{\lambda} \leq \alpha \tag{2}$$

and

$$\beta < (e\lambda)^{-2/\eta}. \tag{3}$$

Then there is an integer $n_0 = n_0(\lambda, \eta, g)$ such that for every $n \geq n_0$ there is a graph H of order n, girth at least g and independence number at most αn such that every subgraph of H with $\ell \leq \beta n$ vertices contains at most $(1+\eta)\ell$ edges.

Remark. Condition (2) is satisfied if we take

$$\lambda = (3/\alpha)\log(1/\alpha).$$

PROOF: Let us consider the space of random graphs $\mathcal{G}(n,p)$ with $p = \lambda/n$. Let $0 < \alpha_0 < \alpha$ be such that

$$1 + \log(l/\alpha_0) < \lambda\alpha_0/2. \tag{4}$$

Inequality (2) implies that we can choose such an α_0.

In order to avoid unnecessary clutter, in what follows, we shall drop the integrality signs (in particular, we shall write $\alpha_0 n$ instead of $\lceil \alpha_0 n \rceil$); this slight inaccuracy will not endanger the validity of the arguments. Also, as usual, we shall assume that n is large enough to make our inequalities hold.

(*i*) The probability that for some ℓ, $4 \leq \ell \leq 1/\eta$, some ℓ-set spans at least $\ell+1$ edges is at most

$$\sum_{\ell=4}^{\lfloor 1/\eta \rfloor} \binom{n}{\ell}\binom{\binom{\ell}{2}}{\ell+1}\left(\frac{\lambda}{n}\right)^{\ell+1} = O(n^{-1}).$$

Also, very crudely, for $1/\eta < \ell < \beta n$, the expected number of ℓ-sets spanning at least $(1+\eta)\ell$ edges is less than

$$\begin{aligned}
&\binom{n}{\ell}\binom{\binom{\ell}{2}}{(1+\eta)\ell}\left(\frac{\lambda}{n}\right)^{(1+\eta)\ell} \\
&\leq \left[\left(\frac{en}{\ell}\right)\left(\frac{e\ell}{2(1+\eta)}\right)^{1+\eta}\left(\frac{\lambda}{n}\right)^{1+\eta}\right]^{\ell} \\
&\leq \left[e^2(\ell/n)^{\eta}\lambda^{1+\eta}\right]^{\ell} = S_\ell.
\end{aligned}$$

Since $e^2\beta^\eta\lambda^2 < 1$,

$$\sum_{\ell=\lceil 1/\eta \rceil}^{\lfloor \beta n \rfloor} S_\ell = O(n^{-1}).$$

Hence, *with probability $1 - O(n^{-1/2})$, every set of $\ell \leq \beta n$ vertices spans at most $(1+\eta)\ell$ edges.*

(*ii*) Let $I = I(G_{n,p})$ be the number of independent sets of $\lceil \alpha_0 n \rceil$ vertices. Note that

$$\begin{aligned}
\mathbb{E}(I) &= \binom{n}{\alpha_0 n}\left(1 - \frac{\lambda}{n}\right)^{\binom{\alpha n}{2}} \\
&\leq \left(\left(\frac{e}{\alpha_0}\right)e^{-\lambda\alpha_0/2}\right)^{\alpha_0 n} = \gamma_0^{\alpha_0 n}
\end{aligned}$$

Inequality (4) implies that $\gamma_0 < 1$, so the probability that $G_{n,p}$ contains an independent set of $\alpha_0 n$ vertices is exponentially small.

(*iii*) Call a cycle in $G_{n,p}$ *short* if its length is less than g. The expected number of short cycles is less than

$$\sum_{\ell=3}^{g-1} \frac{n^\ell}{l}\left(\frac{\lambda}{n}\right)^{\ell} = \sum_{\ell=3}^{g-1}\frac{\lambda^\ell}{\ell} = s.$$

Therefore, with probability $1 - O(n^{-1/2})$, $G_{n,p}$ has at most $n^{1/2}$ short cycles. Deleting an edge of each of these cycles, we get a graph of girth at least g.

Consequently, with probability $1 - O(n^{-1/2})$, a random graph $G_{n,p}$ is such that no set of $\ell \leq \beta n$ vertices spans more than $(1+\eta)\ell$ edges, and if we delete an edge from each short cycle then the independence number of the new graph $H = G'_{n,p}$ is at most

$$\alpha_0 n + \sqrt{n} < \alpha n.$$

This graph H has the required properties. □

Now we establish some basic properties of χ_f. Note that if G has a k-coloring with color classes $U_1, \ldots, U_k$ then $\mathcal{C} = \{U_1, \ldots, U_k\}$ is also a fractional k-coloring of G. Consequently, $\chi_f(G) \leq \chi(G)$. It is easy to check that a path with at least one edge has fractional chromatic number 2. The proof of the next lemma is left to the reader.

Lemma 4 *1. If C_k denotes the cycle of length k then $\chi_f(C_{2\ell}) = 2$ and $\chi_f(C_{2\ell}+1) = (2\ell+1)/\ell$.*

2. If $|V(G_1) \cap V(G_2)| \leq 1$ then $\chi_f(G_1 \cup G_2) = \max\{\ \chi_f(G_1),\ \chi_f(G_2)\ \}$.

The next Lemma concerns χ_f of a graph that contains a long path (the vertices on the path's interior have no edges outside the path edges).

Lemma 5 *Let $\ell \geq 2$ and let G be a graph obtained by adding a path $x_0x_1\ldots x_{\ell+1}$ to a graph G', where $x_0, x_{\ell+1} \in V(G')$ and $x_i \notin V(G')$ for $1 \leq i \leq \ell$. Then $\chi_f(G) \leq \max\{\chi_f(G'), \frac{2\ell}{\ell-1}\}$.*

PROOF: Let $\mathcal{C}' = \{U'_1, \ldots, U'_N\}$ be a multiset of independent sets in G' such that every vertex of G' is in at least γN of these sets. Let $\mathcal{C}'_1, \ldots \mathcal{C}'_\ell$ be copies of the multiset $2\mathcal{C}' = \{U'_1, U'_1, U'_2, U'_2, \ldots, U'_N, U'_N\}$. For $1 \leq i \leq \ell$ and $j \neq i$, let $\mathcal{C}_j$ be obtained from $\mathcal{C}'_j$ by extending each $U'_h \in \mathcal{C}'_j$ to an independent set U_h in $G \setminus \{x_j\}$ (i.e., the graph with x_j removed) in the trivial way. Thus every x_k, $k \neq j$, is in at least N sets of $\mathcal{C}_j$. Then $2\mathcal{C} = \cup_{j=1}^{\ell} \mathcal{C}_j$ is a multiset of $2\ell N$ independent sets of G such that every vertex of G' is in at least $2\ell\gamma N$ of these sets and every vertex x_i, $1 \leq i \leq \ell$, is in at least $(\ell - 1)N$ of these sets. Consequently, every vertex of G is in at least

$$\min\{\gamma, \frac{\ell - 1}{2\ell}\}$$

proportion of the sets $2\mathcal{C}$. This shows that $\chi_f(G) \leq \max\{\frac{1}{\gamma}, \frac{2\ell}{\ell-1}\}$, completing the proof. □

For a positive real $k > 1$, call a graph *k-sparse* if it has no subgraph with ℓ vertices and more than $k\ell/2$ edges. This concept is closely related to that of degeneracy: if k is a natural number and $\epsilon > 0$ then a $(\frac{k+1}{2} - \epsilon)$-sparse graph is k-degenerate; conversely, every k-degenerate graph is $2k$-sparse.

Lemma 6 *Let $\ell \geq 1$ be an integer and $0 < \eta < \frac{1}{3\ell-1}$, and let G be a 2-connected $(1+\eta)$-sparse graph which is not a cycle. Then G contains a path of length $\ell+1$ whose internal vertices have degree 2 in G.*

PROOF: Suppose that G has n vertices and does not contain a path of length $\ell+1$ with ℓ internal vertices of degree 2 in G. As every 2-connected graph with more edges than vertices, G is made up of a certain $k \geq 2$ number of branch-vertices (i.e., vertices of degree at least 3) and induced paths joining them, say $P_1, \ldots, P_m$, where $m \geq \lceil 3k/2 \rceil$. Write $\ell_i \leq \ell$ for the length of P_i. Then

$$1 + \eta \geq \frac{e(G)}{n} = \frac{\sum_{i=1}^{m} \ell_i}{k + \sum_{i=1}^{m} (\ell_i - 1)} \geq \frac{\ell}{\frac{k}{m} + \ell - 1} \geq 1 + \frac{1}{3\ell - 1},$$

contradicting our assumption that $0 < \eta < \frac{1}{3\ell-1}$. □

Lemma 7 *Let $h \geq 2$ be an integer and $0 < \eta < \frac{1}{3h+2}$. Then every $(1+\eta)$-sparse graph G of girth at least $2h$ has $\chi_f(G) \leq 2 + \frac{1}{h}$.*

PROOF: We use induction on the number of vertices. The base case is trivial. Assume the statement is true when the number of vertices is at most m and G is a graph with $m+1$ vertices. If it is not 2-connected, it has a vertex v whose removal disconnects the graph and hence we can complete the inductive step using part 2 of Lemma 4. So assume G is 2-connected. If it is a cycle then its length must be at least $2h$, and hence χ_f is at most $2 + \frac{1}{h}$ by part 1 of Lemma 4. So assume G is not a cycle. But then, by Lemma 6, G contains a path of length $h+2$ whose internal vertices have degree 2 in G. Let G' be the graph obtained from G by deleting these internal vertices (together with the edges incident with them). Since, by assumption, $\chi_f(G') \leq 2 + \frac{1}{h}$, by Lemma 5 we have $\chi_f(G) = \chi_f(G') \leq 2 + \frac{1}{h}$. This completes the induction and the Lemma is proved. □

Now we can prove the main theorem.

PROOF:(Theorem 1) Set $h = \lceil 1/\delta \rceil$, $g = 2h$ and $\eta = \frac{1}{3h+3}$. Choose $\lambda > e^2$ and $\beta > 0$ to satisfy inequalities (2) and (3). Let H be a graph of order n whose existence is guaranteed by Lemma 3. Thus, H has independence number at most αn, and if G is a subgraph of H with at most βn edges then G is $(1+\eta)$-sparse and has girth at least $g = 2h$. Hence, by Lemma 7, $\chi_f(G) \leq 2 + \frac{1}{h} \leq 1 + \delta$, completing the proof. □

3 The second family

The *defect* of an inequality $a^T \cdot x \geq b$, where b is an integer vector, is $2b - \sum_i a_i$. This number is always non-negative [14]. We generalize the results in Section 2 to families of relaxations containing inequalities of defect at most ϵn.

Since adding two inequalities gives a new inequality whose defect is the sum of their defects, it suffices in any linear system to look at the defect of the facets. In a relaxation for vertex cover set, these facets must be of the type $a^T \cdot x \geq b$, where $a \in Z_+^n, b \in Z_+$.

Let H be the graph constructed in Theorem 1 for $\alpha, \delta > 0$ arbitrary constants. Let the defect of some relaxation be at most ϵn where $\epsilon = \beta\delta/(2+\delta)$. We show that the vector x_δ with all coordinates $(1+\delta)/(2+\delta)$ is feasible.

There are two types of facets. If $\sum_i a_i \leq \beta n$ then the constraint only involves βn variables and so the feasibility of the vector x_δ follows as in Section 2. If $\sum_i a_i > \beta n$ then the feasibility of this vector follows by direct substitution:

$$\sum_i a_i \frac{1+\delta}{2+\delta} = \sum_i a_i \frac{\delta}{2(2+\delta)} + \frac{1}{2}\sum_i a_i \quad (5)$$

$$> \frac{\beta n \delta}{2(2+\delta)} + \frac{1}{2}\sum_i a_i \quad (6)$$

$$= \frac{\epsilon n}{2} + \frac{1}{2}\sum_i a_i \geq b. \quad (7)$$

4 The third family

Lovász and Schrijver [14] present a general construction for obtaining tighter relaxations for a 0/1 optimization starting from an arbitrary relaxation. The idea is to "lift" to n^2 dimensions and then project back to the n-space. We illustrate using the standard relaxation for Vertex Cover.

$$x_i + x_j \geq 1 \qquad \forall \{i,j\} \in E \qquad (\textit{Edge constraint}) \quad (8)$$

In this relaxation x_i's are real numbers in $[0,1]$, and we wish to tighten the relaxation to force the the x_i's to $0,1$. To this end, we can introduce any constraints satisfied by 0/1 vertex covers. For instance, the x_i's can be required to satisfy, for every odd-cycle C,

$$\sum_{i \in C} x_i \geq \frac{|C|+1}{2} \qquad (\textit{Odd-cycle constraint}) \quad (9)$$

Many other families of inequalities are known, but a complete listing will probably never be found because of complexity reasons. Lovász and Schrijver give an automatic method to generate all valid inequalities.

Notice, 0/1-valued vertex covers satisfy

$$x_i^2 = x_i \qquad \forall i \quad (10)$$

$$(1-x_i)(1-x_j) = 0 \qquad \forall \{i,j\} \in E \quad (11)$$

We can derive inequalities starting from only (10) and (11) and using the following two rules to eliminate nonlinear terms:

$$\alpha_1, \ldots, \alpha_m \geq 0 \Rightarrow c_0 + \sum_{i=1}^{k} c_i \alpha_i \geq 0 \ \ \forall c_0, \ldots, c_m \geq 0$$

(Linear combination)

$$\alpha > 0 \Rightarrow x_i \alpha \geq 0, \ \ (1-x_i)\alpha \geq 0 \ \forall i$$

(Restricted multiplication)

It can be shown that every valid inequality is generated in at most n rounds. The set of inequalities derivable in one round are exactly the odd-cycle inequalities, but no exact characterization exists for subsequent rounds. However, we do know that the set of inequalities derivable in $O(1)$ rounds has a polynomial-time separation oracle. For details see [14].

To understand our results, however, the reader only needs to know the next Lemma, taken from [14]. The notation uses homogenized inequalities. Let FRAC(G) be the cone in $\Re^{n+1}$ that contains a vector $(x_0, x_1, \ldots, x_n)$ iff it satisfies the edge constraints $x_i + x_j \geq x_0$ for each edge $\{i,j\} \in G$. (All cones below will be in $\Re^{n+1}$ and we're interested in the slice cut out by the hyperplane $x_0 = 1$.) Denote by $N^r(\text{FRAC}(G))$ the feasible cone of all inequalities obtained from r rounds of the [LS] lifting procedure. The lemma defines the effect of one round.

Lemma 8 ([14]) *If K is a cone in $\Re^{n+1}$, then $\overline{\alpha} \in N(K)$ iff there is an $(n+1) \times (n+1)$ symmetric matrix Y satisfying (notation: e_i is the ith unit vector and hence Ye_i is the ith column of Y)*

1. *$Ye_0 = diag(Y) = \overline{\alpha}$.*
2. *For each i, both Ye_i and $Y(e_0 - e_i)$ are in K.*

Let y_γ denote the vector $(1, \frac{1}{2}+\gamma, \frac{1}{2}+\gamma, \ldots, \frac{1}{2}+\gamma)$.

Theorem 9 *The vector y_γ is in $N^r(\text{FRAC}(G))$ provided girth$(G) \geq \frac{4r^2}{\gamma}$.*

The following result is essentially due to Erdős [7]; see Bollobás [3], Theorem 4, Ch VII.

Theorem 10 *For any $\alpha > 0$ there is an $n_0(\alpha)$ such that for every $n \geq n_0(\alpha)$ there are graphs on n vertices that girth at least $\log n / 3\log(1/\alpha)$ but no independent set of size greater than αn.*

In such graphs, the $1/2+\gamma$ vector is feasible for $N^r(\mathrm{FRAC}(G))$ for r up to $\Omega(\sqrt{\gamma \log n/\log(1/\alpha)})$, and hence the integrality gap is at least $2(1-\alpha)/(1+2\gamma)$.

4.1 Proof of Theorem 9

We prove the Theorem by induction, where the inductive hypothesis requires a set of vectors other than just y_γ to be in $N^m(\mathrm{FRAC}(G))$ for $m \le r$. These vectors are essentially all-$(1/2+\gamma)$, except possibly for a few vertices and small neighborhoods around those vertices, for which the vector could take arbitrary nonegative values so long as it satisfies edge constraints.

Definition 1 *Let $S \subseteq \{1,\ldots,n\}$ and R be a positive integer and $\gamma > 0$. Then a nonnegative vector $(\alpha_0, \alpha_1, \ldots, \alpha_n)$ with $\alpha_0 = 1$ is an (S,R,γ)-*vector *if the entries satisfy all edge constraints, and $\alpha_j = \frac{1}{2} + \gamma$ for each $j \notin \cup_{w\in S} Ball(w,R)$. Here $Ball(w,R)$ denotes the set of vertices within distance R of w.*

These vectors have a simple but important property:

Lemma 11 *If $\overline{\alpha}$ is an (S,R,γ)-vector, then along any path of length L that contains at least $2/\gamma$ edges whose endpoints lie outside $\cup_{w\in S} Ball(w,R)$, the sum of the α_i's along the path exceeds $\lceil \frac{L}{2} \rceil + 1$.*

PROOF: The vector $\overline{\alpha}$ satisfies all edge constraints, and for edges whose endpoints do not lie in $\cup_{w\in S} Ball(w,R)$, it oversatisfies them by 2γ. There are at least $2/\gamma$ such edges on the path. □

The following Corollary is immediate:

Corollary 12 *If $\overline{\alpha}$ is an (S,R,γ)-vector, and $girth(G) > 2(R+1)\cdot|S| + 2/\gamma$, then $\overline{\alpha}$ satisfies all odd-cycle constraints.*

Now we are ready to prove the Main Theorem.

PROOF:(Main Theorem) Let $\gamma = \rho$, $R_r = 0$ and $R_m = R_{m+1} + 2/\gamma$. We note that the girth is large enough that $R_m + 2/\gamma \le \mathrm{girth}(G)/r$ for all m.

We prove the following claim by induction; note that the Theorem is a subcase of the claim for $m = r$.

INDUCTIVE CLAIM FOR $N^m(\mathrm{FRAC}(G))$: *For every set S of $r-m$ vertices, every (S,R_m,γ)-vector is in $N^m(\mathrm{FRAC}(G))$.*

Base case $m = 0$. Trivial since the (S,R_m,γ)-vector $\overline{\alpha}$ satisfies edge constraints and hence lies in $\mathrm{FRAC}(G)$.

Proof for $m+1$ assuming truth for m: Let $\overline{\alpha}$ be an (S,R_{m+1},γ)-vector where $|S| = r-m-1$. Since the inductive hypothesis guarantees that every $(S\cup\{i\}, R_m, \gamma)$-vector is in $N^m(\mathrm{FRAC}(G))$, to show $\overline{\alpha} \in N^{m+1}(\mathrm{FRAC}(G))$ it suffices by Lemma 8 to exhibit an $(n+1)\times(n+1)$ times matrix Y that is symmetric and satisfies

A : $Ye_0 = diag(Y) = \overline{\alpha}$,

B : For each i such that $\alpha_i = 0$, we have $Ye_i = 0$. For each i such that $\alpha_i = 1$ we have $Ye_0 = Ye_i$. Otherwise Ye_i/α_i and $Y(e_0-e_i)/(1-\alpha_i)$ are $(S\cup\{i\}, R_m, \gamma)$-vectors.

We represent conditions A and B as a linear program, and show that this linear program is feasible, thus proving the existence of Y. The proof technique is similar to the proof in [LS] that $N(\mathrm{FRAC}(G))$ contains exactly those vectors that satisfy odd-cyle constraints.

The notation below assumes symmetry, namely, Y_{ij} is really $Y_{\{i,j\}}$.

The edge constraints for Ye_k/α_k and $Y(e_0-e_k)/(1-\alpha_k)$ require

$$\alpha_k \le Y_{ik} + Y_{jk} \le \alpha_k + \alpha_j + \alpha_i - 1 \qquad \forall k \in \{1,\ldots,n\}, \{i,j\} \in E \quad (12)$$

$$Y_{kk} = \alpha_k \qquad \forall k \in \{1,\ldots,n\} \quad (13)$$

Now we write constraints implying for every i that if $t \notin \cup_{w\in S\cup\{i\}} Ball(w, R_{m+1} + 2/\gamma)$ then in Ye_i/α_i and $Y(e_0-e_i)/(1-\alpha_i)$, the tth coordinate is $1/2+\gamma$.

We say a pair of vertices i,t is a *distant pair* if $t \notin \cup_{w\in S\cup\{i\}} Ball(w, R_{m+1}+2/\gamma)$. (Notice, such a t satisfies $\alpha_t = 1/2+\gamma$.) For such a pair, we require

$$Y_{it} = \alpha_i\alpha_t = \alpha_i(\frac{1}{2}+\gamma). \quad (14)$$

By definition, distant pairs have the property that every path in the graph connecting them contains at least $2/\gamma - 1$ edges whose endpoints were not in $\cup_{w\in S} Ball(w, R_{m+1})$, and hence are oversatisfied in vector $\overline{\alpha}$ by 2γ.

To show that the desired Y exists we prove the consistency of the family of constraints in (12), (13), (14). This uses a combinatorial version of Farkas' Lemma adapted from [LS].

Lemma 13 (Special case of Farkas' Lemma) *Let $H = (W, F)$ be a graph and let two values $0 \le a(u,v) \le b(u,v)$ be associated with each edge (u,v) in F. Let $U \subseteq W$ be also given, and let $\beta(u) \ge 0$ be associated with each $u \in U$. Then the linear system*

$$a(u,v) \le y_u + y_v \le b(u,v) \qquad \forall \{u,v\} \in F \quad (15)$$
$$y_u \ge 0 \qquad (u \in W) \quad (16)$$
$$y_u = \beta(u) \qquad (u \in U) \quad (17)$$

has no solution if and only if there exists a walk $v_0, v_1, \ldots, v_p$ (the walk could visit a node multiple times) in the graph H such that one of the following holds:

1. *p is odd and $b(v_0,v_1) - a(v_1,v_2) + b(v_2,v_3) - \cdots + b(v_{p-1},v_p) < 0$;*
2. *p is even, $v_0 = v_p$ and $b(v_0,v_1) - a(v_1,v_2) + b(v_2,v_3) - \cdots - a(v_{p-1},v_p) < 0$;*
3. *p is odd and $v_0, v_p \in U$ and $b(v_0,v_1) - a(v_1,v_2) + b(v_2,v_3) - \cdots + b(v_{p-1},v_p) - \beta(v_0) - \beta(v_p) < 0$;*
4. *p is even, $v_0, v_p \in U$ and $b(v_0,v_1) - a(v_1,v_2) + b(v_2,v_3) - \cdots - a(v_{p-1},v_p) - \beta(v_0) + \beta(v_p) < 0$;*
5. *p is odd, $v_0, v_p \in U$ and $-a(v_0,v_1) + b(v_1,v_2) - a(v_2,v_3) + \cdots - a(v_{p-1},v_p) + \beta(v_0) + \beta(v_p) < 0$;*
6. *p is even, $v_p \in U$ and $b(v_0,v_1) - a(v_1,v_2) + b(v_2,v_3) - \cdots - a(v_{p-1},v_p) + \beta(v_p) < 0$.*

In our system, W is the set of all pairs i,j; two pairs (i,k) and (j,k) are connected by an edge in H iff $\{i,j\} \in E$. For such a pair, $b(ij,jk) = \alpha_i + \alpha_j + \alpha_k - 1$, $a(ij,jk) = \alpha_j$. Set U consists of two types of pairs (a) (i,i) for which $\beta(i,i) = \alpha_i$. (b) (i,j) where i,j are a *distant pair*. One of them, say j, satisfies $\alpha_j = 1/2 + \gamma$, in which case $\beta(ij) = \alpha_i(1/2+\gamma)$.

To prove feasibility we have to rule out all six cases mentioned in the Lemma. The cases where both v_0, v_p are not distant pairs are ruled out as in [LS], whose proof only needs that the α_i's satisfy the odd-cycle constraints, which they do thanks to Corollary 12. So assume that at least one of v_0, v_p is a distant pair.

Following [LS], for (ik,jk) in F call $\{i,j\} \in E$ the *bracing edge*. We call a pair (i,i) a *diagonal pair*.

Assume that we have one of the cases 3,4,5. If only one of v_0, v_p is a distant pair, and the other is a diagonal pair, then the distant pair is connected by the walk formed by the bracing edges. By definition of distant pair, every path in the graph that connects them contains at least $2/\gamma - 1$ edges that are oversatisfied by 2γ in vector $\overline{\alpha}$. Then the arguments of [LS] together with Lemma 11 rule out all of the cases listed in the Lemma statement. (The reader can also verify this using the ideas below.)

So assume both v_0, v_p are distant pairs. For convenience let us pick case 3; the others are similar.

Let w be a vertex in G. Vertex w could become an element of v_i for $i \ge 1$, stay an element of subsequent v_j's for a while and then drop out. Say this is repeated $f(w)$ times. Then the contribution of α_w to the sum is $f(w)\alpha_w$. Then the sum in condition case 3 is

$$\sum_w f(w)\alpha_w + \alpha_i + \alpha_j + \alpha_r + \alpha_t - \frac{p+1}{2} - \beta(ij) - \beta(rt) \quad (18)$$

(Note that i,j,r,t may appear in the interior of the walk, and also have an $f()$ count; the extra terms are due to their appearance in the first and last edge.)

Every w that is not an element of v_0 and v_p appears in exactly $2f(w)$ bracing edges in the walk. Summing the edge constraints for all p bracing edges appearing in the walk gives

$$2\sum_w f(w)\alpha_w + \alpha_i + \alpha_j + \alpha_r + \alpha_t \ge p. \quad (19)$$

Since both ij and rt are distant pairs, assume i,r are such that $\alpha_i = \alpha_r = 1/2 + \gamma$ and that j,t are connected by a path of bracing edges. Then $\beta(ij) = \alpha_j(1/2+\gamma)$ and $\beta(rt) = \alpha_t(1/2+\gamma)$. Also, i,r must be connected by a path of bracing edges, and all edges on this path are oversatisfied by 2γ in $\overline{\alpha}$. If this path has nonzero length, then this contribution of 2γ greatly simplifies the calculations below. The nontrivial case is when the path has zero length (i.e., $i = r$) and hence the path connecting j,t has odd length.

Let D denote $\alpha_j + \alpha_t$. There are two cases: $D > 1$ and $D \le 1$. If $D > 1$ then actually we

can strengthen (19) to

$$2\sum_w f(w)\alpha_w + \alpha_i + \alpha_j + \alpha_r + \alpha_t \geq p + D - 1. \quad (20)$$

(This just says that if the two endpoints sum to $D > 1$ then some edge(s) along the path must be oversatisfied by $D-1$.) Now rewrite the sum in (18) as

$$\sum_w f(w)\alpha_w + 2 \cdot \frac{1}{2}(\alpha_i + \alpha_j + \alpha_r + \alpha_t) - \alpha_i\alpha_j - \alpha_r\alpha_t - \frac{p+1}{2} \quad (21)$$

$$\geq \frac{p+D-1}{2} + \frac{1}{2}(\alpha_i + \alpha_j + \alpha_r + \alpha_t) - \alpha_i\alpha_j - \alpha_r\alpha_t - \frac{p+1}{2} \quad (22)$$

by substituting values,

$$= \frac{D-2}{2} + \frac{1}{2}(1 + 2\gamma + \alpha_j + \alpha_t) - (\frac{1}{2} + \gamma)(\alpha_j + \alpha_t) \quad (23)$$

$$= \frac{D-1}{2} + \gamma + \frac{D}{2} - D(\frac{1}{2} + \gamma) \quad (24)$$

$$= (D-1)(\frac{1}{2} - \gamma) \qquad \geq 0 \quad (25)$$

We conclude that the sum in case 3 cannot be negative if $D > 1$. If $D \leq 1$ on the other hand, then we can also rewrite the sum in (18) using (19)

$$\sum_w f(w)\alpha_w + 2 \cdot \frac{1}{2}(\alpha_i + \alpha_j + \alpha_r + \alpha_t) - \alpha_i\alpha_j - \alpha_r\alpha_t - \frac{p+1}{2} \quad (26)$$

$$\geq \frac{p}{2} + (\frac{1}{2} + \gamma) + \frac{D}{2} - D(\frac{1}{2} + \gamma) - \frac{p+1}{2} \quad (27)$$

$$= \gamma(1 - D) \qquad \geq 0 \quad (28)$$

Thus irrespective of the value of D, the sum in case 3 cannot be negative.

Case 6, when v_p is a distant pair, can be ruled out by a similar arguments. □

5 Discussion

As mentioned earlier, the interesting open problems are to extend our techniques to problems other than vertex cover and to semidefinite relaxations instead of linear relaxations. We also feel that the lower-bound for the [LS] procedure should extend to more than $\sqrt{\log n}$ rounds but the argument seems to need some property other than high girth.

We also note that the integrality gaps proven in Section 2 are strong enough (namely, they apply to LPs that we do not know how to solve in $2^{o(n)}$ time) that they may be seen as complementary to PCP-based results. Even if it were shown using PCPs that $(2-\epsilon)$-approximation to vertex cover is NP-hard, the proof would probably involve even more complex reductions than those in [6]. Thus it might reduce 3SAT formulae of size n to vertex cover on graphs of size n^c, where c is astronomical. Even if we assume 3SAT has no $2^{o(n)}$ time algorithms, such a reduction would not rule out integrality gap of 1.1 (say) for the relaxations in Section 2.

In other words, even in a world with PCP-based results, our methods may be useful for ruling out subexponential approximation algorithms that use linear programming approaches.

Acknowledgements

The first author would like to thank Noga Alon for drawing his attention to the Erdős results of [1, 8] and their tightenings, and to Subhash Khot for useful discussions.

References

[1] N. Alon and J. Spencer. The probabilistic method (2nd ed.). John Wiley and Sons, NY, 2000.

[2] S. Arora and C. Lund. Hardness of approximations. In [13].

[3] Bollobás, B., *Modern Graph Theory*, Graduate Texts in Mathematics, vol. **184**, Springer-Verlag, New York, 1998, xiv+394 pp.

[4] M. Bonet, T. Pitassi, and R. Raz. Lower bounds for cutting planes proofs with small coeffcients. *Journal of Symbolic Logic*, 62:708–728, 1997.

[5] V. Chvatal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4:305–337, 1973.

[6] I. Dinur and S. Safra. The importance of being biased. *Proc. ACM STOC*, 2002.

[7] P. Erdős. Graph theory and probability. *Canadian J. Math.* **11**: 34–38, 1959.

[8] P. Erdős. On circuits and subgraphs of chromatic graphs. *Mathematika* **9**: 170-175, 1962.

[9] U. Feige and R. Krauthgamer. The probable value of the Lovász-Schrijver relaxations for maximum independent set. Weizmann Institute Technical Report MCS-02-01, January 2002.

[10] M. Goemans. Worst-case comparison of valid inequalities for the TSP. *Mathematical Programming*, **69**: 335–349, 1995.

[11] M.X. Goemans and L. Tuncel. When does the positive semidefiniteness constraint help in lifting procedures? *Mathematics of Operations Research*, to appear.

[12] M. Grotschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*. Springer-Verlag, 1993.

[13] D. Hochbaum, ed. *Approximation Algorithms for NP-hard problems.* PWS Publishing, Boston, 1996.

[14] L. Lovász and A. Schrijver. Cones of matrices and setfunctions, and 0-1 optimization. *SIAM Journal on Optimization*, 1:166–190, 1990.

[15] D. Hochbaum. Approximating Covering and Packing Problems: Set Cover, Vertex Cover, Independent Set, and Related Problems. In [13].

[16] C.H. Papadimitriou and S. Vempala. On the Approximability of the Traveling Salesman Problem, *Proc. 32nd ACM Symposium on Theory of Computing*, 126–133, 2000.

[17] A. A. Razborov. Lower bounds on the monotone complexity of some boolean functions. *Dokl. Akad. Nauk SSSR*, 281(4):798–801, 1985. (In Russian); English translation in *Soviet Math. Dokl.* 31:354–357, 1985.

[18] D. Shmoys. Cut problems and their application to divide-and-conquer. In [13].

[19] V. V. Vazirani. *Approximation Algorithms.* Springer Verlag, 2001.

[20] L. A. Wolsey. Heuristic analysis, linear programming and branch and bound. *Mathematical Programming Study*, **13**:121–134, 1980.

[21] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. Syst. Sci.*, **43** (1991), pp. 441–466.