THEORY OF COMPUTING, Volume 2 (2006), pp. 19–52
http://theoryofcomputing.org

Proving Integrality Gaps without Knowing the Linear Program

Sanjeev Arora* Béla Bollobás† László Lovász Iannis Tourlakis

Received: June 1, 2005; published: February 3, 2006.

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.

ACM Classification: G.1.6, F.1.3

AMS Classification: 68Q17, 90C05, 90C57

Key words and phrases: linear programming, inapproximability, approximation algorithms, NP-hard problems, integrality gaps, lift-and-project, vertex cover

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 [31], Hochbaum [22]). For instance, a simple 2-approximation algorithm for VERTEX COVER solves the following relaxation: minimize $\sum_{i \in V} x_i$ such that $x_i + x_j \geq 1$ for all $\{i, j\} \in E$. 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 [21]. 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 the minimum VERTEX COVER in G and the cost of the optimum fractional solution.

*Supported by David and Lucile Packard Fellowship, NSF Grants CCR-009818 and CCR-0205594. Work done while visiting Microsoft Research and the CS Dept at UC Berkeley.

†Research supported by NSF grant DSM 9971788 and DARPA grant F33615-01-C-1900

 DOI: 10.4086/toc.2006.v002a002

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 never-ending, 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 guarantees 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 [11], 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 [25], yet decades of work has failed to yield a relaxation with integrality gap better than 1.5 [32] (or $4/3$, if one believes a well-known conjecture [17]). For graph expansion and related graph problems essentially no hardness results exist yet we only know relaxations with integrality gap $\Omega(\log n)$ (Shmoys [29]).

When decades of work has failed to turn up tighter relaxations, one should seriously investigate the possibility that *no tighter relaxations exist*. However, proving such a statement may be related to P vs. NP, since linear programming is complete for P.[1] Thus it seems necessary to work with subfamilies of linear relaxations. An integrality gap result for a large subfamily of relaxations may then be viewed as a lower bound for a restricted computational model, analogous say to lower bounds for monotone circuits [27] and for proof systems [5]. An example is Yannakakis's result [33] that representing TSP (the exact version) using a symmetric linear program requires exponentially many constraints—this ruled out some approaches to $P = NP$ that were being tried at the time.

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 of $2 - o(1)$. 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 any minimum vertex cover must contain almost all the vertices (in particular, it must contain $(1-\alpha)n$ vertices where $\alpha > 0$ is very small), yet the all-$1/2$ solution (or something close to it) is feasible for each inequality. Since the complement of a vertex cover is an independent set, and vice versa, our results may also be trivially rephrased to say that the integrality gap of the INDEPENDENT SET problem for our three families of LPs is unbounded, even though the graphs witnessing these gaps have independent sets of size $\Omega(n)$.

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. Some such 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

[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 problem [33] whether $P = NP$ implies that the VERTEX COVER polytope has a polynomial size description (where additional variables are allowed).

Lovász and Schrijver [24].

The first family consists of linear programs that can include *arbitrary* inequalities on any set of εn variables, for some small constant $\varepsilon > 0$.

The second family consists of linear programs containing inequalities with low *defect*. Usually one defines "defect" for facets of the INDEPENDENT SET polytope (see for instance [24, 23]); here we will make an analogous definition for the VERTEX COVER polytope (i.e., the convex hull of all integra vertex covers): The defect of a VERTEX COVER polytope facet $a^T x \geq b$, where a is a vector of integers and b an integer, is defined to be $2b - \sum_i a_i$. The defects of such facets are always non-negative [24]. Linear programs in the second family are allowed those inequalities defining facets of the VERTEX COVER polytope with 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 $O(\log n)$ rounds of a "lift-and-project" construction of Lovász and Schrijver [24]. This is a method that underlies semidefinite relaxations used in many recent approximation algorithms starting with Goemans and Williamson [19]. The *LS* procedure over many rounds generates tighter and tighter linear relaxations for $0/1$ optimization problems. It is more round-efficient than classical cutting planes procedures such as Gomory-Chvátal [8] 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. In general, one can optimize over the set of inequalities obtained after r rounds in $n^{O(r)}$ time. We show that at least $\Omega(\log n)$ rounds of the *LS* procedure (the LP version, not the semidefinite version) are necessary to reduce the integrality gap below $2 - o(1)$.[2] Note that characterizing the set of inequalities obtained in even $O(1)$ rounds has proved difficult; even the case of 2 rounds is open.

For the first family, better integrality gaps can be obtained for INDEPENDENT SET than those trivially implied by our results for VERTEX COVER. We show that for linear programs where each inequality uses at most $n^{\varepsilon(1-\gamma)}$ variables (here $\varepsilon, \gamma > 0$ are any small constants), the integrality gap for INDEPENDENT SET is $n^{1-\varepsilon}$. This is essentially tight since constraints using n^{ε} variables can clearly approximate INDEPENDENT SET within a factor of $n^{1-\varepsilon}$.

Our techniques seem applicable to problems other than VERTEX COVER (and INDEPENDENT SET) and have been the subject of future work [6, 1, 30]. These developments are discussed in the related work section below. However, several open problems remain. For example, extending our ideas to semidefinite relaxations as well as to the semidefinite programming analogue of the Lovász-Schrijver procedure remains a difficult and interesting open problem. We discuss this and other open problems further in Section 5.

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 $(2 - \varepsilon)$-approximation to VERTEX COVER is NP-hard, the proof would probably involve even more complex reductions than those in [11]. Thus it

[2] **Erratum:** The conference version of this paper [3] argued $\Omega(\sqrt{\log n})$ rounds of the *LS* procedure were needed to reduce the integrality gap for VERTEX COVER below $2 - o(1)$. However, Cheriyan and Qian [7] observed that the argument in [3] was incomplete. In the current paper we give a new (complete) proof of the *LS* round lower bound. Independently, Qian [26] also provides a fix for the proof in [3]. However, our new proof has the advantage of showing that in fact at least $\Omega(\log n)$ rounds of LS tightenings are needed to reduce the integrality gap below $2 - o(1)$.

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 an 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.

Related work A few authors have viewed the Lovász-Schrijver procedure as a proof system and shown that $\Omega(n)$ rounds are required to derive certain simple inequalities (e.g., Goemans and Tunçel [18], Cook and Dash [9]). However, these papers do not consider the issue of how the *integrality gap* improves (or fails to improve) after a few rounds of the *LS* procedure. A recent (and independent) paper by Feige and Krauthgamer [16] considers the question of integrality gaps, but for the maximum CLIQUE problem on a random graph with edge probabilities 1/2. They show that $\Omega(\log n)$ rounds of LS_+, the semi-definite version of Lovász and Schrijver's lift-and-project procedure, are necessary and sufficient to reduce the integrality gap to 1 (with high probability over the choice of the graph). However, this result does not directly give any lower bound on the approximability of VERTEX COVER, since in their graphs both the minimum (integral) vertex cover and the optimal value of the relaxations considered are about n.

Subsequently to our work there have appeared several papers proving integrality gaps for relaxations using both the LP and SDP versions of the Lovász-Schrijver lift-and-project method. Buresh-Oppenheim et al. [6] show that $\Omega(n)$ rounds of LS_+ are needed to obtain relaxations for MAX-kSAT, $k \geq 5$, with integrality gaps less than $(2^k-1)/2^k - \varepsilon$. Alekhnovich et al. [1], building upon [6], show that $\Omega(n)$ rounds of LS_+ are needed to obtain relaxations for MAX-3SAT with integrality gaps less than $7/8-\varepsilon$. In addition they showed that $\Omega(n)$ rounds of LS_+ are needed to obtain relaxations for SET-COVER and rank-k hypergraph VERTEX COVER with integrality gaps less than $(1-\varepsilon)\ln n$ and $k-1-\varepsilon$, respectively. Note that PCP-based results (such as those of Håstad [20], Feige [15] and Dinur et al. [10]) already ruled out non-trivial polynomial-time approximation algorithms for these problems (assuming $\mathrm{P} \neq \mathrm{NP}$). However, they did not rule out slightly subexponential approximation algorithms (defined as those running in 2^{n^c} time for $c < 1$) for the reasons mentioned earlier, namely, the blowup in instance size caused by the PCP-based reductions.

Tourlakis [30], building on techniques used in the current paper, proved that $\Omega(\log\log n)$ rounds of *LS* are needed to obtain relaxations for rank-k hypergraph VERTEX COVER with integrality gaps less than $k-\varepsilon$.

2 The first family

In this section we prove integrality gaps for linear programs in $\{x_1, x_2, \ldots, x_n\}$ for both VERTEX COVER and INDEPENDENT SET where the programs allow any constraint of the form $a^T x \leq b$ such that the coefficient vector a is nonzero for at most εn coordinates. In other words, each constraint involves at most εn variables. Such linear programs 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 $\mathrm{P} = \mathrm{NP}$. We only require that all 0/1 vertex covers and 0/1 independent sets in the graph are feasible for the VERTEX COVER and INDEPENDENT SET relaxations, respectively.

The natural candidate graph for exhibiting integrality gaps for these relaxations would be one where the largest independent set has size at most αn for some small $\alpha > 0$, but every induced subgraph on εn vertices has an independent set of size nearly $\varepsilon n/2$. However, it turns out that the local property we need for our graph is somewhat stronger: all small induced subgraphs have small *fractional chromatic number* which we define below.

We will construct the required graph by the probabilistic method in Theorem 2.3. This result appears to be new, although it fits in a line of results starting with Erdős [13] showing that the chromatic number of a graph cannot be deduced from "local considerations" (see also Alon and Spencer [2], p.130).

Definition 2.1. A *fractional γ-coloring* of a graph G is a multiset $\mathcal{C} = \{U_1, \ldots, U_N\}$ of independent sets of vertices (for some N) such that every vertex is in at least N/γ members of $\mathcal{C}$. The *fractional chromatic number* of G is

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

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, $\chi_f(G) \le \chi(G)$.

Remark 2.2. If $\chi_f(G) = \gamma$ and $\{U_1, \ldots, U_N\}$ is a fractional γ-coloring for G, we will usually assume without loss of generality that each vertex of G (by deleting it from a few of the U_i if necessary) is in *exactly* N/γ sets.

Note that strictly speaking, having $\chi_f(G) = \gamma$ does not guarantee that there exists a fractional γ-coloring for G; it only guarantees a fractional $(\gamma+\varepsilon)$-coloring for all $\varepsilon > 0$. Nevertheless, in the interest of keeping our notation clean, we will always assume that a fractional γ-coloring does exist (in particular, we will only consider rational γ). This slight inaccuracy will not affect the validity of our arguments.

Theorem 2.3. *Let $0 < \alpha, \delta < 1/2$ be constants. Then there exist constants $\beta = \beta(\alpha, \delta) > 0$ and $n_0 = n_0(\alpha, \beta, \delta)$ such that for every $n \ge n_0$ 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 2.3 with α, δ arbitrarily close to 0 and let β be as given by the theorem.

Theorem 2.4. *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, since any independent set is the complement of a vertex cover, and vice versa, the integrality gap is at least $(1-\alpha) \cdot \frac{2+\delta}{1+\delta}$.*

Proof. It suffices to show that the all-$\frac{1+\delta}{2+\delta}$ vector is feasible for any set of constraints $A_I \cdot x \le b_I$ where $I \subseteq \{1, \ldots, n\}$ has size at most βn.

So fix any subset I of at most βn vertices and let $\{U_1, \ldots, U_N\}$ be a fractional $(2+\delta)$-coloring for I such that each vertex in I is in exactly a $1/(2+\delta)$ fraction of the U_i's (see Remark 2.2). Note that 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. By definition, the characteristic vector of any such vertex cover extension obeys $A_I \cdot x \le b_I$. So since these constraints only involve variables from I, it follows that any vector in $\mathbb{R}^n$ that has $1_{I \setminus U_i}$ (the characteristic vector of $I \setminus U_i$) in the coordinates corresponding to I is also feasible for $A_I \cdot x \le b_I$.

Consider the vectors $v_1, v_2, \ldots, v_N \in \mathbb{R}^n$ where v_i is equal to $1_{I \setminus U_i}$ in those coordinates corresponding to I and is $(1+\delta)/(2+\delta)$ 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 a $1 - 1/(2+\delta) = (1+\delta)/(2+\delta)$ fraction of the vertex covers, this average is the all-$\frac{1+\delta}{2+\delta}$ vector. Thus this vector satisfies $A_I \cdot x \leq b_I$, as desired. $\square$

Note that the same construction can also be used to prove integrality gaps for linear relaxations for INDEPENDENT SET:

Corollary 2.5. *Every* INDEPENDENT SET *linear relaxation for H (where H is the same graph as above) where each constraint in the relaxation has at most βn variables has integrality gap at least $\frac{1}{\alpha(2+\delta)}$.*

Proof. Let I be any subset of at most βn vertices and let $\{U_1, \ldots, U_N\}$ be a fractional $(2+\delta)$-coloring for I such that each vertex in I is in exactly a $1/(2+\delta)$ fraction of the U_i's (see Remark 2.2). Now define vectors $v_1, v_2, \ldots, v_N \in \mathbb{R}^n$ as follows: Let v_i equal 1_{U_i} in those coordinates corresponding to I but have v_i equal $1/(2+\delta)$ outside I. Then each v_i is feasible for all constraints involving variables only from I. But then, the average of the v_i's, i.e., the vector with all coordinates $1/(2+\delta)$, is also feasible for these constraints. $\square$

Denote the size of the maximum independent set in a graph G by $\alpha(G)$. The above argument in fact yields the following more general theorem.

Theorem 2.6. *Let G be a graph on n vertices such that every subgraph induced by a set of at most $\beta(n)$ vertices has fractional chromatic number $\leq C$. Then the vector with all coordinates $\frac{1}{C}$ is feasible for any linear relaxation of the* INDEPENDENT SET *constraints for G in which each relaxed constraint involves at most $\beta(n)$ variables. Consequently, the integrality gap for the relaxation is at least $\frac{n}{\alpha(G)C}$.*

This suggests we can obtain larger integrality gaps for INDEPENDENT SET if we further limit the number of variables in each constraint. In Section 2.2 below we show that this is indeed the case by exhibiting graphs for which Theorem 2.6 yields the following:

Theorem 2.7. *Fix $\varepsilon, \gamma > 0$. Then there exists a constant $n_0 = n_0(\varepsilon, \gamma)$ such that for every $n \geq n_0$ there exists a graph G with n vertices for which the integrality gap of any linear relaxation for* INDEPENDENT SET *in which each constraint uses at most $n^{\varepsilon(1-\gamma)}$ variables is at least $n^{1-\varepsilon}$.*

2.1 Proof of Theorem 2.3

The proof uses standard random graph theory supplemented with a couple of new ideas. Let us recall the standard part (see [2]). If we pick a random graph G using the familiar $\mathcal{G}(n,p)$ model and choose p appropriately, then the largest independent set in G has size at most αn and yet the induced subgraph on every subset of βn vertices has an independent 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 cycles (i.e., has high girth). Finally, we will show that these induced subgraphs on βn vertices also satisfy a sparsity condition; this latter property appears to be previously unknown.

We then show in Lemma 2.13 that every high girth graph satisfying this sparsity condition has fractional chromatic number $2+\delta$ on every induced subgraph with at most βn vertices. The proof uses induction on the subgraph size. The main idea in the inductive step is to exhibit a long path inside every subgraph using Lemma 2.12. Peeling away the path gives a smaller subgraph that is colored (fractionally) using the inductive assumption. Lemma 2.11 is then used to extend this fractional coloring, completing the induction.

Now we give details. The next lemma concerns the "standard random graph theory" mentioned above, together with the new sparsity condition.

Lemma 2.8. *Given real numbers α, η with $0 < \alpha < 1/250$ and $0 < \eta < 1/2$, let $\lambda > e^2$ and $\beta > 0$ be such that*

$$2\frac{\log\lambda}{\lambda} \le \alpha \tag{2.1}$$

and

$$\beta < (e\lambda)^{-2/\eta} \ . \tag{2.2}$$

Let $g \ge 3$ be an integer such that $g \le \log n/(3\log\lambda)$. Then there is an integer $n_0 = n_0(\lambda, \eta, g)$ such that for every $n \ge n_0$ 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 $\ell \le \beta n$ vertices contains at most $(1+\eta)\ell$ edges.

Remark 2.9. Condition (2.1) is satisfied if we take

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

Proof of Lemma 2.8. Let us consider the space of random graphs $\mathcal{G}(n,p)$ with $p = \lambda/n$. We will show that a graph $G_{n,p}$ drawn randomly from this space, modulo a few small alterations, satisfies with high probability the three properties required of H in the statement of the lemma.

Let $0 < \alpha_0 < \alpha$ be such that

$$1 + \log(1/\alpha_0) < \lambda\alpha_0/2 \ . \tag{2.3}$$

Inequality (2.1) 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.

1. The probability that for some ℓ, $4 \le \ell \le 1/\eta$, some ℓ-set in $G_{n,p}$ spans at least $\ell+1$ edges is at most

$$\begin{aligned}\sum_{\ell=4}^{1/\eta}\binom{n}{\ell}\binom{\binom{\ell}{2}}{\ell+1}\left(\frac{\lambda}{n}\right)^{\ell+1} &\le \sum_{\ell=4}^{1/\eta}\left(\frac{en}{\ell}\right)^{\ell}\left(\frac{e\frac{\ell(\ell+1)}{2}}{\ell+1}\right)^{\ell+1}\left(\frac{\lambda}{n}\right)^{\ell+1} \\ &= \frac{1}{en}\sum_{\ell=4}^{1/\eta}\ell\left(\frac{e^2\lambda}{2}\right)^{\ell+1} \\ &= O(n^{-1}) \ .\end{aligned}$$

Similarly, the probability that for some ℓ, $1/\eta < \ell < \beta n$, some ℓ-set spans at least $(1+\eta)\ell$ edges is at most

$$\begin{aligned}\sum_{\ell=1/\eta}^{\beta n}\binom{n}{\ell}\binom{\binom{\ell}{2}}{(1+\eta)\ell}\left(\frac{\lambda}{n}\right)^{(1+\eta)\ell} &\le \sum_{\ell=1/\eta}^{\beta n}\left[\left(\frac{en}{\ell}\right)\left(\frac{e\ell}{2(1+\eta)}\right)^{1+\eta}\left(\frac{\lambda}{n}\right)^{1+\eta}\right]^{\ell} \\ &\le \sum_{\ell=1/\eta}^{\beta n}\left[e^2(\ell/n)^{\eta}\lambda^{1+\eta}\right]^{\ell} \ . \qquad (2.4)\end{aligned}$$

We bound (2.4) by first splitting the sum into two quantities and bounding each of them: Letting $C = e^2\lambda^{1+\eta}$ we have that,

$$\sum_{\ell=1/\eta}^{\sqrt{n}}\left[e^2(\ell/n)^{\eta}\lambda^{1+\eta}\right]^{\ell} \le \sum_{\ell=1/\eta}^{\sqrt{n}}\left(\frac{C}{\sqrt{n}}\right)^{\ell} = \left(\frac{C}{\sqrt{n}}\right)^{1/\eta+1}\frac{1-(C/\sqrt{n})^{\sqrt{n}-1/\eta}}{1-(C/\sqrt{n})} = O(n^{-1}) \ .$$

On the other hand, (2.2) implies that $D = e^2\beta^{\eta}\lambda^2 < 1$, and hence,

$$\sum_{\ell=\sqrt{n}+1}^{\beta n}\left[e^2(\ell/n)^{\eta}\lambda^{1+\eta}\right]^{\ell} \le \sum_{\ell=\sqrt{n}}^{\beta n} D^{\ell} \le \frac{D^{\sqrt{n}+1}}{1-D} = O(n^{-1}) \ .$$

So, (2.4) is at most $O(n^{-1})$.

Hence, the probability that *all* ℓ-sets, $\ell \le \beta n$, in $G_{n,p}$ span at most $(1+\eta)\ell$ edges is at least $1-O(n^{-1})$.

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

$$\mathbb{E}(I) = \binom{n}{\alpha_0 n}\left(1-\frac{\lambda}{n}\right)^{\binom{\alpha_0 n}{2}} \le \left(\left(\frac{e}{\alpha_0}\right)e^{-\lambda\alpha_0/2}\right)^{\alpha_0 n} = \gamma_0^{\alpha_0 n} \ .$$

Inequality (2.3) implies that $\gamma_0 < 1$, so the probability that $G_{n,p}$ contains an independent set of $\alpha_0 n$ vertices is exponentially small.

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

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

By Markov's inequality, $G_{n,p}$ has at most $n^{1/2}$ short cycles with probability at least $1-O(n^{-1/6})$. Deleting an edge from each of these cycles then gives a graph of girth at least g.

Consequently, with probability $1-O(n^{-1/6})$, $G_{n,p}$ has no set of $\ell \le \beta n$ vertices spanning more than $(1+\eta)\ell$ edges, and moreover, 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. It is easy to check that a path with at least one edge has fractional chromatic number 2. In particular, a graph has fractional chromatic number less than 2 if and only if it is an independent set. The proof of the next lemma is left to the reader.

Lemma 2.10. *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 $|V(G_1) \cap V(G_2)| \le 1$ then $\chi_f(G_1 \cup G_2) = \max\{\chi_f(G_1),\ \chi_f(G_2)\}$.

The next Lemma concerns the fractional chromatic number of a graph that contains a long path (i.e., the vertices on the path's interior have no edges outside the path edges).

Lemma 2.11. *Let $\ell \ge 2$ and let G be a graph obtained by adding a path $x_0x_1 \dots x_{\ell+1}$ to a graph G', where $x_0, x_{\ell+1} \in V(G')$ and $x_i \notin V(G')$ for $1 \le i \le \ell$. Then $\chi_f(G) \le \max\{\chi_f(G'), \frac{2\ell}{\ell-1}\}$.*

Proof. Let $\chi_f(G') = 1/\gamma$ and suppose first that $\gamma > 1/2$. Then G is an independent set and the lemma follows since paths have fractional chromatic number 2.

So assume $\gamma \le 1/2$. By Remark 2.2 we can assume without loss of generality that there exists a multiset $\mathcal{C}' = \{U'_1, \dots, U'_N\}$ of independent sets in G' such that every vertex of G' is in *exactly* γN of these sets. So since $x_0 \in G'$ and $\gamma \le 1/2$, there exists a multiset $\mathcal{A}$ containing exactly $N/2$ sets from $\mathcal{C}'$ such that no set in $\mathcal{A}$ contains x_0. Similarly, there exists a multiset $\mathcal{B}$ of $N/2$ sets taken from $\mathcal{C}'$ such that no set in $\mathcal{B}$ contains $x_{\ell+1}$.

Fix i, $1 \le i \le n$. We will define a colouring $\mathcal{C}_i$ for $G \setminus \{x_i\}$ (i.e., G with x_i removed) by extending the sets U'_h in $\mathcal{C}'$ to independent sets U_h in $G \setminus \{x_i\}$. Moreover, our colouring will have the property that each x_j, $1 \le j \le \ell$, $j \ne i$ will be in exactly half the sets of $\mathcal{C}_i$. Our approach will be as follows: Fix a set $U'_h \in \mathcal{C}'$. If $U'_h \in \mathcal{A}$ we will then add to U_h every other node in the path fragment from x_1 to x_{i-1} starting with x_1: That is, $x_1, x_3, x_5, \dots$ will be in U_h, but $x_2, x_4, \dots$ will not. If instead $U'_h \notin \mathcal{A}$ then $x_2, x_4, x_6, \dots$ will be in U_h, but $x_1, x_3, \dots$ will not. Similarly, we will decide which of the nodes $x_{i+1}, x_{i+2}, \dots, x_\ell$ to add to U_h depending on whether or not U'_h is in $\mathcal{B}$. Since $\mathcal{A}$ and $\mathcal{B}$ each contain exactly half the sets of $\mathcal{C}'$, it will follow that each x_j, $1 \le j \le \ell$, $j \ne i$, is in exactly half the sets of $\mathcal{C}_i$.

Formally, the exact construction is as follows: Fix $U'_h \in \mathcal{C}'$. For $1 \le j \le i$, if $U'_h \in \mathcal{A}$ then add x_j to U_h if j is odd; if instead $U'_h \notin \mathcal{A}$ then add x_j to U_h if j is even. For $i \le j \le \ell$, if $U'_h \in \mathcal{B}$ then add x_j to U_h if $\ell - j$ is even; if instead $U'_h \notin \mathcal{B}$ then add x_j to U_h if $\ell - j$ is odd.

Let $\mathcal{C} = \cup \mathcal{C}_i$. This multiset of ℓN sets is then a fractional colouring for G. Note that every vertex of G' is in $\gamma \ell N$ sets of $\mathcal{C}$. Moreover, every x_i, $1 \le i \le \ell$, is in $\frac{1}{2}(\ell-1)N$ sets. Consequently, every vertex of G is in at least a $\min\{\gamma, \frac{\ell-1}{2\ell}\}$ fraction of the sets of $\mathcal{C}$. Hence, $\chi_f(G) \le \max\{\frac{1}{\gamma}, \frac{2\ell}{\ell-1}\}$. □

For real $k > 1$ call a graph *k-sparse* if it has no subgraph with ℓ vertices and more than $k\ell$ edges. Hence, sparsity quantifies (half of) the maximum average degree of subgraphs. This concept is closely related to that of *degeneracy*: Recall that a graph is *k-degenerate* if every subgraph has a vertex of degree at most k. Hence, if k is a natural number and $\varepsilon > 0$, then a $(\frac{k+1}{2} - \varepsilon)$-sparse graph is k-degenerate; conversely, every k-degenerate graph is k-sparse.

Recall that a graph is *k-connected* if there does not exist a set of $k-1$ vertices whose removal disconnects the graph. By length of a path we will mean the number of edges.

Lemma 2.12. *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 at least $\ell+1$ 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. Since G is a 2-connected graph with more edges than vertices, G consists of a certain $k \geq 2$ number of branch-vertices (i.e., vertices of degree at least 3) and the induced paths joining them, say $P_1, \dots, P_m$, where $m \geq \lceil 3k/2 \rceil$, and all internal nodes in these paths have degree 2. Let $\ell_i \leq \ell$ denote the length of P_i. Then,

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

and so, $m - k \leq \eta n$. On the other hand,

$$n = k + \sum_{i=1}^{m} (\ell_i - 1) \leq k + m(\ell - 1) \ ,$$

and hence $m - k \leq \eta(k + m(\ell-1))$. But then, since $m \geq \lceil 3k/2 \rceil$, it follows that $\eta \geq 1/(3\ell-1)$, a contradiction. □

Lemma 2.13. *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 $\chi_f(G) \leq 2 + \frac{2}{h}$.*

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 2.10. 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 2.10. So assume G is not a cycle. But then, by Lemma 2.12, 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). By the induction hypothesis, $\chi_f(G') \leq 2 + \frac{2}{h}$, and so by Lemma 2.11 we have $\chi_f(G) = \max\{\chi_f(G'), 2 + \frac{2}{h}\} \leq 2 + \frac{2}{h}$. This completes the induction and the Lemma is proved. □

We can now prove Theorem 2.3.

Proof of Theorem 2.3. Set $h = \lceil 2/\delta \rceil$, $g = 2h$ and $\eta = \frac{1}{3h+3}$. Choose $\lambda > e^2$ and $\beta > 0$ to satisfy inequalities (2.1) and (2.2). Let H be a graph of order n whose existence is guaranteed by Lemma 2.8. Thus, H has independence number at most αn, and if G is a subgraph of H with at most βn vertices then G is $(1+\eta)$-sparse and has girth at least $g = 2h$. Hence, by Lemma 2.13, $\chi_f(G) \leq 2 + \frac{2}{h} \leq 2 + \delta$, completing the proof. □

2.2 Proof of Theorem 2.7

Throughout this section, log will denote base-2 logarithms.

By Theorem 2.6, to obtain a large integrality gap we need to construct graphs where the independence and local fractional chromatic numbers are as small as possible. One way to do this is using graph products.

Definition 2.14. The *inclusive graph product* $G \times H$ of two graphs G and H is the graph on $V(G \times H) = V(G) \times V(H)$ where $\{(x,y),(x',y')\} \in E(G \times H)$ iff $(x,x') \in E(G)$ or $(y,y') \in E(H)$. The notation G^k indicates the graph resulting by taking the k-fold inclusive graph product of G with itself.

The key observation is that $\alpha(G \times H) = \alpha(G) \times \alpha(H)$ and $\chi_f(G \times H) = \chi_f(G)\chi_f(H)$ (the former fact is easy; for the latter see [14] for a proof). Moreover, if all sets of size at most βn have fractional chromatic number C in G, then all sets of size at most βn in G^k have fractional chromatic number C^k. So taking products of a graph with itself drives down the relative sizes of both the independence and local fractional chromatic numbers. However, since the resulting graph is much larger, the fractional chromatic number is small only for negligibly sized subgraphs. To get around this we instead consider an appropriately chosen (small) random subgraph of G^k. The particular construction we use is due to Feige [14]. By choosing each vertex of G^k independently at random with probability $\alpha(G)^{-k}$ and analyzing the resulting induced subgraph, Feige proves the following theorem (we sketch a proof below for completeness; see [14] for details):

Theorem 2.15 (Feige [14]). *There exists an integer n_0 such that for every graph G on $n \geq n_0$ vertices and any integer k, there exists a graph G_k such that:*

1. *G_k is a vertex induced subgraph of G^k.*
2. $\frac{1}{2}\left(\frac{n}{\alpha(G)}\right)^k \leq |V(G_k)| \leq 2\left(\frac{n}{\alpha(G)}\right)^k$.
3. $\alpha(G_k) \leq \frac{k\alpha(G)\ln n}{\ln(k\alpha(G)\ln n)}$.

Proof. (Sketch) Select each vertex of G^k independently and at random with probability $\alpha(G)^{-k}$. Let $\hat{G}$ be the induced subgraph of G^k obtained by this process. We show that $\hat{G}$ satisfies the above three properties with high probability.

By construction, $\hat{G}$ is an induced subgraph of G^k. Moreover, the probability that $|V(\hat{G})|$ deviates by more than a factor of 2 from its expectation is negligible. For the last property, fix a maximal independent set I in G^k. The expected number of vertices from I in $\hat{G}$ is at most 1. Chernoff bounds sharply bound the probability that more than $\frac{k\alpha(G)\ln n}{\ln(k\alpha(G)\ln n)}$ vertices of I survive in $\hat{G}$. The last property can now be seen to hold with high probability by observing that G contains at most $n^{\alpha(G)}$ maximal independent sets and by observing that all maximal independent sets in G^k are the direct product of k maximal independent sets in G. In particular, the probability that more than $\frac{k\alpha(G)\ln n}{\ln(k\alpha(G)\ln n)}$ vertices of any maximal independent set of G^k survive in $\hat{G}$ can be shown to go to 0 as n grows. □

Our strategy for proving Theorem 2.7 will then be as follows: We will start with a graph G where both the independence number and local fractional chromatic number are already small (such a graph will exist by Theorem 2.3) and then apply Feige's randomized graph product to it.

Now the details. Fix arbitrarily small constants $\alpha, \delta > 0$ and $n > 0$ such that $n \geq n_0$ where n_0 is from Theorem 2.15. Provided that n is chosen sufficiently large, Theorem 2.3 implies that there exists a graph G on n vertices such that $\alpha(G) \leq \alpha n$ and such that for some constant $\beta > 0$, all induced subgraphs of G with at most βn vertices have chromatic number $\leq 2 + \delta$.

Fix an arbitrarily small constant $d > 0$ and let G_k be the graph given by Theorem 2.15 for $k = d \log n$. Let $N = |V(G_k)|$. Note that $N = \Theta(\alpha^{-k}) = \Theta(n^{d \log(1/\alpha)})$. On the other hand, all subsets of G_k of size at most

$$\beta n = \Theta\left(N^{\frac{1/d}{\log(1/\alpha)}}\right) \tag{2.5}$$

have fractional chromatic number $\leq (2+\delta)^k$.

By Theorem 2.6 it follows that any linear relaxation of the independent set constraints for G_k where the relaxed constraints contain at most βn variables has integrality gap (the $\tilde{\Theta}$ notation indicates asymptotic order up to logarithmic factors):

$$\Theta\left(\frac{\alpha^{-k}}{(2+\delta)^k \frac{k\alpha n \ln n}{\ln(k\alpha n \ln n)}}\right) = \tilde{\Theta}\left(n^{d(\log(1/\alpha) - \log(2+\delta)) - 1}\right) = \tilde{\Theta}\left(N^{1 - \frac{1/d + \log(2+\delta)}{\log(1/\alpha)}}\right). \tag{2.6}$$

Since we can take α and δ to be arbitrarily small in Theorem 2.3 (provided n is large enough), and since $d > 0$ can also be chosen arbitrarily small, it follows that we can simultaneously make (2.5) more than $N^{\varepsilon(1-\gamma)}$ and (2.6) more than $N^{1-\varepsilon}$. The theorem follows.

3 The second family

For an n-vertex graph G, let $\mathrm{VC}(G)$ denote the convex hull of all integral vertex covers for G, i.e., the convex hull of all 0/1 vectors $x \in \mathbb{R}^n$ satisfying $x_i + x_j \geq 1$ for all edges $\{i, j\}$ in G. All non-trivial facets of the polytope $\mathrm{VC}(G)$ can be expressed in the form $a^T x \geq b$ where $a \in \mathbb{Z}_+^V$ and $b \in \mathbb{Z}_+$. (By non-trivial we exclude facets of the form $x_i \geq 0$ and $x_i \leq 1$, and require that at least two coordinates of a are non-zero.) Note moreover that the non-trivial facets of *any* relaxation for $\mathrm{VC}(G)$ lying in $[0,1]^n$ must also be of the form $a^T x \geq b$ where $a \in \mathbb{Z}_+^V$ and $b \in \mathbb{Z}_+$.

While $\mathrm{VC}(G)$ requires exponentially many non-trivial facets to completely specify, it may be that a smaller subset of these facets yields a linear relaxation with integrality gap less than $2 - \varepsilon$ for some $\varepsilon > 0$. In this section we consider relaxations defined by those facets of $\mathrm{VC}(G)$ having low defect.

The *defect* of a facet $a^T x \geq b$ of $\mathrm{VC}(G)$ is defined to be $2b - \sum_i a_i$. It follows from the proof of an analogous result for the Independent Set polytope by Lovász and Schrijver [24] that this quantity is always non-negative for facets of $\mathrm{VC}(G)$. For more about defects of facets see [24] and also Lipták and Lovász [23].

We now generalize the results of Section 2 to any linear relaxation for VERTEX COVER defined by facets of $\mathrm{VC}(G)$ with defect at most εn:

Theorem 3.1. *For all $\gamma > 0$ there exists a constant $\varepsilon > 0$ such that the integrality gap is at least $2 - \gamma$ for any relaxation for* VERTEX COVER *consisting of inequalities of the form $a^T x \geq b$ ($a \in Z_+^n, b \in Z_+$) and defect at most εn.*

Proof. Let $\alpha, \delta > 0$ be constants such that $(1-\alpha)\frac{2+\delta}{1+\delta} \geq 2-\gamma$, and let H be the graph constructed in Theorem 2.3 for these constants. Let the defect of our relaxation be at most εn where $\varepsilon = \beta\delta/(2+\delta)$. The theorem will follow by showing that the vector x_δ with all coordinates $(1+\delta)/(2+\delta)$ is feasible.

There are two types of facets $a^T x \geq b$. 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 Theorem 2.4. If $\sum_i a_i > \beta n$ then the feasibility of x_δ follows by direct substitution:

$$\sum_i a_i \frac{1+\delta}{2+\delta} = \sum_i a_i \frac{\delta}{2(2+\delta)} + \frac{1}{2}\sum_i a_i > \frac{\beta n \delta}{2(2+\delta)} + \frac{1}{2}\sum_i a_i = \frac{\varepsilon n}{2} + \frac{1}{2}\sum_i a_i \geq b \ .$$

□

4 The third family

Consider the standard relaxation for VERTEX COVER:

$$x_i + x_j \geq 1 \qquad \forall \{i,j\} \in E \qquad (\textit{Edge constraint}) \tag{4.1}$$

In this relaxation the x_i's are real numbers in $[0,1]$. Suppose we wish to tighten the relaxation to force the x_i's to be $0/1$ in any optimal solution. To this end, we could 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,

$$\sum_{i \in C} x_i \geq \frac{|C|+1}{2} \qquad (\textit{Odd-cycle constraint}) \tag{4.2}$$

Many other families of inequalities satisfied by 0/1 vertex covers are known, but a complete listing will probably never be found because of complexity reasons.

Lovász and Schrijver [24] give an automatic method for generating over many rounds all valid inequalities. More generally, they give a method for obtaining tighter and tighter relaxations for any $0/1$ optimization problem starting from an arbitrary relaxation. The idea is to "lift" to n^2 dimensions and then project back to n-space. This is why the procedure is called "lift-and-project" or "lifting." The motivation is to try to simulate the power of quadratic programs. Solving quadratic programs is of course NP-hard since adding the constraints $x_i(1-x_i) = 0$ to a linear relaxation forces 0/1 answers. For example, all 0/1 vertex covers satisfy

$$x_i^2 = x_i \tag{4.3}$$

$$(1-x_i)(1-x_j) = 0 \qquad \forall \{i,j\} \in E \ . \tag{4.4}$$

To linearly simulate these constraints, we can introduce new linear variables Y_{ij} to "represent" the products $x_i x_j$ and then demand that the "lifted" variables satisfy $x_i = Y_{ii}$ and $1 - x_i - x_j + Y_{ij} = 0$ for all edges $\{i,j\}$. We can then take positive linear combinations of these constraints to eliminate all "quadratic" terms and obtain constraints using only the original variables x_i.

Formally, given a relaxation

$$a_r^T x \geq b \qquad r = 1,2,\ldots,m \tag{4.5}$$

$$0 \leq x_i \leq 1 \qquad i = 1,2,\ldots,n \ , \tag{4.6}$$

one round of *LS* produces a system of inequalities in $(n+1)^2$ variables Y_{ij} for $i, j = 0, 1, \ldots, n$. As already mentioned, the intended "meaning" is that $Y_{ij} = x_i x_j$ and $Y_{00} = 1, Y_{0i} = x_i x_0 = x_i$, and $Y_{00} = 1$ so every quadratic expression in the x_i's can be viewed as a linear expression in the Y_{ij}'s. This is how the quadratic inequalities below should be interpreted. The following inequalities are derived in one round:

$$
\begin{aligned}
(1-x_i)a_r^T x &\geq (1-x_i)b & \qquad & \forall i = 1, \ldots, n, \quad \forall r = 1, \ldots, m \\
x_i a_r^T x &\geq x_i b & & \forall i = 1, \ldots, n, \quad \forall r = 1, \ldots, m \\
x_i x_i &= x_i & & \forall i = 1, 2, \ldots, n
\end{aligned}
$$

The last constraint corresponds to the fact that $x_i^2 = x_i$ for 0/1 variables. Since any positive combination of the above inequalities is also implied, we can use such combinations to eliminate all non-linear terms.

Lovász and Schrijver show that every inequality valid for the integral hull is generated in at most n rounds. Moreover, they show that the set of inequalities derivable in one round for the VERTEX COVER relaxation are exactly the odd-cycle inequalities. To illustrate, we now show how to derive in one round the odd-cycle inequality $x_1 + x_2 + x_3 \geq 2$ for a triangle on nodes $\{1, 2, 3\}$ starting from the edge constraints (4.1). One round of *LS* generates the following inequalities (amongst others):

$$(1-x_1)(x_1+x_2) \geq 1-x_1 \tag{4.7}$$

$$(1-x_2)(x_2+x_3) \geq 1-x_2 \tag{4.8}$$

$$(1-x_3)(x_1+x_3) \geq 1-x_3 \tag{4.9}$$

$$x_1(x_2+x_3) \geq x_1 \tag{4.10}$$

$$x_2(x_1+x_3) \geq x_2 \tag{4.11}$$

Adding inequality (4.7) twice to the sum of the remaining four inequalities and then simplifying using the rule $x_i^2 = x_i$ gives $x_1 + x_2 + x_3 \geq 2$ as desired.

No exact characterization exists for the inequalities derivable in subsequent rounds. However, we do know that the set of inequalities derivable in $O(1)$ rounds has a polynomial-time separation oracle. For more details see [24].

To understand our results, the reader only needs to know the next Lemma taken from [24] and which gives an alternate characterization of *LS* liftings useful for proving lower bounds. The notation uses homogenized inequalities. Let $\mathrm{FR}(G)$ be the cone in $\mathbb{R}^{n+1}$ that contains a vector $(x_0, x_1, \ldots, x_n)$ iff it satisfies $0 \leq x_i \leq x_0$ for all i as well as the edge constraints $x_i + x_j \geq x_0$ for each edge $\{i, j\} \in G$. All cones below will be in $\mathbb{R}^{n+1}$ and we are interested in the slice cut out by the hyperplane $x_0 = 1$. Denote by $N^r(\mathrm{FR}(G))$ the feasible cone of all inequalities obtained from r rounds of the *LS* lifting procedure. Let e_i denote the ith unit vector so that Ye_i denotes the ith column of Y. The next lemma defines the effect of one round.

Lemma 4.1 ([24]). *If K is a cone in $\mathbb{R}^{n+1}$, then $x \in N^m(K)$ iff there is an $(n+1) \times (n+1)$ symmetric matrix Y satisfying*

1. *$Ye_0 = diag(Y) = x$.*
2. *For $1 \leq i \leq n$, both Ye_i and $Y(e_0 - e_i)$ are in $N^{m-1}(K)$.*

Following [6] we will call the matrix Y witnessing that $x \in N^m(K)$ in the above lemma a *protection matrix* since it "protects" x for one round of *LS* tightening.

In practice, we will only be concerned with showing that vectors $x \in \mathbb{R}^{n+1}$ with $x_0 = 1$ survive a round of lifting. For such points, we have the following corollary of Lemma 4.1:

Corollary 4.2. *Let K be a cone in $\mathbb{R}^{n+1}$ and suppose $x \in \mathbb{R}^{n+1}$ where $x_0 = 1$. Then $x \in N^m(K)$ iff there is an $(n+1) \times (n+1)$ symmetric matrix Y satisfying*

1. *$Ye_0 = diag(Y) = x$.*
2. *For $1 \leq i \leq n$: If $x_i = 0$ then $Ye_i = \vec{0}$; If $x_i = 1$ then $Ye_i = x$; Otherwise, Ye_i/x_i, $Y(e_0 - e_i)/(1 - x_i)$ both lie in the projection of $N^{m-1}(K)$ onto the hyperplane $x_0 = 1$.*

Our main theorem for this section is the following:

Theorem 4.3. *For all $\varepsilon > 0$ there exists an integer n_0 and a constant $\delta(\varepsilon) > 0$ such that for all $n \geq n_0$, there exists an n vertex graph G for which the integrality gap of $N^r(\mathrm{FR}(G))$ for any $r \leq \delta(\varepsilon) \log n$ is at least $2 - \varepsilon$.*

The proof of Theorem 4.3 relies on the following two theorems. The first (which also follows as a subcase from the arguments used to prove Lemma 2.8) is essentially due to Erdős [12]; see Bollobás [4], Theorem 4, Ch VII. The second, Theorem 4.5, will be proved in Section 4.2 with an overview of the argument first given in Section 4.1.

Theorem 4.4. *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 with girth at least $\log n/(3\log(1/\alpha))$ but no independent set of size greater than αn.*

Let y_γ denote the vector $(1, \frac{1}{2} + \gamma, \frac{1}{2} + \gamma, \ldots, \frac{1}{2} + \gamma)$ where $0 < \gamma < \frac{1}{2}$.

Theorem 4.5. *Let $G = (V, E)$ have $\mathrm{girth}(G) \geq 16r/\gamma$. Then $y_\gamma \in N^r(\mathrm{FR}(G))$.*

Proof of Theorem 4.3. Let $\gamma = \varepsilon/8$ and $\alpha = \varepsilon/4$, and let n_0 be the constant from Theorem 4.4 for this α. For $n \geq n_0$, let G be the n-vertex graph given by Theorem 4.4. Finally, let $\delta(\varepsilon) = \frac{\varepsilon}{384\log(4/\varepsilon)}$. Then by Theorem 4.5, y_γ is in $N^r(\mathrm{FR}(G))$ for all $r \leq \delta(\varepsilon)\log n$, and hence, the integrality gaps for all these polytopes is at least $2(1-\alpha)/(1+2\gamma) \geq 2 - \varepsilon$. □

4.1 Intuition for Theorem 4.5

Lemma 4.1 (and Corollary 4.2) suggest using induction to prove Theorem 4.5. We first will identify for each j some large set of vectors within each polytope $N^j(\mathrm{FR}(G))$ called the "palette" for $N^j(\mathrm{FR}(G))$. In stage j of the induction we will show the following: For each vector x in the palette for $N^j(\mathrm{FR}(G))$, there exists a protection matrix Y such that for all $i \in [n]$ the vectors Ye_i and $Y(e_0 - e_i)$ all lie in the palette for the previous polytope $N^{j-1}(\mathrm{FR}(G))$ (Figure 1). The condition that such a protection matrix exists can be expressed as an LP. Hence, to show that a protection matrix exists for each x in the palette for $N^j(\mathrm{FR}(G))$ we show using Farkas's lemma that the LP is feasible. The theorem then follows since our definition for the palette for $N^r(\mathrm{FR}(G))$ will ensure that it contains y_γ.

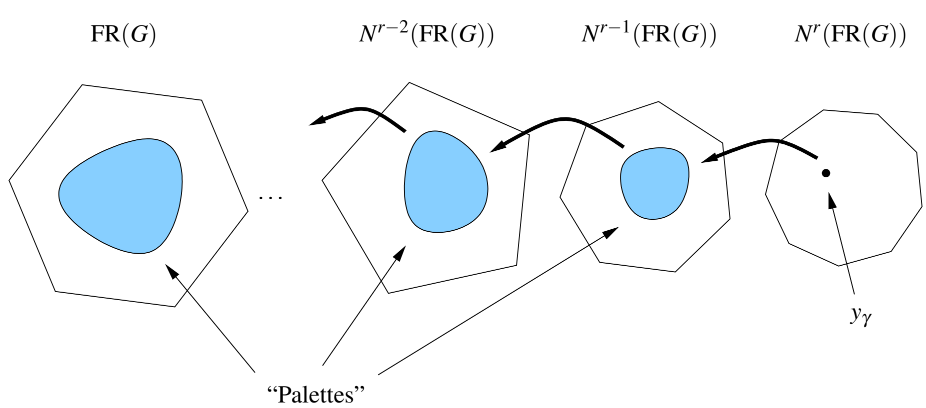

Figure 1: Chain of dependencies in the proof of Theorem 4.5: Each palette is contained in its respective polytope because some other palette is contained in the previous polytope.

Since our protection matrices will be found using LP duality, we will pick the simplest palettes possible in order to ensure that our LPs are also as simple as possible (and hence easy to analyze). To understand what desirable properties the palette vectors should have, let us look at the simpler problem of showing that $y_\gamma \in N(\mathrm{FR}(G))$ (rather than showing $y_\gamma \in N^r(\mathrm{FR}(G))$) and make some observations about the constraints the conditions in Corollary 4.2 force upon a protection matrix for y_γ.

To that end, consider the projected "columns" $Ye_i/y_\gamma^{(i)}$ and $Y(e_0-e_i)/(1-y_\gamma^{(i)})$ of Y (from condition 2 of Corollary 4.2). These vectors must satisfy the edge constraints. As will be shown in Section 4.4 (see equation (4.14)), the constraints forcing this are given by the following constraint:

$$\alpha_i \le Y_{ij} + Y_{ik} \le \alpha_i + (\alpha_j + \alpha_k - 1) \qquad \forall i \in \{1,\dots,n\}, \forall \{j,k\} \in E \ . \tag{4.12}$$

Fix i. If j_1 is adjacent to i, then (4.12) implies $\frac{1}{2}+\gamma \le Y_{ii}+Y_{ij_1} \le \frac{1}{2}+3\gamma$. Since Y is a protection matrix for y_γ, it must satisfy $Y_{ii} = y_\gamma^{(i)} = \frac{1}{2}+\gamma$. Hence, $0 \le Y_{ij_1} \le 2\gamma$. Now consider a node j_2 at distance 2 from j_1. Then (4.12), together with the fact that $0 \le Y_{ij_1} \le 2\gamma$ for all j_1 adjacent to i, imply that $\frac{1}{2}-\gamma \le Y_{ik} \le \frac{1}{2}+3\gamma$. In turn, for a node j_3 at distance 3 from i we must have $0 \le Y_{ij_3} \le 4\gamma$; and for a node j_4 at distance 4 from i we have $\frac{1}{2}-3\gamma \le Y_{ij_4} \le \frac{1}{2}+3\gamma$. So as j gets further and further from i, the constraints on Y_{ij} implied by (4.12) get looser and looser so that for nodes j sufficiently far from i (distance $2/\gamma$ more than suffices) no constraint on Y_{ij} is implied. So intuitively, for such j we should be able to choose Y_{ij} such that node j remains $\frac{1}{2}+\gamma$ in both $Ye_i/y_\gamma^{(i)}$ and $Y(e_0-e_i)/(1-y_\gamma^{(i)})$. Note that the fact that the coordinates of y_γ are $\frac{1}{2}+\gamma$ instead of $\frac{1}{2}$ is crucial in ensuring that the effects of the edge constraints die out as we get further away from node i. Note also that we have implicitly assumed that our graph has girth larger than $2/\gamma$ so that two nodes cannot be connected by two paths of different lengths both less than $2/\gamma$—intuitively this is why Theorem 4.5 requires large girth. We should also mention that we have simplified things by ignoring constraints required by Corollary 4.2 forcing the projected "columns" to lie in $[0,1]^{n+1}$: these tighten the above constraints on the Y_{ij} a bit but the intuition given above is mostly unchanged.

In any case, the above suggests that to prove $y_\gamma \in N(\mathrm{FR}(G))$ we could use a palette consisting of all vectors in $\mathrm{FR}(G)$ which are $\frac{1}{2}+\gamma$ everywhere except perhaps on some ball of radius $2/\gamma$ in G. As such, we can add "palette constraints" to the LP defining Y forcing all nodes j distant from i to be $\frac{1}{2}+\gamma$ in both $Ye_i/y_\gamma^{(i)}$ and $Y(e_0-e_i)/(1-y_\gamma^{(i)})$. In fact, since Y must also be symmetric, the actual constraints we will add will force the following: for all pairs of nodes i,j with distance at least $2/\gamma$ between them, the jth nodes in $Ye_i/y_\gamma^{(i)}$ and $Y(e_0-e_i)/(1-y_\gamma^{(i)})$, and the ith nodes $Ye_j/y_\gamma^{(j)}$ and $Y(e_0-e_j)/(1-y_\gamma^{(j)})$ must all be $\frac{1}{2}+\gamma$.

The proof of Theorem 4.5 will use generalized versions of the above palette: The palettes for each polytope $N^j(\mathrm{FR}(G))$ will consist of vectors from $\mathrm{FR}(G)$ that are $\frac{1}{2}+\gamma$ except in a few neighbourhoods (see Definition 4.6 in Section 4.2 for the precise statement). For a vector x in the palette for $N^j(\mathrm{FR}(G))$ the LP used to find a protection matrix Y for x will have two types of constraints: constraints that force Y to satisfy the conditions in Corollary 4.2 and constraints that force the "columns" Ye_i/x_i and $Y(e_0-e_i)/(1-x_i)$ to belong to the "palette" for $N^{j-1}(\mathrm{FR}(G))$.

The palettes we will use will have the following property: The diameter of the largest neighbourhood H in G such that H consists entirely of nodes with values not equal to $\frac{1}{2}+\gamma$ will grow linearly with the number of rounds. Hence, our method is limited to proving integrality gaps for at most $O(\log n)$ rounds since only graphs with girth $O(\log n)$ yield large integrality gaps.[3]

4.2 Proof of Theorem 4.5

The theorem will be proved by induction where the inductive hypothesis requires a set of vectors other than just y_γ to be in $N^m(\mathrm{FR}(G))$ for $m \le r$ (the "palettes" from Section 4.1). These vectors will be essentially all-$(\frac{1}{2}+\gamma)$, except possibly for a few small neighborhoods where the vector can take arbitrary nonnegative values so long as the edge constraints are satisfied. Let $Ball(w,R)$ denote the set of vertices within distance R of w in G.

Definition 4.6. Let $S \subseteq \{1,\ldots,n\}$, R be a positive integer and $\gamma > 0$. Then a nonnegative vector $(\alpha_0,\alpha_1,\ldots,\alpha_n) \in [0,1]^{n+1}$ with $\alpha_0 = 1$ is an (S,R,γ)-*vector* if the entries satisfy the edge constraints and if for each $w \in S$ there exists a positive integer R_w such that

1. $\sum_{w\in S}(R_w + \frac{2}{\gamma}) \le R$
2. For distinct $w,w' \in S$, $Ball(w,R_w) \cap Ball(w',R_{w'}) = \emptyset$
3. $\alpha_j = \frac{1}{2}+\gamma$ for each $j \notin \cup_{w\in S} Ball(w,R_w)$

We will say that the integers $\{R_w\}_{w\in S}$ *witness* that α is an (S,R,γ)-vector.

Let $R^{(r)} = 0$ and let $R^{(m)} = R^{(m+1)} + \frac{4}{\gamma}$ for $0 \le m < r$. Note that $4R^{(m)} \le \mathrm{girth}(G)$ for $0 \le m \le r$. To prove Theorem 4.5 we will prove the inductive claim below. Since the set of $(\emptyset, R^{(r)},\gamma)$-vectors consists precisely of the vector y_γ, the theorem will then follow as a subcase of the case $m = r$.

Inductive Claim for $N^m(\mathrm{FR}(G))$: For every set S of at most $r-m$ vertices, every $(S,R^{(m)},\gamma)$-vector is in $N^m(\mathrm{FR}(G))$.

Base case $m = 0$. Trivial since $(S,R^{(0)},\gamma)$-vectors satisfy the edge constraints for G.

Proof for $m+1$ assuming truth for m. Let α be an $(S,R^{(m+1)},\gamma)$-vector where $|S| \le r-m-1$. To show that $\alpha \in N^m(\mathrm{FR}(G))$ it suffices to find a protection matrix Y for α satisfying the properties of Corollary 4.2. We exploit the structure of (S,R,γ)-vectors and prove some important structural properties of these vectors in Lemma 4.7, which then enables us to argue that such a protection matrix exists thereby completing the induction step.

Note first that α is trivially an $(S \cup i, R^{(m)},\gamma)$-vector for any $i \in G$. Lemma 4.7, which we now state and prove in Section 4.3 below, says that for appropriate sets S', $|S'| \le r-m$, α is also an $(S',R^{(m)},\gamma)$-vector enjoying crucial additional structural properties.

Lemma 4.7. *Let i be such that $\alpha_i \notin \{0,1\}$. Then there exists a set $S_i \subseteq \{1,\ldots n\}$, $|S_i| \le r-m$, and positive integers $\{R_w^{(m)}\}_{w\in S_i}$ such that,*

1. *α is an $(S',R^{(m)},\gamma)$-vector with witnesses $\{R_w^{(m)}\}_{w\in S_i}$*
2. *$i \in \cup_{w\in S_i} Ball(w,R_w^{(m)})$*

[3] In the conference version of this paper [3], the palettes were picked such that the diameter of the largest neighbourhood grew *quadratically* in the number of rounds, thereby yielding integrality gaps only for $O(\sqrt{\log n})$ rounds.

3. *For each* $\ell \notin \cup_{w \in S_i} Ball(w, R_w^{(m)})$*, any path between* i *and* ℓ *in* G *contains at least* $\frac{2}{\gamma}$ *consecutive vertices* ℓ *such that* $\alpha_\ell = \frac{1}{2} + \gamma$

By the induction hypothesis, for any $S' \subseteq \{1, \dots, n\}$ such that $|S'| \le r - m$, every $(S', R^{(m)}, \gamma)$-vector is in $N^m(\mathrm{FR}(G))$. Hence, to show that $\alpha \in N^{m+1}(\mathrm{FR}(G))$ it suffices by Corollary 4.2 to exhibit an $(n+1) \times (n+1)$ symmetric protection matrix Y that satisfies:

A. $Ye_0 = diag(Y) = \alpha$,

B. For each i such that $\alpha_i = 0$, we have $Ye_i = 0$; for each i such that $\alpha_1 = 1$, we have $Ye_0 = Ye_i$; otherwise, Ye_i/α_i and $Y(e_0 - e_i)/(1 - \alpha_i)$ are $(S_i, R^{(m)}, \gamma)$-vectors, where S_i as well as the integers $\{R_w^{(m)}\}_{w \in S_i}$ witnessing that these vectors are $(S_i, R^{(m)}, \gamma)$-vectors are given by Lemma 4.7 for i.

We will complete the proof of the induction step (and hence of Theorem 4.5) by showing in Section 4.4 below that a matrix Y exists satisfying conditions **A** and **B**.

4.3 Proof of Lemma 4.7

Let $\{R_w^{(m+1)}\}_{w \in S}$ witness that α is an $(S, R^{(m+1)}, \gamma)$-vector and let $C = \cup_{w \in S} Ball(w, R_w^{(m+1)})$. There are two cases depending on whether $Ball(i, \frac{2}{\gamma})$ intersects C or not.

In the first (easy) case, $Ball(i, \frac{2}{\gamma})$ does not intersect C. Then let $S_i = S \cup \{i\}$, let $R_i^{(m)} = \frac{2}{\gamma}$, and let $R_w^{(m)} = R_w^{(m+1)}$ for $w \in S$. It is easy to see that the conditions of the lemma are satisfied by these choices.

So consider the second case where $Ball(i, \frac{2}{\gamma})$ does intersects C. Let

$$T_1 = \left\{ w \in S : i \in Ball\left(w, R_w^{(m+1)} + \frac{2}{\gamma} \right) \right\} \quad .$$

That is, T_1 consists of all points in S whose balls, slightly enlarged, contain i. Note that it may be that $i \in S$, in which case $i \in T_1$.

Now let

$$D = \bigcup_{w \in T_1} Ball\left(w, R_w^{(m+1)} + \frac{2}{\gamma} \right) \quad .$$

Since $\sum_{w \in S} (R_w^{(m+1)} + \frac{2}{\gamma}) \le R^{(m+1)} < \frac{1}{2}\,\mathrm{girth}(G) - \frac{2}{\gamma}$, it follows that D is a tree. Let q be a longest path in D and let w_1 be a node in the middle of this path. Then certainly,

$$D \subseteq Ball\left(w_1, \sum_{w \in T_1} \left(R_w^{(m+1)} + \frac{2}{\gamma} \right) \right) \quad .$$

We will now increase the size of this "big ball" around w_1 (perhaps also moving its centre in the process) until there are no points $w \in S$ outside the "big ball" for which $Ball(w, R_w^{(m+1)} + \frac{2}{\gamma})$ intersects the "big ball". We do this as follows:

Suppose $Ball(w_1, \sum_{w\in T_1}(R_w^{(m+1)} + \frac{2}{\gamma}))$ intersects $Ball(w', R_{w'}^{(m+1)} + \frac{2}{\gamma})$ for some $w' \in S\backslash T_1$. Add w' to T_1 and call the new set T_2. Reasoning as before, there exists $w_2 \in G$ such that,

$$\bigcup_{w\in T_2} Ball\left(w, R_w^{(m+1)} + \frac{2}{\gamma}\right) \subseteq Ball\left(w_2, \sum_{w\in T_2}\left(R_w^{(m+1)} + \frac{2}{\gamma}\right)\right) .$$

In general, at stage j if $Ball(w_j, \sum_{w\in T_j}(R_w^{(m+1)} + \frac{2}{\gamma}))$ intersects $Ball(w', R_{w'}^{(m+1)} + \frac{2}{\gamma})$ for some $w' \in S\backslash T_j$, then add w' to T_j, call the new set T_{j+1}, and find a new $w_{j+1} \in G$ (using again the same arguments as before) such that,

$$\bigcup_{w\in T_{j+1}} Ball\left(w, R_w^{(m+1)} + \frac{2}{\gamma}\right) \subseteq Ball\left(w_{j+1}, \sum_{w\in T_{j+1}}\left(R_w^{(m+1)} + \frac{2}{\gamma}\right)\right) .$$

Continue in this way until the first stage k for which no point w' in $S\backslash T_k$ such that $Ball(w', R_{w'}^{(m+1)} + \frac{2}{\gamma})$ intersects $Ball(w_k, \sum_{w\in T_k}(R_w^{(m+1)} + \frac{2}{\gamma}))$. Let $T = T_k$ and $u = w_k$.

We can now define S_i and $\{R_w^{(m)}\}_{w\in S_i}$: Let $S_i = (S\backslash T) \cup \{u\}$. For $w \in S\backslash T$, let $R_w^{(m)} = R_w^{(m+1)}$; let

$$R_u^{(m)} = \frac{2}{\gamma} + \sum_{w\in T}\left(R_w^{(m+1)} + \frac{2}{\gamma}\right) .$$

To complete the proof of the lemma we need to show that α is an $(S_i, R^{(m)}, \gamma)$-vector witnessed by these $\{R_w^{(m)}\}$ and that the remaining two conditions in the statement of the lemma are satisfied.

Note first that

$$\begin{aligned}\sum_{w\in S_i}\left(R_w^{(m)} + \frac{2}{\gamma}\right) &= \sum_{w\in S\backslash T}\left(R_w^{(m)} + \frac{2}{\gamma}\right) + \left(R_u^{(m)} + \frac{2}{\gamma}\right) \\ &= \sum_{w\in S\backslash T}\left(R_w^{(m+1)} + \frac{2}{\gamma}\right) + \left(\sum_{w\in T}\left(R_w^{(m+1)} + \frac{2}{\gamma}\right) + \frac{4}{\gamma}\right) \\ &= \sum_{w\in S}\left(R_w^{(m+1)} + \frac{2}{\gamma}\right) + \frac{4}{\gamma} \le R^{(m+1)} + \frac{4}{\gamma} = R^{(m)} .\end{aligned}$$

The inequality above follows from the fact that α is an $(S, R^{(m+1)}, \gamma)$-vector witnessed by the integers $R_w^{(m+1)}$. Hence α satisfies condition (1) of being an $(S_i, R^{(m)}, \gamma)$-vector witnessed by the integers $\{R_w^{(m)}\}$.

Next note that by construction, $Ball(u, R_u^{(m)})$ does not intersect $\cup_{w\in S\backslash T} Ball(w, R_w^{(m)})$. Moreover, since α is an $(S, R^{(m+1)}, \gamma)$-vector witnessed by the integers $R_w^{(m+1)}$, it follows that

$$Ball(w, R_w^{(m)}) \cap Ball(w', R_{w'}^{(m)}) = \emptyset$$

for distinct $w, w' \in S\backslash T$. Also, by construction, $\alpha_j = \frac{1}{2} + \gamma$ for all $j \notin \cup_{w\in S_i} Ball(w, R_w^{(m)})$. Hence α satisfies conditions (2) and (3) of being an $(S_i, R^{(m)}, \gamma)$-vector witnessed by the integers $\{R_w^{(m)}\}$.

Next note that by construction, we have on one hand that

$$\bigcup_{w\in T} Ball\left(w, R_w^{(m+1)} + \frac{2}{\gamma}\right) \subseteq Ball\left(u, R_u^{(m)} - \frac{2}{\gamma}\right) \ .$$

On the other hand, $Ball(u, R_u^{(m)})$ does not intersect $\cup_{w\in S\setminus T} Ball(w, R_w^{(m+1)})$. Since α is an $(S, R^{(m+1)}, \gamma)$-vector witnessed by the integers $R_w^{(m+1)}$, it thus follows from the definition of such vectors that for all vertices k in $Ball(u, R_u^{(m)}) \setminus Ball(u, R_u^{(m)} - \frac{2}{\gamma})$, we have $\alpha_k = \frac{1}{2} + \gamma$. Hence condition (3) of the lemma holds.

Finally, condition (2) of the lemma holds since by construction $i \in \bigcup_{w\in T} Ball(w, R_w^{(m+1)} + \frac{2}{\gamma})$. The lemma follows.

4.4 Existence of Y

We will show that Y exists by representing conditions **A** and **B** as a linear program and then showing that the program is feasible. This approach was first used in [24] and subsequently in the conference version of this paper.

Our notation will assume symmetry, namely, Y_{ij} will represent $Y_{\{i,j\}}$. Condition **A** requires that:

$$Y_{kk} = \alpha_k, \quad \forall k \in \{1, \ldots, n\} \ . \tag{4.13}$$

Condition **B** requires that the vectors Ye_i/α_i and $Y(e_0 - e_i)/(1-\alpha_i)$ are $(S_i, R^{(m)}, \gamma)$-vectors. In particular, we need constraints on the variables Y_{ij} forcing these vectors to satisfy both the edge constraints as well as the extra structural properties enjoyed by $(S_i, R^{(m)}, \gamma)$-vectors.

The following constraints imply that Ye_i/α_i and $Y(e_0 - e_i)/(1-\alpha_i)$ satisfy the edge constraints: For all $i \in \{1, \ldots, n\}$ and all $\{j,k\} \in E$:

$$\alpha_i \le Y_{ij} + Y_{ik} \le \alpha_i + (\alpha_j + \alpha_k - 1) \ , \tag{4.14}$$

To see that the above inequalities force Ye_i/α_i and $Y(e_0 - e_i)/(1-\alpha_i)$ to satisfy the edge constraints note first that Ye_i/α_i satisfies the edge constraint for some edge $\{j,k\}$ iff the jth and kth coordinates of Ye_i/α_i sum to at least 1. In equations, this requires $Y_{ij}/\alpha_i + Y_{ik}/\alpha_i \ge 1$, or equivalently $\alpha_i \le Y_{ij} + Y_{ik}$ for the edge $\{j,k\}$. Similarly, the equation $Y_{ij} + Y_{ik} \le \alpha_i + (\alpha_j + \alpha_k - 1)$ implies that $Y(e_0 - e_i)/(1-\alpha_i)$ satisfies the edge constraint for edge $\{j,k\}$.

Let (i,t) be a pair of vertices such that $\alpha_i, \alpha_t \notin \{0,1\}$. Let $S_i \subseteq \{1, \ldots, n\}$ be the set, and $\{R_w^{(m)}\}_{w\in S_i}$ the witnesses given by Lemma 4.7 for i. Then i,t are called a *distant pair* if $t \notin \cup_{w\in S_i} Ball(w, R_w^{(m)})$. (Note then that $\alpha_t = \frac{1}{2} + \gamma$.) To ensure that Ye_i/α_i and $Y(e_0 - e_i)/(1-\alpha_i)$ are $(S_i, R^{(m)}, \gamma)$-vectors witnessed by $\{R_w^{(m)}\}_{w\in S_i}$ (as required by condition **B**) it suffices to ensure that the tth coordinates of Ye_i/α_i and $Y(e_0 - e_i)/(1-\alpha_i)$ are $\frac{1}{2} + \gamma$ for all distant pairs (i,t). In particular, for all such pairs,

$$Y_{it} = \alpha_i \alpha_t = \alpha_i(\frac{1}{2} + \gamma) \ . \tag{4.15}$$

Remark 4.8. By Lemma 4.7, distant pairs have the property that every path in G that connects them contains at least $2/\gamma$ consecutive vertices k such that $\alpha_k = \frac{1}{2} + \gamma$. In particular, any such path contains $2/\gamma - 1$ consecutive edges whose endpoints are "oversatisfied" by α by 2γ.

Finally, $(S_i, R^{(m)}, \gamma)$-vectors must lie in $[0,1]^{n+1}$. The following constraints imply that Ye_i/α_i and $Y(e_0 - e_i)/(1-\alpha_i)$ are in $[0,1]^{n+1}$:

$$0 \le Y_{ij} \le \alpha_i, \qquad \forall i,j \in \{1,\dots,n\}, i \neq j \tag{4.16}$$
$$-Y_{ij} \le 1 - \alpha_i - \alpha_j, \qquad \forall i,j \in \{1,\dots,n\}, i \neq j \tag{4.17}$$

Constraints (4.13)–(4.17) suffice to force Y to satisfy conditions **A** and **B**. We will not directly analyze these constraints but instead analyze the following four constraint families which imply constraints (4.13)–(4.17) but are also in a cleaner form:

$$Y_{ij} \le \beta(i,j), \qquad \forall i,j \in \{1,\dots,n\} \tag{4.18}$$
$$-Y_{ij} \le \delta(i,j), \qquad \forall i,j \in \{1,\dots,n\} \tag{4.19}$$
$$Y_{ij} + Y_{ik} \le a(i,j,k), \qquad \forall \{j,k\} \in E \tag{4.20}$$
$$-Y_{ij} - Y_{ik} \le b(i,j,k), \qquad \forall \{j,k\} \in E \tag{4.21}$$

Here (1) $\beta(i,j) = \alpha_i\alpha_j$ if i,j is a distant pair and $\beta(i,j) = \min(\alpha_i, \alpha_j)$ otherwise; (2) $\delta(i,j) = -\alpha_i$ if $i = j$, $\delta(i,j) = -\alpha_i\alpha_j$ if i,j is a distant pair, and $\delta(i,j) = 1 - \alpha_i - \alpha_j$ otherwise; (3) $a(i,j,k) = \alpha_i + (\alpha_j + \alpha_k - 1)$; and (4) $b(i,j,k) = -\alpha_i$. Note that since $\alpha \in [0,1]^{n+1}$, $\beta(i,j) + \delta(i,j) \ge 0$.

To prove the consistency of constraints (4.18)–(4.21), a special combinatorial version of Farkas's lemma will be used similar to that used in [24] and the conference version of this paper [3]. Before giving the exact combinatorial form we require some definitions.

Let $H = (W,F)$ be the graph where $W = \{Y_{ij} : i,j \in \{1,\dots,n\}\}$ (i.e., there is a vertex for each variable $Y_{i,j}$) and the edges F consist of all pairs $\{Y_{ij}, Y_{ik}\}$ such that $\{j,k\} \in E$. Vertices in W labelled Y_{ii} are called *diagonal*. Given an edge $\{Y_{ij}, Y_{ik}\}$ in H, call i its *bracing node* and $\{j,k\} \in E$ its *bracing edge*. An edge $\{i,j\}$ in G is called *overloaded* if $\alpha_i = \alpha_j = \frac{1}{2} + \gamma$. An edge $\{Y_{ij}, Y_{ik}\}$ in H is *overloaded* if its bracing edge is overloaded.

Let p be a walk $v_0, v_1, \dots, v_r$ on H and let $e_1, \dots, e_r$ be the edges in H traversed by this walk. An *alternating sign assignment* (P,N) for p assigns either all the odd or all the even indexed edges of p to the set P with the remaining edges assigned to N. Given an alternating sign assignment (P,N) for p, an endpoint of p is called *positive* (*negative*, respectively) if it is incident to an edge in P (N, respectively). We will be particularly concerned with the positive diagonal endpoints of a walk.

Given a path p in H with an alternating sign assignment (P,N), let

$$S_1^{(p;P,N)} = \sum_{\{Y_{ij},Y_{ik}\} \in P} a(i,j,k) + \sum_{\{Y_{ij},Y_{ik}\} \in N} b(i,j,k) \quad . \tag{4.22}$$

Suppose the endpoints of p are labelled by $Y_{ij}, Y_{k\ell}$. Define $S_2^{(p;P,N)}$ to be $D+E$ where D is $\delta(i,j)$ if Y_{ij} is a positive endpoint and is $\beta(i,j)$ otherwise; and E is $\delta(k,\ell)$ if $Y_{k\ell}$ is a positive endpoint and is $\beta(k,\ell)$ otherwise. Let $S^{(p;P,N)} = S_1^{(p;P,N)} + S_2^{(p;P,N)}$.

$$
\begin{aligned}
-Y_{12} &\le -\alpha_1\alpha_2 \\
Y_{12} + Y_{13} &\le \alpha_1 + (\alpha_2 + \alpha_3 - 1) \\
-Y_{13} - Y_{34} &\le -\alpha_3 \\
Y_{34} + Y_{35} &\le \alpha_3 + (\alpha_4 + \alpha_5 - 1) \\
-Y_{35} - Y_{56} &\le -\alpha_5 \\
Y_{56} &\le \alpha_5\alpha_6 \\
-Y_{77} &\le -\alpha_7 \\
Y_{77} + Y_{78} &\le \alpha_7 + (\alpha_7 + \alpha_8 - 1) \\
-Y_{78} &\le 1 - \alpha_7 - \alpha_8
\end{aligned}
$$

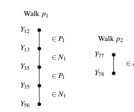

Figure 2: A positive integer linear combination of the constraints where the LHS is 0, and which corresponds to two walks p_1 and p_2 in H with alternating sign assignments (P_1, N_1) and (P_2, N_2), respectively.

Lemma 4.9 (Special case of Farkas's Lemma). *The constraints on the variables Y_{ij} are unsatisfiable iff there exists a walk p on H and an alternating sign assignment (P,N) for p such that $S^{(p;P,N)}$ is negative.*

Proof. Note first that by Farkas's lemma, constraints (4.18)–(4.21) are unsatisfiable iff there exists a positive rational linear combination of them where the LHS is 0 and the RHS is negative.

Now suppose that there exists a path p in H and an alternating sign assignment (P,N) such that $S^{(p;P,N)} < 0$. Consider the following linear integer combination of the constraints: (1) For each edge $\{Y_{ij}, Y_{ik}\} \in p$, if $\{Y_{ij}, Y_{ik}\} \in P$, add the constraint $Y_{ij} + Y_{ik} \le a(i,j,k)$; if $\{Y_{ij}, Y_{ik}\} \in N$, add the constraint $-Y_{ij} - Y_{ik} \le b(i,j,k)$; (2) For each endpoint Y_{ij} of p, if it is a negative endpoint add the constraint $Y_{ij} \le \beta(i,j)$; if it is a positive endpoint add the constraint $-Y_{ij} \le \delta(i,j)$. But then, for this combination of constraints the LHS equals 0 while the RHS equals $S^{(p;P,N)} < 0$. So by Farkas's lemma the constraints are unsatisfiable.

Now assume on the other hand that the constraints are unsatisfiable. So there exists a positive rational linear combination of the constraints such that the LHS is 0 and the RHS is negative. In fact, by clearing out denominators, we can assume without loss of generality that this linear combination has *integer* coefficients. Hence, as $\beta(i,j) + \delta(i,j) \ge 0$ for all i, j, our combination must contain, without loss of generality, constraints of type (4.20) and (4.21). Moreover, since the LHS is 0, for each Y_{ij} appearing in the integer combination there must be a corresponding occurrence of $-Y_{ij}$. But then, it is easy to see that the constraints in the integer linear combination can be grouped into a set of paths $\{p_i\}$ in H each with its own alternating sign assignment such that the RHS of the linear combination equals $\sum S^{(p_i;P_i,N_i)}$ (for an example, see Figure 2). But then, since the RHS is negative, it must be that at least one of the paths p in the set is such that $S^{(p;P,N)} < 0$. The lemma follows. □

So to show that the constraints for the matrix Y are consistent, we will show that $S^{(p;P,N)} \ge 0$ for any walk p on H and any alternating sign assignment (P,N) for p.

To that end, fix a walk p on H and an alternating sign assignment (P,N) for p. To simplify notation we drop the superscript $(p;P,N)$ from $S_1^{(p;P,N)}$, $S_2^{(p;P,N)}$ and $S^{(p;P,N)}$. Let $v_0, v_1, \ldots, v_r$ be the nodes visited by p in H (a node may be visited multiple times) and let $e_1, \ldots, e_r$ be the edges in H traversed by p. We divide our analysis into three cases depending on whether none, one or both endpoints of p are positive diagonal. We will show that in any of these cases $S \geq 0$.

We first note three easy facts about p used below:

Proposition 4.10. *Let C be the subgraph of G induced by the bracing edges for $e_1, \ldots, e_n$. Then,*

1. *Subgraph C consists of at most two connected components;*
2. *If p visits a diagonal node, then C is connected; Moreover, if v_0 is diagonal and $v_r = Y_{st}$, then C contains a path in G from s to t;*
3. *If p visits at least two diagonal nodes then C contain a cycle.*

Proof. We sketch a proof of the first fact; the other two are similar.

Consider the edges $e_1, \ldots, e_r$ in order. As long as the bracing node in successive edges does not change, then the bracing edges of these successive edges form a path p' in G. If the bracing node changes, say at edge e_i in p, the bracing edge for e_i now starts a new path p'' in G. Moreover the last vertex w in G visited by p' is the bracing node for e_i. The bracing edges of the edges following e_i in p now extend p'' in G until an edge e_j is encountered with a new bracing node. But then, the bracing edge for e_j must contain w. Hence, the bracing edge for e_j now extends path p' in G. Continuing this argument we see that each time the bracing node changes we go back and forth from having the bracing edges contributing to the paths p' and p'' in G. Fact (1) follows (also see Figure 3). □

Case 1: No endpoint of p is positive diagonal

Suppose the endpoints v_0, v_r of p are labelled by Y_{ab} and Y_{cd}, respectively, and consider the following sum S_2': If v_0 is a negative endpoint, then it contributes $\alpha_a\alpha_b$ to S_2'; otherwise it contributes $-\alpha_a\alpha_b$. Similarly, if v_r is negative, then it contributes $\alpha_c\alpha_d$ to S_2' and otherwise it contributes $-\alpha_c\alpha_d$. Since $\alpha \in [0,1]^{n+1}$ and neither endpoint is positive diagonal, it follows that $S_2 \geq S_2'$. So to prove that $S \geq 0$ in this case, it suffices to show $S_1 + S_2' \geq 0$.

To that end, consider the following sum:

$$\sum_{\{Y_{ij},Y_{ik}\}\in P} (-\alpha_i\alpha_j - \alpha_i\alpha_k) + \sum_{\{Y_{ij},Y_{ik}\}\in N} (\alpha_i\alpha_j + \alpha_i\alpha_k) \quad . \tag{4.23}$$

By definition of an alternating sign assignment it follows that (4.23) telescopes and equals S_2'. Hence,

$$S \geq S_1 + S_2' = \sum_{\{Y_{ij},Y_{ik}\}\in P} (a(i,j,k) - (\alpha_i\alpha_j + \alpha_i\alpha_k)) + \sum_{\{Y_{ij},Y_{ik}\}\in N} (b(i,j,k) + (\alpha_i\alpha_j + \alpha_i\alpha_k)) \tag{4.24}$$

$$= \sum_{\{Y_{ij},Y_{ik}\}\in P} (1-\alpha_i)(\alpha_j + \alpha_k - 1) + \sum_{\{Y_{ij},Y_{ik}\}\in N} \alpha_i(\alpha_j + \alpha_k - 1) \quad . \tag{4.25}$$

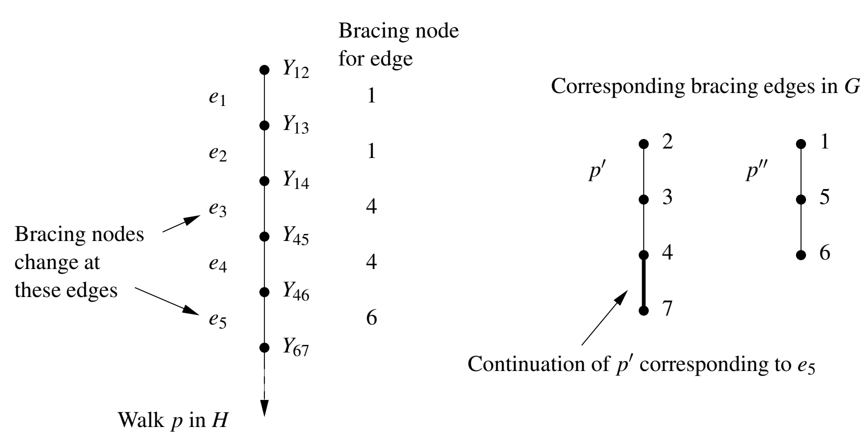

Figure 3: A walk p in H and the corresponding pair of walks p', p'' in G formed by the bracing edges in p. The walks p', p'' could meet, e.g., if p visits a diagonal vertex in H.

Now the bracing edges for all edges in P and N are in G. Moreover, α satisfies the VERTEX COVER edge constraints (4.1) for G. Hence, $\alpha_j + \alpha_k \geq 1$ for all edges $\{Y_{ij}, Y_{ik}\} \in P \cup N$. But then, since we always have $0 \leq \alpha_i \leq 1$, it follows that all summands in (4.25) are at least 0 and hence, $S \geq 0$ as desired.

Case 2: One endpoint of p is positive diagonal

Assume without loss of generality that v_0 is the positive endpoint and is labelled Y_{11}, and suppose the other endpoint v_r is labelled Y_{st}. There are two subcases:

Subcase 1: $\{s,t\}$ is a distant pair: By Proposition 4.10, if C is the subgraph of G induced by the bracing edges for $e_1, \ldots, e_n$, then there is a path p' in C (and hence in G) from s to t. So since s,t are distant, Remark 4.8 implies that p' contains at least $2/\gamma - 1$ *consecutive* overloaded edges.

We first define some notation to refer to the summands appearing in (4.25) which will also be important in this subcase: For an edge $e = \{Y_{ij}, Y_{ik}\}$ in our path p,

$$\zeta(e) = \begin{cases} (1-\alpha_i)(\alpha_j + \alpha_k - 1), & \text{if } \{Y_{ij}, Y_{ik}\} \in P \\ \alpha_i(\alpha_j + \alpha_k - 1), & \text{if } \{Y_{ij}, Y_{ik}\} \in N \end{cases}$$

As noted in Case 1, $\zeta(e) \geq 0$ for all $e \in p$.

In Case 1 we showed that $S \geq 0$ by first defining a sum S_2' such that $S_2 \geq S_2'$ and then noting that $S_1 + S_2' = \sum_{e \in p} \zeta(e)$. Unfortunately, in the current subcase, since p contains a positive diagonal endpoint, it is no longer true that $S_2 \geq S_2'$. However, it is easy to see that $S_2 \geq S_2' - (\alpha_1 - \alpha_1^2)$. In particular, $S \geq \sum_{e \in p} \zeta(e) - (\alpha_1 - \alpha_1^2)$ for the current subcase. So since $\zeta(e) \geq 0$ always, to show that $S \geq 0$ in the current subcase, it suffices to show that for "many" edges e in p, $\zeta(e)$ is "sufficiently large" so that $\sum_{e \in p} \zeta(e) \geq \alpha_1 - \alpha_1^2$. The existence of these edges in p will follow from the existence of the $2/\gamma - 1$ consecutive overloaded edges in p'.

Assume without loss of generality that $2/\gamma - 1 = 4q$ for some integer q and let $f_1, \ldots, f_{4q}$ be, in order, the $4q$ consecutive overloaded edges in p' (recall that p' is the path from s to t in G and defined by the bracing edges of p). Let $U = \{e_{i_1}, \ldots, e_{i_{4q}}\}$ be the set of edges in p whose bracing edges correspond to $f_1, \ldots, f_{4q}$ (where e_{i_j} corresponds with f_j). Note that the edges in U need not occur consecutively in p. However, using arguments similar to those used in Proposition 4.10 we can prove the following fact:

Fact 4.11. The edges of p' can be divided into two consecutive walks p_1' and p_2' (i.e., all edges in p_1' and p_2' are consecutive and all edges in p_2' either all occur before or after all edges in p_1') such that if $U_i \subseteq U$ denotes the edges of p whose bracing edges form the walk p_i', then the order in p of the edges U_1 is the same as the order of the corresponding bracing edges in p_1', while the order in p of the edges U_2 is the reverse of the order of the corresponding bracing edges in p_2'.

Example 4.12. Suppose $p = Y_{11}$-Y_{12}-Y_{13}-Y_{16}-Y_{46}-Y_{56}. The corresponding walk p' is 5-4-1-2-3-6 and the division guaranteed by the above Fact has $p_1' = $ 1-2-3-6, $p_2' = $ 5-4-1.

Let p_1', p_2' be the division of p' and U_1, U_2 the corresponding subsets of U for these paths, respectively, guaranteed by Fact 4.11 for p'. Without loss of generality, assume that the length of p_1' is at least $2q$. In particular, assume without loss of generality that $i_1 < \cdots < i_{2q}$. (If instead p_2' has length greater than $2q$, then we assume without loss of generality that $i_{2q+1} > \cdots > i_{4q}$ and the arguments below are modified accordingly.)

Let $B = \{1,3,5,\ldots,2q-1\}$. Fix some $j \in B$ and consider the pair $e_{i_j}, e_{i_{j+1}}$ of edges from U. Suppose $e_{i_j} = \{Y_{ab}, Y_{ac}\}$, $e_{i_{j+1}} = \{Y_{uv}, Y_{uw}\}$ where $u \neq a$. Since the bracing edges for these two edges are consecutive in p', all edges e_ℓ such that $i_j < \ell < i_{j+1}$ have the same bracing node (say x) and moreover, this bracing node is different from the bracing nodes in e_{i_j} or $e_{i_{j+1}}$. So we have $x = c = v$ (Figure 4).

Let $Z_j = \sum \zeta(e)$, where the sum is over $e \in \{e_{i_j}, e_{i_j+1}, e_{i_j+2}, \ldots, e_{i_{j+1}-1}, e_{i_{j+1}}\}$ (i.e., over the edges e_{i_j} and $e_{i_{j+1}}$, and all edges between them in p).

Claim 4.13. $Z_j \geq 2\gamma/3$.

Since $j \in B$ was arbitrary and $|B| = q$, the claim implies $S_1 + S_2' \geq q(2\gamma/3) \geq 1/3 - \gamma/6$. So since $\gamma < 1/2$ and $\alpha_1 - \alpha_1^2 \leq \frac{1}{4}$ for $\alpha_1 \in [0,1]$, it follows that $S_1 + S_2' \geq \alpha_1 - \alpha_1^2$, completing the proof that $S \geq 0$ in this subcase.

Proof of Claim 4.13. Suppose $d = i_{j+1} - i_j - 1$ is odd (the case where d is even is similar). Moreover, assume that $e_{i_j}, e_{i_{j+1}} \in P$ (the case where they are both in N is similar). Let $\alpha_a + \alpha_u = 1 + D$. Since e_{i_j} and $e_{i_{j+1}}$ are overloaded,

$$\zeta(e_{i_j}) + \zeta(e_{i_{j+1}}) = 2\gamma(2 - \alpha_u - \alpha_a) = 2\gamma(1-D) \ . \tag{4.26}$$

If $D \leq \frac{2}{3}$, then (4.26) is greater than $2\gamma/3$, and hence so is Z_j. So assume $D > \frac{2}{3}$.

Note that the bracing edges of $e_{i_j+1}, e_{i_j+2}, \ldots, e_{i_{j+1}-1}$ form a path p'' from a to u of length d in G. Let $g_1, \ldots, g_{d+1}$ be the nodes on p'' where $g_1 = a$, $g_{d+1} = u$ (Figure 4). Since α satisfies the VERTEX COVER edge constraints (4.1) for G, $\sum_{k=1}^{d+1} \alpha_{g_k} \geq (d+1)/2$. In fact, we must have that $\sum_{k=1}^{d+1} \alpha_{g_k} \geq (d+1)/2 + D$ (this just says that since the endpoints of p'' sum to $1 + D$ then some edge(s) along p'' must be oversatisfied by D). But then,

$$Z_j \geq \sum_{k=1}^{(d+1)/2} \zeta(e_{i_j+2k-1}) = \alpha_c \left(\sum_{k=1}^{d+1} \alpha_{g_k} - \frac{d+1}{2} \right) \geq \left(\frac{1}{2} + \gamma \right) D > \frac{2\gamma}{3} \ .$$

□

Subcase 2: $\{s,t\}$ is not a distant pair: Let S_{st} be the contribution of Y_{st} to S_2 (i.e., $S_{st} = \delta(s,t)$ if v_r is a positive endpoint and $S_{st} = \beta(s,t)$ if v_r is negative). Since the contribution of Y_{11} to S_2 is $-\alpha_1$, it follows that $S_2 = S_{st} - \alpha_1$.

For an edge $e_\ell = \{Y_{ij}, Y_{ik}\} \in p$, let

$$T_\ell = \begin{cases} a(i,j,k), & \text{if } e_\ell \in P \\ b(i,j,k), & \text{if } e_\ell \in N \end{cases}$$

Recall that $v_0, v_1 \ldots, v_r$ are the nodes visited by the walk p and that e_i denotes the edge traversed between v_{i-1} and v_i. Note then that $S_1 = \sum_{\ell=1}^{r} T_\ell$. Moreover, recall that we have assumed without loss of generality that $v_0 = Y_{11}$ and $v_r = Y_{st}$. So since $e_1 \in P$, the following claim implies $S_{st} + \sum_{\ell=1}^{r} T_\ell \geq \alpha_1$, and hence that $S \geq 0$ in this subcase.

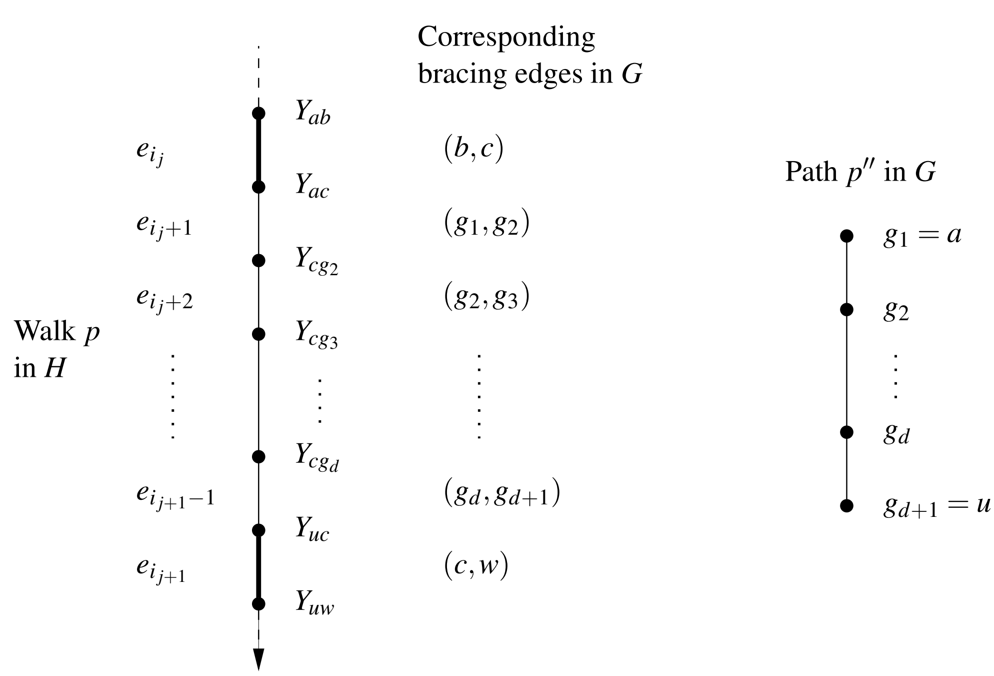

Figure 4: A portion of a walk p in H in which the bracing node (in this case c) does not change between edges $e_{i_j}, e_{i_{j+1}}$, together with the path p'' of bracing edges in G for the portion of p with bracing node c.

Claim 4.14. *Let $1 \le q \le r$ and suppose $v_{q-1} = Y_{ij}$, $v_q = Y_{ik}$ (i.e., $e_q = \{Y_{ij}, Y_{ik}\}$). Then $S_{st} + \sum_{\ell=q}^{r} T_\ell$ is at least $\min(\alpha_i, \alpha_j)$ if $e_q \in P$ and is at least $\min(0, 1 - \alpha_i - \alpha_j)$ if $e_q \in N$.*

Proof. By "backward" induction on q. For the base case $q = r$, assume without loss of generality that $v_{q-1} = Y_{sj}$, so that $e_q = \{Y_{sj}, Y_{st}\}$. If $e_q \in N$, then $T_1 = -\alpha_s$ so that $T_1 + S_{st} = -\alpha_s + \min(\alpha_s, \alpha_t)$. Since α satisfies the edge constraints (4.1), it follows that $\alpha_j + \alpha_t \ge 1$ for the bracing edge $\{j, t\}$. Hence, $T_1 + S_{st} \ge \min(0, 1 - \alpha_j - \alpha_s)$. If instead $e_q \in P$, then $T_1 = \alpha_s + (\alpha_j + \alpha_t - 1)$ so that

$$T_1 + S_{st} = [\alpha_s + (\alpha_j + \alpha_t - 1)] + (1 - \alpha_s - \alpha_t) = \alpha_j \ .$$

The base case $q = r$ follows.

Assume the claim holds for e_q and consider $e_{q-1} = \{Y_{ij}, Y_{ik}\}$ where $v_{q-2} = Y_{ij}$ and $v_{q-1} = Y_{ik}$. If $e_{q-1} \in N$, then $T_{q-1} = -\alpha_i$. Moreover, $e_q \in P$ and by induction,

$$S_{st} + \sum_{\ell=q}^{r} T_\ell \ge \min(\alpha_i, \alpha_k) \ .$$

Since α satisfies the edge constraints (4.1), it follows that $\alpha_j + \alpha_k \ge 1$ for the bracing edge $\{j, k\}$. Hence,

$$S_{st} + \sum_{\ell=q-1}^{r} T_\ell \ge \min(0, 1 - \alpha_i - \alpha_j) \ .$$

If instead $e_{q-1} \in P$, then $T_{q-1} = \alpha_i + (\alpha_j + \alpha_k - 1)$. Moreover, $e_q \in N$ and by induction,

$$S_{st} + \sum_{\ell=q}^{r} T_\ell \ge \min(0, 1 - \alpha_i - \alpha_k) \ .$$

So since $\alpha_j + \alpha_k \ge 1$, it follows that

$$S_{st} + \sum_{\ell=q-1}^{r} T_\ell \ge \min(\alpha_i, \alpha_j) \ .$$

The claim follows for e_{q-1}. ☐

Case 3: Both endpoints of p are positive diagonal

Since p contains two diagonal vertices, Proposition 4.10 implies that there is a cycle C in the subgraph of G induced by the bracing edges corresponding to the edges in p. Since $\mathrm{girth}(G) \ge 4R^{(m)}$, it follows that C contains a distant pair. But then, as there are two different paths between this pair along C, Remark 4.8 implies that there are two subpaths p'_1 and p'_2 in C *each* consisting of $2/\gamma$ overloaded edges.

Recall that in subcase 1 of Case 2 where there was one positive diagonal vertex, one such subpath was used to argue that $S \ge 0$ in that subcase. In the current case where there are *two* positive diagonals and the *two* subpaths p'_1 and p'_2, the same argument then implies that $S \ge 0$ for the current case also.

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 $\log n$ rounds but the argument seems to need some property other than high girth.

As mentioned in our related work section, since the appearance of the conference version of this paper, a few other papers [6, 1, 30] have addressed questions introduced here. However, the techniques in all the above papers do not seem to apply to graph VERTEX COVER. Furthermore, they also do not apply to a lift-and-project method of Sherali-Adams [28] that was contemporaneous to Lovász-Schrijver.

Acknowledgements

The first author would like to thank Noga Alon for drawing his attention to the Erdős results of [2, 13] and their tightenings, and to Subhash Khot for useful discussions.

We would also like to thank Joseph Cheriyan and Fei Qian for letting us know about the error in the proof of the *LS* lower bound in the conference version of this paper.

Finally we would like to thank the anonymous referees for their careful reading of our manuscript and for their many suggestions on improving this paper.

References

[1] * Mikhail Alekhnovich, Sanjeev Arora, and Iannis Tourlakis: Towards strong nonapproximability results in the Lov ász-Schrijver hierarchy. In *Proceedings of the 37th Annual Symposium on the Theory of Computing*, New York, May 2005. ACM Press. [STOC:1060590.1060634]. 1, 1, 5

[2] * Noga Alon and Joel H. Spencer: *The probabilistic method*. Wiley, New York, 1992. 2, 2.1, 5

[3] * Sanjeev Arora, Béla Bollobás, and László Lovász: Proving integrality gaps without knowing the linear program. In *Proceedings of the 43rd Symposium on Foundations of Computer Science (FOCS-02)*, pp. 313–322, Los Alamitos, November 16–19 2002. [FOCS:10.1109/SFCS.2002.1181954]. 2, 3, 4.4

[4] * Béla Bollobás: *Modern graph theory*, volume 184 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1998. 4

[5] * Maria Bonet, Toniann Pitassi, and Ran Raz: Lower bounds for cutting planes proofs with small coefficients. *The Journal of Symbolic Logic*, 62(3):708–728, September 1997. 1

[6] * JOSHUA BURESH-OPPENHEIM, NICOLA GALESI, SHLOMO HOORY, AVNER MAGEN, AND TONIANN PITASSI: Rank bounds and integrality gaps for cutting planes procedures. In *Proceedings: 44th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2003, 11–14 October 2003, Cambridge, Massachusetts*, pp. 318–327. pub-IEEE, 2003. 1, 1, 4, 5

[7] * JOSEPH CHERIYAN AND FEI QIAN: Personal communication, 2005. 2

[8] * VASEK CHVÁTAL: Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Math.*, 4:305–337, 1973. 1

[9] * WILLIAM COOK AND SANJEEB DASH: On the matrix-cut rank of polyhedra. *Mathematics of Operations Research*, 26(1):19–30, 2001. 1

[10] * IRIT DINUR, VENKATESAN GURUSWAMI, SUBHASH KHOT, AND ODED 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*, pp. 595–601, New York, NY, USA, 2003. ACM Press. [STOC:780542.780629]. 1

[11] * IRIT DINUR AND SHMUEL SAFRA: On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1), 2005. 1

[12] * PAUL ERDÖS: Graph theory and probability. *Canadian Journal of Mathematics*, 11:34–38, 1959. 4

[13] * PAUL ERDÖS: On circuits and subgraphs of chromatic graphs. *Mathematika*, 9:170–175, 1962. 2, 5

[14] * URIEL FEIGE: Randomized graph products, chromatic numbers, and the Lovász ϑ-function. *Combinatorica*, 17(1):79–90, 1997. [Combinatorica:x785787h43724566]. 2.2, 2.15

[15] * URIEL FEIGE: A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998. [JACM:285055.285059]. 1

[16] * URIEL FEIGE AND ROBERT KRAUTHGAMER: The probable value of the Lovász–Schrijver relaxations for maximum independent set. *SIAM Journal on Computing*, 32(2):345–370, 2003. [SICOMP:10.1137/S009753970240118X]. 1

[17] * MICHEL X. GOEMANS: Worst-case comparison of valid inequalities for the TSP. *Mathematical Programming*, 69:335–349, 1995. [MathProg:n25193135v44l625]. 1

[18] * MICHEL X. GOEMANS AND LEVENT TUNÇEL: When does the positive semidefiniteness constraint help in lifting procedures? *Mathematics of Operations Research*, 26(4):796–815, 2001. 1

[19] * MICHEL X. GOEMANS AND DAVID P. WILLIAMSON: .878-approximation algorithms for MAX CUT and MAX 2SAT. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing, STOC'94 (Montréal, Québec, Canada, May 23-25, 1994)*, pp. 422–431, New York, 1994. ACM Press. [STOC:195058.195216]. 1

[20] * JOHAN HÅSTAD: Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001. [JACM:258533.258536]. 1

[21] * DORIT HOCHBAUM: Approximating covering and packing problems: Set cover, vertex cover, independent set, and related problems. In *Approximation Algorithms for NP-hard Problems*. PWS Publishing Company, 1997. 1

[22] * Dorit Hochbaum, editor. *Approximation Algorithms for NP-hard Problems*. PWS Publishing Company, 1997. 1

[23] * LÁSZLÓ LIPTÁK AND LÁSZLÓ LOVÁSZ: Critical facets of the stable set polytope. *Combinatorica*, 21(1):61–88, 2001. [Combinatorica:8hmkpcvb5qaw5dnv]. 1, 3

[24] * LÁSZLÓ LOVÁSZ AND ALEXANDER SCHRIJVER: Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, May 1991. [SIOPT:01/0801013]. 1, 3, 4, 4, 4.1, 4.4, 4.4

[25] * CHRISTOS H. PAPADIMITRIOU AND SANTOSH VEMPALA: On the approximability of the Traveling Salesman problem. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing, STOC'2000 (Portland, Oregon, May 21-23, 2000)*, pp. 126–133, New York, 2000. ACM Press. [STOC:335305.335320]. 1

[26] * FEI QIAN: The integrality gap of the minimum vertex cover problem. Master's thesis, University of Waterloo, 2003. 2

[27] * ALEXANDER 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. 1

[28] * HANIF D. SHERALI AND WARREN 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, August 1990. [SIDMA:03/0403036]. 5

[29] * DAVID B. SHMOYS: Cut problems and their application to divide-and-conquer. In *Approximation Algorithms for NP-hard Problems*. PWS Publishing Company, 1997. 1

[30] * IANNIS TOURLAKIS: Towards optimal integrality gaps for hypergraph vertex cover in the Lovász-Schrijver hierarchy. In *8th. International Workshop on Approximation Algorithms for Combinatorial Optimization Problems and 9th International Workshop on Randomization and Computation*, pp. 233–244, 2005. 1, 1, 5

[31] * VIJAY V. VAZIRANI: *Approximation Algorithms*. Springer-Verlag, Berlin, 2001. 1

[32] * LAURENCE A. WOLSEY: Heuristic analysis, linear programming, and branch and bound. *Mathematical Programming Studies*, 13:121–134, 1980. 1

[33] * MIHALIS YANNAKAKIS: Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, December 1991. [JCSS:10.1016/0022-0000(91)90024-Y]. 1

AUTHORS

Sanjeev Arora [About the author]
professor
Department of Computer Science
Princeton University
Princeton, NJ 08544
arora@cs.princeton.edu
http://www.cs.princeton.edu/~arora

Béla Bollobás [About the author]
professor
University of Memphis, TN 38152, U.S.A.,
and Trinity College, Cambridge, U.K.
bollobas@msci.memphis.edu and
b.bollobas@dpmms.cam.ac.uk

László Lovász [About the author]
senior research associate
Microsoft Research One Microsoft Way, Redmond, WA 98052-6399
lovasz@microsoft.com

Iannis Tourlakis [About the author]
graduate student
Department of Computer Science
Princeton University
Princeton, NJ 08544
itourlak@cs.princeton.edu
http://www.cs.princeton.edu/~itourlak

About the Authors

Sanjeev Arora obtained his Ph.D. from the University of California at Berkeley in 1994 under the supervision of Umesh Vazirani. He has been at Princeton University since 1994 where he is professor of Computer Science. His research area is Theoretical Computer Science, specifically, Computational Complexity, uses of randomness in computation, Probabilistically Checkable Proofs (PCPs), computing approximate solutions to NP-hard problems, and geometric embeddings of metric spaces. He was a co-winner of the 1995 ACM Dissertation award. For his work on PCPs he was a co-winner of the 2001 Gödel Prize. He spends his spare time with his two kids, and doing black and white photography in his darkroom.

Béla Bollobás is a student of Paul Erdős, and received doctorates from Budapest, Hungary, and Cambridge, England. He has been in Cambridge since 1969, where he is a *Senior Research Fellow* of Trinity College, and for the past decade he has held the *Jabie Hardin Chair of Excellence in Graph Theory and Combinatorics* at the University of Memphis. He works mostly in Combinatorics, Graph Theory and Percolation Theory, and has published over 300 research papers and eight books, including *Extremal Graph Theory*, *Random Graphs*, and *Modern Graph Theory*. He has had over thirty Ph.D. students. He is a Foreign Member of the Hungarian Academy of Sciences, and was an invited speaker at the ICM in 1998.

László Lovász received his doctorate from the Eötvös Loránd University in Budapest, Hungary, where his advisor was Tibor Gallai. He is a member of the Hungarian Academy of Sciences. He received the George Pólya Prize of the Society for Industrial and Applied Mathematics (1979), the Delbert Ray Fulkerson Prize of the American Mathematical Society and the Mathematical Programming Society (1982), the Brouwer Medal of the Dutch Mathematical Society (1993) and the Wolf Prize (1999). He is editor-in-chief of Combinatorica and editor of 12 other Journals. His field of research is mainly in combinatorial optimization, algorithms, complexity, and random walks on graphs. He has written 4 research monographs and 3 textbooks, and about 200 research papers.

Iannis Tourlakis obtained his B.Sc. in 1998 and M.Sc. in 2000 from the University of Toronto where his M.Sc. advisor was Stephen Cook. He is currently completing his Ph.D. at Princeton University under the supervision of Sanjeev Arora. When not thinking about research (mainly lower bounds in convex optimization as well as problems in derandomization) he likes to fool around with whatever musical instrument he can get his hands on, preferably a violin or piano. Failing that, like a true Canadian he'll settle for playing or watching some hockey.