Sherali-Adams relaxations of the matching polytope

Claire Mathieu[*] Alistair Sinclair[†]

November 17, 2008

Abstract

We study the Sherali-Adams lift-and-project hierarchy of linear programming relaxations of the matching polytope. Our main result is an asymptotically tight expression $1 + 1/k$ for the integrality gap after k rounds of this hierarchy. The result is derived by a detailed analysis of the LP after k rounds applied to the complete graph K_{2d+1}. We give an explicit recurrence for the value of this LP, and hence show that its gap exhibits a "phase transition," dropping from close to its maximum value $1 + \frac{1}{2d}$ to close to 1 around the threshold $k = 2d - \sqrt{d}$. We also show that the *rank* of the matching polytope (i.e., the number of Sherali-Adams rounds until the integer polytope is reached) is exactly $2d - 1$.

[*]Computer Science Department, Brown University, Providence RI 02912, U.S.A. Email: claire@cs.brown.edu.

[†]Computer Science Division, University of California, Berkeley CA 94720-1776, U.S.A. Email: sinclair@cs.berkeley.edu.

1 Introduction

Background. Recent years have seen an explosion of interest in hierarchies of linear or semidefinite relaxations of 0-1 integer programs, such as those due to Sherali and Adams [26], Balas, Ceria and Cornuejols [5], Lovász and Schrijver [22] and Lasserre [19, 20]. (For an excellent discussion and comparison of these methods, see the article of Laurent [21].) Given a convex polytope $P_0 \subseteq R^n$, the goal is to maximize a linear function f over the associated integer polytope $P = \mathrm{conv}(P_0 \cap \{0,1\}^n$. The above methods construct a sequence $P_0 \supseteq P_1 \supseteq P_2 \supseteq \cdots \supseteq P_n = P$ of successive relaxations of P such that the nth relaxation P_n is equal to P. The relaxations are either linear or (in the case of Lasserre and one variant of Lovász and Schrijver) semidefinite, and (under suitable assumptions about P) have the property that that a linear function can be optimized over P_k in time $n^{O(k)}$, which is polynomial for any fixed k. The relaxations are constructed by "lifting" the current P_k to a higher dimensional space, tightening it by adding further linear or semidefinite constraints that are satisfied by all 0-1 vectors, and then projecting back down to R^n. For this reason, the methods are often referred to as "lift-and-project" algorithms.

Interest in these methods has come from at least three distinct communities. First, in polyhedral combinatorics, the structure of the successive relaxations P_k is of intrinsic interest. In particular, one may naturally ask about the *rank* of P, i.e., the minimum number of rounds k for which $P_k = P$, or the rank of any particular linear inequality known to be satisfied by P. Second, in computational complexity there has recently been a substantial sequence of results proving for several classical combinatorial problems that, even for $k = \Omega(n)$, the kth relaxation P_k has a large integrality gap.[1] The motivation for these results is that the various lift-and-project schemes encompass most known sophisticated approximation algorithms for NP-hard problems such as Sparsest Cut and Maximum Satisfiability; therefore, a large integrality gap after a linear (or even logarithmic) number of rounds rules out (unconditionally) a very wide class of efficient approximation algorithms. Third, in the area of proof complexity, the various hierarchies can be viewed as sequences of proof systems with the goal of proving that the integer polytope P is empty (which may be equivalent to, e.g., showing that a given formula is unsatisfiable). The inclusion of new constraints corresponds to the derivation of new inequalities from previous ones in the proof. Again, the power of the proof system is related to the properties of the relaxation P_k after k rounds. We briefly summarize the relevant literature on these three directions in the Related Work section below.

Results. In this paper, we study the integrality gap of the Sherali-Adams hierarchy for the classical *matching polytope*. The Sherali-Adams scheme, in addition to being the earliest of the lift-and-project methods, is also the strongest of the linear versions and has a particularly simple description as well as certain other advantages (see [21]). As is well known, the matching polytope is defined for any finite graph $G = (V, E)$ by the variables $\{x_1, \ldots, x_{|E|}\}$ and constraints $0 \le x_e \le 1$ and $\sum_{e:u\in e} x_e \le 1$ for all $u \in V$; the goal is to maximize $f(x) = \sum_e x_e$. The kth Sherali-Adams relaxation P_k is obtained by multiplying each of these constraints by a multiplier of the form $\prod_{e\in I} x_e \prod_{f\in J}(1 - x_f)$ for disjoint subsets $I, J \subseteq E$ with $|I \cup J| = k$, linearizing the resulting monomials by introducing new variables, and projecting back down to $|E|$ dimensions. Our main result is a precise estimate for the integrality gap after k rounds as a function of k, which is tight up to lower order terms. This is expressed in the following theorem:

Theorem 1.1 *As k tends to infinity, the integrality gap of the kth round of the Sherali-Adams hierarchy for maximum matching is $\alpha_k = 1 + 1/k + o(1/k)$.*

Theorem 1.1 follows from a detailed analysis of the sequence of relaxations P_k applied to complete graphs K_{2d+1} of odd cardinality; it is easy to see that, for any k, the integrality gap is always

[1] The integrality gap is the ratio between the optimum of f over P_k and the optimum of f over P.

attained on such a graph. More precisely, we study the *integrality ratio* $g_k \equiv g_k(K_{2d+1})$, i.e., the ratio of the value of the kth Sherali-Adams relaxation applied to K_{2d+1} to that of the optimum (which is clearly d, the size of a maximum matching in K_{2d+1}). For the standard LP relaxation P_0 this value is well known to be $g_0 = 1 + 1/2d$. We show first that it remains at exactly this value for $0 \le k \le d-1$, and also that it reaches 1 when $k = 2d-1$. In other words, the Sherali-Adams relaxations make no progress in the first $d-1$ rounds, and achieve the integer optimum after $2d-1$ rounds. Between these two extremes, we observe a perhaps surprising behavior: g_k exhibits a "phase transition" in that it switches suddenly from close to its maximum value $1 + 1/2d$ to close to 1 in the neighborhood of the threshold $k = 2d - \sqrt{d}$. The following theorem makes this statement precise:

Theorem 1.2

(i) *If* $k < d$ *then* $g_k(K_{2d+1}) = 1 + 1/2d$.

(ii) *If* $d \le k \le 2d - \omega(d^{1/2})$ *then* $1 + 1/2d - o(1/d) \le g_k(K_{2d+1}) \le 1 + 1/2d$.

(iii) *If* $2d - o(d^{1/2}) \le k < 2d-1$ *then* $1 \le g_k(K_{2d+1}) \le 1 + o(1/d)$.

(iv) *If* $k \ge 2d-1$ *then* $g_k(K_{2d+1}) = 1$.

Theorem 1.1 follows easily from this result and the fact that α_k is non-increasing. However, Theorem 1.2 carries more detailed information about the Sherali-Adams hierarchy. Our analysis also shows as a byproduct that the integrality ratio is strictly larger than 1 for $k < 2d-1$, which implies that the *rank* of the matching polytope (i.e., the number of Sherali-Adams rounds needed to reach the integer polytope) is exactly $2d-1$.

Theorem 1.3 *For* $n = 2d+1$*, the Sherali-Adams rank of the matching polytope, in the worst case over all* n*-vertex graphs, is* $2d-1$.

Theorem 1.3 answers for the Sherali-Adams hierarchy a question initially posed by Lovász and Schrijver about the rank of the matching polytope in the LS_+ hierarchy, which was answered by Stephen and Tunçel [27].

Our analysis proceeds by showing that, for each k, the Sherali-Adams constraints on K_{2d+1} are all captured by a much simpler family of multipliers of the form $\prod_{e\in I} x_e \prod_{f\in J}(1-x_f)$, where I is a *matching* and J is a *star* disjoint from I. (We call these "standard multipliers.") This simplification allows us to explicitly write down the Sherali-Adams linear program for any k (see Theorem 4.1), and then to express its solution exactly in the form of a recurrence relation (Lemma 4.1). While this recurrence does not appear to have a closed-form solution, we are able to bound its value quite tightly and hence show that it has the behavior claimed in Theorem 1.2.

Related work. The various lift-and-project hierarchies are placed in a common framework and compared by Laurent [21], who shows among other things that the Sherali-Adams hierarchy is stronger (i.e., gives a tighter relaxation at any given level) than LS (the linear programming version of the Lovász-Schrijver hierarchy) but incomparable with LS_+ (i.e., LS with added semidefinite constraints); the Lasserre hierarchy is stronger than all the others.

The matching polytope was first studied in the lift-and-project context by Lovász and Schrijver [22], who posed the problem of determining the *rank* (i.e., the minimum number of rounds until the integer polytope is reached) for complete graphs K_n. For $n = 2d+1$, they showed that the rank lies between $2d$ and $2d^2 - 1$ in the LS hierarchy, and is at most d in the LS_+ hierarchy. Stephen and Tunçel [27] subsequently proved that the LS_+-rank is exactly d, and Goemans and Tunçel [15] improved the upper bound on LS-rank to d^2. Aguilera, Bianchi and Nasini [1] show that the LS-rank is strictly larger than d, and also that the rank in the weaker Balas-Ceria-Cornuéjols hierarchy is exactly d^2. We note that these results say very little about the Sherali-Adams hierarchy (other than the weak upper bound of d^2 on the rank inherited from LS), and do not directly address the more detailed question of how the integrality gap behaves as a function of k.

A question similar to ours, but for a different problem and for the LS_+ hierarchy, was asked by Feige and Krauthgamer [12]. They consider the independent set problem on a *random* graph $G \in \mathcal{G}_{n,1/2}$, and show that the value of the SDP relaxation after k rounds of LS_+ is almost surely about $\sqrt{n/2^k}$.

Arora *et al.* [3, 4] were the first to propose using lift-and-project hierarchies as a model of computation, in order to obtain strong evidence for the hardness of approximating optimization problems. They showed in particular that the integrality gap for Vertex Cover remains at least $2-\varepsilon$ after $\Omega_\varepsilon(\log n)$ rounds of LS. Since then there has been a flurry of activity, proving larger gaps after fewer rounds for Vertex Cover and several other classical NP-hard optimization problems; see, e.g., [2, 8, 9, 10, 13, 14, 24, 25, 28]. Most of this work has focused on the LS and LS_+ hierarchies; exceptions are [10, 13], which consider Sherali-Adams, and [24] which considers Lasserre. We mention also the recent work of Chlamtac [11], who uses the Lasserre hierarchy explicitly to derive improved approximation algorithms for coloring and independent set in 3-uniform hypergraphs.

Finally, we briefly mention a third body of work that views the lift-and-project hierarchies as proof systems. A recent paper of Pitassi and Segerlind [23] proves exponential size lower bounds for tree-like LS_+ proofs of unsatisfiability for several important classes of CNFs, and also shows that tree-like LS_+ proofs cannot efficiently simulate certain other standard proof systems. This differs from the aforementioned work in that the lower bounds are for the *size* of the proofs rather than for the rank (which corresponds to depth in the tree-like scenario). It also extends earlier results by Buresh-Oppenheim *et al.* [8] on rank lower bounds, and builds on work of Grigoriev *et al.* [16] and Kojevnikov and Itsykson [18] that proves lower bounds for LS_+ indirectly via the more powerful but complex proof system known as static positivstellensatz refutations.

2 Preliminaries

2.1 The Sherali-Adams hierarchy

We recall the definition of the Sherali-Adams hierarchy of progressively stronger relaxations of integer polytopes [26, 21]. Let $P_0 = L_0 = \{x \in R^n : \forall \ell, 1 \le \ell \le m, a_\ell \cdot x \ge b_\ell\}$ be a convex polytope contained in $[0,1]^n$, defined by m linear constraints, and let $P = \text{conv}(P_0 \cap \{0,1\}^n)$ be the associated 0-1 polytope. Starting from L_0, the Sherali-Adams method constructs a hierarchy of progressively stronger linear relaxations $P_0, P_1, P_2, \cdots$ of P. For $k \ge 1$, the kth iterate P_k in the Sherali-Adams hierarchy is obtained as follows.

First, we multiply each constraint $a_\ell \cdot x - b_\ell \ge 0$ by each product $\prod_{i\in I} x_i \prod_{j \in J}(1-x_j)$ where I, J are disjoint subsets of $\{1, \ldots, n\}$ such that $|I \cup J| = k$, to produce a set of polynomial inequalities. Add to this set all the inequalities $\prod_{i\in I} x_i \prod_{j\in J}(1 - x_j) \ge 0$ where I, J are disjoint subsets of $\{1, \ldots, n\}$ such that $|I \cup J| = \min(k + 1, n)$.

Then, we replace each square x_i^2 by x_i so that each expression is multilinear, and linearize each product monomial $\prod_{\ell \in L} x_\ell$ by replacing it with a new variable y_L (thus $y_{\{i\}} = x_i$): this defines a new, lifted polyhedron[2] L_k in higher-dimensional space R^d, $d = \binom{n}{1} + \cdots + \binom{n}{k+1}$.

Finally, polyhedron P_k is obtained by projecting L_k back onto R^n: $P_k = \{x \in R^n : \exists y \in L_k, \forall i = 1, \ldots, n, y_{\{i\}} = x_i\}$.

[2] The original paper [26] introduces one additional dimension for the purpose of homogeneization, but subsequently intersects the cone with the hyperplane $y_0 = 1$. That is equivalent to the definition used here.

We remark that in the above definition we may equivalently use all multipliers such that $|I \cup J| \leq k$; indeed, any constraint obtained from I, J with $|I \cup J| < k$ can be inferred by taking $i \notin I \cup J$ and adding the constraint for $(I \cup \{i\}, J)$ and the constraint for $(I, J \cup \{i\})$, so such constraints are redundant. (For some proofs it will be more convenient to include these redundant constraints.)

The following basic result is well known [26, 21].

Lemma 2.1 $P_0 \supseteq P_1 \supseteq \cdots P_k \supseteq \cdots \supseteq P_n = P$.

Thus the P_k are indeed progressively stronger relaxations of the integer polytope P, and after at most n rounds we arrive at P itself.

2.2 The matching polytope

Given a graph $G = (V, E)$ with $|V| = n$ and $|E| = m$, any subset of E can be written as a binary vector in $\{0,1\}^m$. Consider the following linear program.

$$\max_x f(x) = \sum_{e \in E} x_e \quad \text{s.t. } x \in L_0 : \begin{cases} \sum_{e: u \in e} x_e & \leq 1 \quad \forall u \in V \\ x_e & \leq 1 \quad \forall e \in E \\ x_e & \geq 0 \quad \forall e \in E \end{cases}$$

Clearly, the polytope L_0 of feasible solutions is contained in $[0,1]^m$, and $P = \text{Conv}(L_0 \cap \{0,1\}^m)$ describes exactly the set of convex combinations of matchings of G.

Starting from L_0, the kth iterate in the Sherali-Adams hierarchy defines the following lifted polyhedron L_k. For every vertex $u \in V$, for every possible I, J disjoint subsets of E with $|I \cup J| = k$, we multiply the constraint $1 - \sum_{v: u \neq v} x_{uv} \geq 0$ by $\prod_{e \in I} x_e \prod_{f \in J}(1 - x_f)$, replace each square x_e^2 by x_e, and replace each monomial $\prod_{\ell \in L} x_\ell$ by a variable y_L, to obtain a linear constraint in y. Add to this set all the inequalities obtained by linearization of $\prod_{i \in I} x_i \prod_{j \in J}(1 - x_j) \geq 0$ where I, J are disjoint subsets of E such that $|I \cup J| = \min(k+1, m)$. If P_k denotes the projection of L_k onto $R^{|E|}$, we have

$$\max\{f(x) \quad \text{s.t. } x \in P_k\} = \max\{\sum_{e \in E} y_{\{e\}} \quad \text{s.t. } y \in L_k\}$$

Abusing notation slightly, we write $f(y) = \sum_{e \in E} y_{\{e\}}$.

The *integrality ratio*[3] $g_k(G)$ of P_k applied to a given graph G is $\max_{x \in P_k} f(x) / \max_{x \in P} f(x)$, which equals $\max_{y \in L_k} f(y) / \max_{x \in P} f(x)$. The *integrality gap* of the kth iterate of the Sherali-Adams hierarchy is $\alpha_k = \sup_G g_k(G)$, which we study as a function of k. By Lemma 2.1 α_k is motonone non-increasing.

2.3 The integrality gap

Our first observation is that the integrality gap is always achieved on a complete graph of odd cardinality. For this we require the notion of a *certificate* (or witness) for maximum matching, given by the following version of the Tutte-Berge formula [7, 6].

Theorem 2.2 [7, 6] *The maximum cardinality of a matching of G equals the minimum of $|S_1| + \sum_{i \geq 2} \lfloor |S_i|/2 \rfloor$ over all partitions $V = S_1 \cup S_2 \cup \cdots \cup S_\ell$ of V into subsets such that every edge either has exactly one endpoint in S_1 or has both endpoints in the same S_i with $i \geq 2$. Such a partition is called a* certificate.

We can deduce the following:

[3] We introduce this term to distinguish the ratio on a particular graph G from the integrality gap, which is a supremum over all G.

Proposition 2.3 *The integrality gap α_k is achieved for G equal to a complete graph of odd cardinality, i.e., $\alpha_k = \sup\{g_k(K_{2d+1}), d \geq 1\}$.*

Proof: Let G be any graph. Let M be a maximum matching of G and $\{S_i\}$ be a certificate that M is maximum, as given by Theorem 2.2. Modify G by adding edges to make every S_i ($i \geq 2$) a complete graph and to make $(S_1, V - S_1)$ a complete bipartite graph. This transformation preserves the certificate, hence M is still an optimal matching, while it further relaxes the linear program. Therefore, such a modification cannot decrease the integrality ratio. Hence, in the definition of α_k, we can restrict our attention to graphs G such that V is partitioned into an independent set S_1, cliques $S_2, S_3, \ldots$, and a complete bipartite graph between S_1 and $V \setminus S_1$. The integer optimum on such a graph has value $|S_1| + \sum_{i\geq 2}\lfloor |S_i|/2\rfloor$. Denoting by $\mathrm{LP}_k(G)$ the value of the k-round Sherali-Adams LP on graph G, we have

$$\mathrm{LP}_k(G) \leq \sum_{e: e\cap S_1 \neq \emptyset} x_e + \sum_i \sum_{e\subseteq S_i} x_e \leq |S_1| + \sum_{i\geq 2} \mathrm{LP}_k(K_{|S_i|}),$$

and therefore

$$g_k(G) \leq \frac{|S_1| + \sum_{i\geq 2}\mathrm{LP}_k(K_{|S_i|})}{|S_1| + \sum_{i\geq 2}\lfloor\frac{|S_i|}{2}\rfloor} \leq \max_i \frac{\mathrm{LP}_k(K_{|S_i|})}{\lfloor\frac{|S_i|}{2}\rfloor} = \max_i g_k(K_{|S_i|})$$

where the last equality follows from the obvious fact that the integer optimum on K_n is $\lfloor n/2 \rfloor$. Hence, in order to compute $\max_G g_k(G)$, it suffices to restrict attention to complete graphs. Finally, note that $g_k(K_{2d+1}) \geq g_k(K_{2d})$ for all d, since adding the extra vertex does not change the integer optimum and can only increase the value of the LP. Hence, to compute α_k, it suffices to restrict attention to complete graphs of odd cardinality. $\square$

Remark: For graphs of any fixed size n, the integrality ratio is also determined by the values $\mathrm{LP}_k(K_j)$, for by the above proof we can write

$$g_k(G) = \max_{\sum_i s_i = n} \frac{\sum_i \mathrm{LP}_k(K_{s_i})}{\sum_i \lfloor\frac{s_i}{2}\rfloor}.$$

By Proposition 2.3, it is enough to study the integrality ratio $g_k(K_{2d+1})$ as a function of k and d. When $k = 0$ we are dealing with the basic LP relaxation, for which it is well known that the integrality ratio is $1 + 1/2d$:

Lemma 2.4 *For every d, we have $g_0(K_{2d+1}) = 1 + \frac{1}{2d}$.*

Proof: The integer optimum is d, and summing all the constraints of the linear program shows that any feasible solution has value at most

$$\sum_e x_e = \frac{1}{2}\sum_v \sum_{e: v\in e} x_e \leq (2d+1)/2.$$

Hence the integrality ratio is at most $1 + 1/2d$. On the other hand, setting $x_e = 1/2d$ for every e clearly gives a feasible solution, and its value is also $\binom{2d+1}{2}\frac{1}{2d} = \frac{2d+1}{2}$. $\square$

Of course, the above is the maximum possible value of the integrality ratio for any fixed d.

Presently we shall see (Corollary 3.10) that in fact $g_k(K_{2d+1})$ remains at its maximum value $1 + 1/2d$ for all $k \leq d - 1$, and also (Corollary 3.11 and Theorem 1.3) that $g_k(K_{2d+1})$ reaches 1 exactly at $k = 2d - 1$. We will also describe in some detail the decrease of the ratio between these two extreme values.

3 An explicit linear program

Our goal in this section is to find a simple, explicit form for the linear program obtained after k rounds of Sherali-Adams lifting applied to $G = K_{2d+1}$.[4] This explicit form can be found in Theorem 3.9 at the end of the section. The key to our analysis is the observation that, among all the multipliers that give rise to Sherali-Adams constraints, only a much smaller special set that we call "standard multipliers" are needed. These have a simple description as follows.

Definition 1 *A* standard multiplier *is a polynomial M in the variables $\{x_e : e \in E\}$, of the form $\prod_{e\in I} x_e \prod_{f\in J}(1 - x_f)$, where the edges of J are a star over some vertex set W and the edges of I are a matching over some vertex set W' disjoint from W.*

We will also need some notation for the following linearization procedure.

Definition 2 *Let C be a polynomial over $\{x_e\}$, and let $\phi(C)$ denote the linear combination of variables $z_1, z_2, \dots$ which is obtained by expanding C, linearizing each monomial (i.e., replacing x_e^m by x_e for each e and each $m > 1$), and replacing each $\prod_{e\in L} x_e$ by 0 if L is not a matching and by $z_{|L|}$ otherwise.*

The key step is to show that only standard multipliers are needed to define the k-round lifted linear program L_k:

Proposition 3.1 *Let $G = K_{2d+1}$. Then the value of L_k equals the value of the following linear program L'_k:*

$$\max_{z_1,z_2,\dots,z_{k+1}} \binom{2d+1}{2} z_1 \text{ subject to}$$

1. *$\forall i > d, \quad z_i = 0$;*
2. *all the constraints of the form $\phi((1 - \sum_{v:u\neq v} x_{uv})M) \geq 0$, where $u \in V$ and M is a standard multiplier of degree at most k over a vertex set not containing u;*
3. *all the constraints of the form $\phi(M) \geq 0$, where M is a standard multiplier of degree at most $k + 1$.*

The proof is through a sequence of lemmas. The first two of these determine the variables in the linear program.

Lemma 3.2 *If I is not a matching then $y_I = 0$.*

Proof: Let $e \in I$. Multiplying the constraint $x_e \geq 0$ by $\prod_{f\in I\setminus\{e\}} x_f$ yields $y_I \geq 0$.

Up to relabeling, assume that $I = \{01, 02\} \cup I'$. Multiplying the constraint $(1 - \sum_{i\neq 1} x_{1i} \geq 0)$ by $x_{01} \prod_{e\in I'} x_e$, replacing x_{01}^2 by x_{01}, and simplifying, yields $-\sum_{i\geq 2} x_{01}x_{0i} \prod_{e\in I'} x_e \geq 0$. For each $i \geq 3$, multiplying the constraint $x_{01} \geq 0$ by $x_{0i} \prod_{e\in I'} x_e$ yields $x_{01}x_{0i} \prod_{e\in I'} x_e \geq 0$. Summing, we obtain $-x_{01}x_{02} \prod_{e\in I'} x_e \geq 0$. Linearizing yields $-y_I \geq 0$.

Hence $y_I = 0$. □

Lemma 3.3 *There exists an optimal solution $y = (y_L) \in L_k$ and associated projection $x = (x_e) \in P_k$ realizing the fractional optimum $\max_{x\in P_k} f(x)$ such that $y_L = z_\ell$ is the same for every set of $\ell = |L|$ edges forming a matching.*

Proof: See the Appendix. □

[4] In this section, we use the definition of the Sherali-Adams construction with $|I \cup J| \leq k$ rather than $|I \cup J| = k$.

The next sequence of lemmas show that the only constraints we need to consider are those associated with standard multipliers.

Lemma 3.4 *Consider the constraint defined by I, J and u:*

$$C = \prod_{e \in I} x_e \prod_{f \in J} (1 - x_f)(1 - \sum_{v: u \neq v} x_{uv}).$$

Without loss of generality, we can assume that I is a matching, that vertex u does not belong to any edge of $I \cup J$, and that the vertices spanned by J are disjoint from the vertices spanned by I.

Proof: If I is not a matching, then no monomial in C is a matching and thus by Lemma 3.2 we have $\phi(C) = 0$. If vertex u belongs to an edge of I, then by Lemma 3.2 again we have $\phi(C) = \phi(C')$, where $C' = \prod_{e \in I} x_e \prod_{f \in J}(1 - x_f)$ is a constraint of the form covered in Lemma 3.5 below.

The remaining two cases are handled by induction on the number of factors in C (this is where we use the fact that we take the variant of the definition where $|I \cup J| < k$). First, if vertex u is an endpoint of an edge $\{u, w\}$ in J, then $C = (1 - x_{uw})(1 - \sum_{v: u \neq v} x_{uv})C'$, and since by Lemma 3.2 the linearization of $(1 - x_{uw})(1 - \sum_{v: u \neq v} x_{uv})$ equals $(1 - \sum_{v: u \neq v} x_{uv})$, we have $\phi(C) = \phi((1 - \sum_{v: u \neq v} x_{uv})C')$. Since the latter argument has one fewer factor than C, we can apply induction to it. Finally, if some vertex w appears in both an edge $\{w_1, w\}$ of I and an edge $\{w_2, w\}$ of J, then $C = x_{w_1 w}(1 - x_{w_2 w})C'$ and by Lemma 3.2 we have $\phi(C) = \phi(x_{w_1 w} C')$ and we can again apply induction. □

Lemma 3.5 *Consider the constraint defined by I and J:*

$$C = \prod_{e \in I} x_e \prod_{f \in J} (1 - x_f).$$

Without loss of generality, we can assume that I is a matching and that the vertices spanned by J are disjoint from the vertices spanned by I.

Proof: Similar to the proof of Lemma 3.4. □

The following straightforward fact will be needed in the proof of the next lemma.

Proposition 3.6 *Let C, D and F be polynomials in $\{x_e\}$ such that the set of vertices spanned by the edges in the support of C or D are disjoint from the set of vertices spanned by the edges in the support of F. If $\phi(C) = \phi(D)$ then $\phi(CF) = \phi(DF)$.*

Lemma 3.7 *Let J be a multiset of edges over some vertex set W (where the same edge can be present several times), and let $C = \prod_{e \in J}(1 - x_e)$. Then there exists a set $C_1, C_2, \ldots$ of standard multipliers over W, and positive coefficients $\lambda_1, \lambda_2, \ldots$, such that $\phi(C) = \sum_i \lambda_i \phi(C_i)$.*

Proof: The proof is by induction over the cardinality of J (degree of C) and over the number t of vertices of W that have more than one adjacent vertex in J. (Note that these adjacent vertices must be distinct. Multiple edges to the same neighbor count as a single adjacency.)

Base case: If $t = 1$, or if $t = 0$ and J spans only two vertices, then J is a star, possibly with some duplicate edges. If there are no duplicate edges, then the conclusion of the lemma holds and we are done. Otherwise, we write $C = \prod_{v \in S}(1 - x_{u_0 v})^{m_v}$, where $m_v \geq 1$ is the multiplicity of edge $\{u_0, v\}$. Observing that the linearization of $(1 - x_e)^2$ equals $(1 - x_e)$, it follows that $\phi(C) = \phi(\prod_{v \in S}(1 - x_{u_0 v}))$ and we are done.

General case: Otherwise, let v_1, v_2 be two vertices that both have neighbors outside $\{v_1, v_2\}$, let A be the multiset of edges from v_1 to neighbors in $V \setminus \{v_2\}$, and let B be the multiset of edges from v_2 to neighbors in $V \setminus \{v_1\}$.

Define B' as the multiset of edges obtained from B by replacing each occurrence of an edge $\{v_2, b\}$ by an occurrence of $\{v_1, b\}$, and define the multiset $J' = (J \setminus B) \cup B'$, where edges are counted with multiplicity. Let $C' = \prod_{e \in J'}(1 - x_e)$.

The matchings of J which do not have both an edge from A and an edge from B are in bijection with the matchings of J'. The other matchings of J have both an edge e_1 from A and an edge e_2 from B. Thus it is easy to check that

$$\phi(C) = \phi(C') + \sum_{e_1 \in A} \sum_{e_2 \in B} \phi(x_{e_1} x_{e_2} \prod_{e \in J \setminus (A \cup B)} (1 - x_e)).$$

Note that J' has the same number of edges as J, but in J' vertex v_2 has at most one adjacent vertex (namely, v_1), so we can apply induction to C'. Now consider $x_{e_1} x_{e_2} \prod_{e \in J \setminus (A \cup B)}(1 - x_e)$. By Lemma 3.5 we have $\phi(x_{e_1} x_{e_2} \prod_{e \in J \setminus (A \cup B)}(1 - x_e)) = \phi(x_{e_1} x_{e_2} \prod_{e \in J''}(1 - x_e))$, where J'' is the set of edges in $J \setminus (A \cup B)$ that have no vertex in common with $e_1 \cup e_2$. Applying induction to J'' (which has smaller degree than J) and using Proposition 3.6 to multiply by $F = x_{e_1} x_{e_2}$ concludes the proof. □

Lemma 3.8 *Let u be a vertex and C, D be polynomials in $\{x_e\}$'s such that the set of vertices spanned by the edges in the support of C or D does not contain u. If $\phi(C) = \phi(D)$ then $\phi(C(1 - \sum_v x_{uv})) = \phi(D(1 - \sum_{v: u \neq v} x_{uv}))$.*

Proof: See the Appendix. □

Armed with the foregoing lemmas, we now prove Proposition 3.1.

Proof of Proposition 3.1: From Lemmas 3.2 and 3.3, we can simplify L_k by defining a new set of variables, with variable z_i denoting the common value of y_I for every matching I of size i and by replacing y_I by 0 whenever I is not a matching. In other words, we take the intersection of the polytope with the subspace of equations $y_I = 0$ for I a non-matching and equations $y_I = y_J$ for I, J matchings of equal size. This transforms L_k into an equivalent linear program with variables $(z_i)_{i \geq 1}$. Since G has $\binom{2d+1}{2}$ edges, the objective function $\sum_{e \in E} y_{\{e\}}$ becomes $\binom{2d+1}{2} z_1$.

Since the maximum matching of G has size d, this implies that $y_I = 0$ whenever $|I| > d$, and therefore $z_i = 0$ for $i > d$. This establishes the first set of constraints.

The second set of constraints is trivially obtained by multiplying the appropriate constraint $(1 - \sum_{v: u \neq v} x_{uv} \geq 0)$ by the appropriate standard multiplier. We now proceed to prove that any other constraints which can be obtained from $1 - \sum_{v: u \neq v} x_{uv} \geq 0$ can be expressed as positive linear combinations of these constraints. By Lemma 3.4 we only need to examine constraints obtained by multiplying $1 - \sum_{v: u \neq v} x_{uv} \geq 0$ by $\prod_{e \in I} x_e \prod_{f \in J}(1 - x_f)$, where I is a matching not containing u, and J spans a set of vertices W which is disjoint from I and from u. Let $C = \prod_{f \in J}(1 - x_f)$. Applying Lemma 3.7 to C, we have

$$\phi(C) = \sum_{I', J'} \alpha_{I', J'} \phi(\prod_{e \in I'} x_e \prod_{f \in J'} (1 - x_f)),$$

where the coefficients $\alpha_{I', J'}$ are non-negative, I' is a matching in W, and J' is a star in W disjoint from I'. By Proposition 3.6 the equality still holds when each term is multiplied by $\prod_{e \in I} x_e$. Finally, by Lemma 3.8 the equality still holds when each term is multiplied by $1 - \sum_{v: u \neq v} x_{uv}$. Hence the constraint $\phi(C) \geq 0$ is a positive linear combination of the constraints described in Proposition 3.1.

Similarly, the third set of constraints is trivially obtained by multiplying the appropriate constraint ($x_e \geq 0$ or $1 - x_e \geq 0$) by the appropriate standard multiplier. Proving that any other constraints that can be obtained from $x_e \geq 0$ or $1 - x_e \geq 0$ are linear combinations of these constraints is analogous to the argument of the previous paragraph (using Lemma 3.5 in place of Lemma 3.4 and omitting the final step involving Lemma 3.8). $\square$

The following theorem writes in algebraic form the constraints of the linear program L'_k defined implicitly in Proposition 3.1. The proof is deferred to the appendix.

Theorem 3.9 *Let $G = K_{2d+1}$. For $k \leq d-1$, the value of the k-round Sherali-Adams lifted linear program is equal to the value of the following linear program:*

$$\max_{z_1, z_2, \ldots, z_{k+1}} \binom{2d+1}{2} z_1 \ \ s.t. \begin{cases} z_j - (2d-2j)z_{j+1} & \geq \ (k-j)(z_{j+1} - (2d-2j-2)z_{j+2}) \ \ (0 \leq j \leq k) \\ z_{k+1} & \geq \ 0, \end{cases}$$

with the special extreme case $z_0 = 1$.

For $k \geq d$, the value of the k-round Sherali-Adams lifted linear program is equal to the value of the following linear program:

$$\max_{z_1, z_2, \ldots, z_{k+1}} \binom{2d+1}{2} z_1 \ \ s.t. \begin{cases} z_{d+1} = \ldots & = \ z_{k+1} = 0 \\ z_j - (2d-2j)z_{j+1} & \geq \ \beta_j(z_{j+1} - (2d-2j-2)z_{j+2}) \ \ (0 \leq j \leq d), \end{cases}$$

with the special extreme case $z_0 = 0$ (j=0) and with the notation $\beta_j = \min(k-j, 2d-2j-1)$.

We close this section by deriving two immediate corollaries about the behavior of the integrality ratio $g_k(K_{2d+1})$ at the lower and upper extremes. (These are parts (i) and (iv) of Theorem 1.2 stated in the Introduction.)

Corollary 3.10 *For $k \leq d-1$ we have $g_k(K_{2d+1}) = 1 + 1/2d$.*

Proof: By Lemma 2.4 it suffices to prove this for $k = d-1$. Define $z_1 = 1/2d$ and $z_{j+1} = 1/((2d-2j)\cdots(2d-4)(2d-2)2d)$. Then $z_{k+1} \geq 0$ and all other constraints are tight, so this is feasible and has value $(2d+1)/2$. Since the integer optimum is d, the result follows. $\square$

Corollary 3.11 *For $k \geq 2d-1$ we have $g_k(K_{2d+1}) = 1$.*

Proof: We know [21, 26] that for k equal to the number of variables in the basic linear program, which in our case is $\binom{2d+1}{2}$, the integrality ratio is 1 and the value of the lifted linear program is equal to the integer optimum. But Proposition 3.1 implies that the value of the lifted linear program is the same for every $k \geq 2d-1$; this follows because the maximum possible degree of a standard multiplier in $G(K_{2d+1})$ is $2d$, so for $k \geq 2d$ no new constraints are added. Hence the ratio is 1 for every $k \geq 2d-1$. $\square$

4 Solving the linear program

Finding the optimal value of the linear program L'_k of Theorem 3.9 is now a purely algebraic problem. The result is given by the following recurrence relation. (Note that, by Corollaries 3.10 and 3.11, we need only consider the range $d \leq k \leq 2d-2$.) We defer the proof to the Appendix.

Lemma 4.1 *Let $d \le k \le 2d-2$. The optimal value of L'_k is*

$$\frac{d(2d+1)}{(k+2d) - 2k(d-1)/\rho_{2d-k-2}},$$

where $(\rho_i)_{0 \le i \le 2d-k-2}$ is given by the recurrence

$$\begin{cases} \rho_0 &= 2(k-d)+3; \\ \rho_i &= 4(k-d)+3+3i - \frac{(2(k-d)+i+1)(2(k-d)+2i)}{\rho_{i-1}} \qquad (\textit{for } i \ge 1). \end{cases}$$

One of our main theorems stated in the Introduction, Theorem 1.3 on the rank of the matching polytope, follows immediately from the above lemma.

Proof of Theorem 1.3: We just need to observe that $g_{2d-2}(K_{2d+1}) > 1$ (since we already know from Corollary 3.11 that $g_{2d-1}(K_{2d+1}) = 1$). This follows by plugging $k = 2d-2$ into Lemma 4.1, so that $\rho_{2d-k-2} = \rho_0$, whence the integrality ratio is easily seen to be $1 + 1/(4d^2-2) > 1$. □

Although we are not aware of any closed form for the solution of the recurrence in Lemma 4.1, by using it we can compute numerically the exact value of the linear program for any fixed number of rounds k and graph size $2d+1$. Additionally, by doing some asymptotic analysis we can prove tight bounds on that solution, and hence prove the threshold behavior for $g_k(K_{2d+1})$ claimed in Theorem 1.2 in the Introduction.

Proof of Theorem 1.2: Parts (i) and (iv) are exactly Corollaries 3.10 and 3.11 respectively.

Now, assume $d \le k < 2d-1$. Recalling the definition of the integrality ratio $g_k(K_{2d+1})$, and using the obvious fact that the integer optimum on K_n is $\lfloor \frac{n}{2} \rfloor$ together with Theorem 3.1 and Lemma 4.1, we see that

$$g_k(K_{2d+1}) = \frac{\text{value}(L_k)}{\lfloor (2d+1)/2 \rfloor} = \frac{\text{value}(L'_k)}{d} = \frac{(2d+1)}{(k+2d) - 2k(d-1)/\rho_{2d-k-2}},$$

with $(\rho_i)_i$ defined as in Lemma 4.1. It only remains to analyze the recurrence relation defining $(\rho_i)_i$. Let ϵ_i be given by $\rho_i = (2(k-d)+2i+2)(1+\epsilon_i)$. The recurrence becomes:

$$\begin{cases} \epsilon_0 &= \frac{1}{2(k-d)+2}; \\ \epsilon_i &= \left(1 - \frac{i+1}{2(k-d)+2i+2}\right) \frac{\epsilon_{i-1}}{1+\epsilon_{i-1}} \equiv r_i \frac{\epsilon_{i-1}}{1+\epsilon_{i-1}} \qquad (\text{for } i \ge 1). \end{cases} \tag{1}$$

Then it is easy to see that the ratio $g_k = g_k(K_{2d+1})$ satisfies

$$1 + \frac{1-(2d+1)\epsilon}{2d} \le g_k = \frac{2d+1}{2d + k\frac{\epsilon}{1+\epsilon}} \le 1 + \frac{1 - \frac{k\epsilon}{1+\epsilon}}{2d}, \tag{2}$$

where $\epsilon \equiv \epsilon_{2d-k-2}$. To prove parts (ii) and (iii) of the theorem, we use (2) to derive upper and lower bounds on g_k in the relevant ranges of k.

Lower bound on g_k for $d \le k \le 2d - \omega(d^{1/2})$

To prove the bound in part (ii), by (2) it suffices to show that $\epsilon = o(1/d)$. Let $k = (2-\gamma)d$ where $\gamma = \omega(d^{-1/2})$. Note that then $2d-k-2 = \gamma d - 2$. Consider the quantity r_i defined in (1). Since r_i decreases monotonically with i, for all i in the range $\frac{\gamma d}{2} \le i \le \gamma d - 2$, we have

$$r_i \le r_{\gamma d/2} = 1 - \frac{\gamma d/2 + 1}{2(1-\gamma)d + \gamma d + 2} \le 1 - \frac{\gamma}{4}.$$

Hence, from the recurrence in (1) we get

$$\epsilon_{2d-k-2} \le \epsilon_0 \prod_{i=\gamma d/2}^{\gamma d-2} r_i \le \epsilon_0 \left(1 - \frac{\gamma}{4}\right)^{\gamma d/2-2} \le \epsilon_0 \exp\left(-\frac{\gamma^2 d}{8} + \frac{\gamma}{2}\right). \tag{3}$$

Now if $k \ge \frac{3}{2}d$ then from (1) $\epsilon_0 \le \frac{1}{2d+2}$, and therefore (3) together with the fact that $\gamma = \omega(d^{-1/2})$ implies that $\epsilon = o(1/d)$. If on the other hand $k < \frac{3}{2}d$ then $\gamma \ge \frac{1}{2}$ and (3) again implies $\epsilon = o(1/d)$. The left-hand side of inequality (2) now completes the proof of part (ii) of the theorem.

Upper bound on g_k for $k \ge 2d - o(d^{1/2})$

To prove the bound in part (iii) of the theorem, by (2) it suffices to show that $\frac{k\epsilon}{1+\epsilon} = 1 - o(1)$. And since $k \ge 2d - o(d)$, it suffices to show $\epsilon = \frac{1-o(1)}{2d}$.

Let $k = 2d - \beta$, where $\beta = o(d^{1/2})$. Then $2d - k - 2 = \beta - 2$. In this case, for $1 \le i \le \beta - 2$ and sufficiently large d we have

$$r_i \ge r_\beta = 1 - \frac{\beta+1}{2d+2} \ge 1 - \frac{\beta}{d} \text{ and } \frac{\epsilon_i}{1+\epsilon_i} \ge \frac{\epsilon_i}{1+\epsilon_0} = \left(1 - \frac{1}{2(d-\beta)+3}\right)\epsilon_i.$$

Thus from the recurrence (1) we get

$$\epsilon \equiv \epsilon_{2d-k-2} \ge \epsilon_0 \prod_{i=1}^{\beta-2} \frac{r_i}{1+\epsilon_0} \ge \frac{1}{2(d-\beta+1)} \left(1 - \frac{1}{2(d-\beta)+3}\right)^{\beta-2} \left(1 - \frac{\beta}{d}\right)^{\beta-2}.$$

Since $\beta = o(d^{1/2})$ we see that the first factor is $\frac{1-o(1)}{2d}$, and the second and third factors are each $1-o(1)$. Hence $\epsilon = \frac{1-o(1)}{2d}$, which in conjunction with the right-hand side of inequality (2) completes the proof of part (iii) of the theorem. $\square$

Finally, our main result on the integrality gap, Theorem 1.1 stated in the Introduction, follows almost immediately from the above theorem.

Proof of Theorem 1.1: By Proposition 2.3, we know that $\alpha_k = \sup\{g_k(K_{2d+1}), d \ge 1\}$.

For a lower bound on α_k, choose $d = d(k)$ such that $k = 2d - \gamma$ where $\omega(\sqrt{d}) < \gamma < o(d)$. This implies that $d = \frac{k}{2} + o(k)$. By part (ii) of Theorem 1.2, we have $\alpha_k \ge g_k(K_{2d+1}) \ge 1 + \frac{1-o(1)}{2d} = 1 + \frac{1}{k} + o(\frac{1}{k})$.

For an upper bound on α_k, note from part (iv) of Theorem 1.2 that $g_k(K_{2d+1}) = 1$ for $d \le \frac{k}{2}$, and hence $\alpha_k \le \max\{g_k(K_{2d+1}) : d \le k/2\}$. But by part (i) of the theorem this is at most $\max\{1 + \frac{1}{2d} : d \le k/2\} = 1 + \frac{1}{k}$. $\square$

References

[1] N.E. Aguilera, S.M. Bianchi and G.L. Nasini. Lift and project relaxations for the matching and related polytopes. *Discrete Applied Mathematics* **134** (2004), pp. 193–212.

[2] M. Alekhnovich, S. Arora and I. Tourlakis. Towards strong non-approximability results in the Lovász-Schrijver hierarchy. *Proceedings of the 37th ACM Symposium on Theory of Computing*, 2005, pp. 294–303.

[3] S. Arora, B. Bollobás and L. Lovász. Proving integrality gaps without knowing the linear program. *Proceedings of the 43rd ACM Symposium on Theory of Computing*, 2002, pp. 313–322.

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

[5] E. Balas, S. Ceria and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming* 58 (1993), pp. 295–324.

[6] C. Berge. Sur le couplage maximum d'un graphe. *Comptes Rendus Hebdomadaires des Sances de l'Acadmie Sciences [Paris]* **247** (1958), pp. 258–259.

[7] J.A. Bondy and U.S.R. Murty. *Graph Theory: An Advanced Course*, Springer, 2007.

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

[9] M. Charikar. "On semidefinite programming relaxations for graph coloring and vertex cover." *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms*, 2002, pp. 616–620.

[10] M. Charikar, K. Makarychev and Y. Makarychev. "Integrality gaps for Sherali-Adams relaxations." Manuscript, 2007.

[11] E. Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, 2007, pp. 691–701.

[12] U. Feige and R. Krauthgamer. The probable value of the Lovász-Schrijver relaxations for maximum independent set. *SIAM Journal on Computing* **32** (2003), pp. 345–370.

[13] W. Fernandez de la Vega and C. Kenyon-Mathieu. Linear programming relaxations of Maxcut. *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms*, 2007.

[14] K. Georgiou, A. Magen, T. Pitassi and I. Tourlakis. Integrality gaps of $2 - o(1)$ for vertex cover SDPs in the Lovász-Schrijver hierarchy. *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, 2007.

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

[16] D. Grigoriev, E. Hirsch and D. Pasechnik. "Complexity of semialgebraic proofs." *Moscow Mathematical Journal* **2** (2002), pp. 647–679.

[17] S.-P. Hong and L. Tunçel. Unification of lower-bound analyses of the lift-and-project rank of combinatorial optimization polyhedra. *Discrete Applied Mathematics* **156** (2008), pp. 25–41.

[18] A. Kojevnikov and D. Itsykson. "Lower bounds for static Lovász-Schrijver calculus proofs for Tseitin tautologies." *Proceedings of ICALP* 2006, pp. 323–334.

[19] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization* **11** (2001), pp. 796–817.

[20] J.B. Lasserre. An explicit exact SDP relaxation for nonlinear 0-1 programs. *Proceedings of the 8th International Conference on Integer Programming and Combinatorial Optimization*, 2001, Springer Lecture Notes in Computer Science **2081**, pp. 293–303.

[21] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research* **28** (2003), pp. 470–496.

[22] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization* **1** (1991), pp. 166–190.

[23] T. Pitassi and N. Segerlind. Exponential lower bounds and integrality gaps for tree-like Lovász-Schrijver procedures. Preprint, 2007. Available as ECCC Report 07-107 at http://eccc.hpi-web.de/eccc-reports/2007/TR07-107/

[24] G. Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. *Proceedings of the 49th IEEE Symposium on Foundations of Computer Science*, 2008.

[25] G. Schoenebeck, L. Trevisan and M. Tulsiani. Tight integrality gaps for Lovász-Schrijver SDP relaxations of vertex cover and max cut. *Proceedings of the 39th ACM Symposium on Theory of Computing*, 2007, pp. 302–310.

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

[27] T. Stephen and L. Tunçel. On a representation of the matching polytope via semidefinite liftings. *Mathematics of Operations Research* **24** (1999), pp. 1–7.

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

Appendix

Proofs omitted from the main text

Proof of Lemma 3.3: Starting from an optimal solution y of L_k, define $n!$ optimal solutions by considering all possible permutations of the vertices: $y^{(\sigma)} = (y_L^{(\sigma)}) = (y_{\sigma(L)})$. By symmetry of K_{2d+1}, $y^{(\sigma)}$ is also an optimal solution of L_k. Averaging over all these solutions defines z. $\square$

Proof of Lemma 3.8: Let $C = \sum_\ell \sum_{M:|M|=\ell} \alpha_M \prod_{e\in M} x_e + C'$, and $D = \sum_\ell \sum_{M:|M|=\ell} \alpha'_M \prod_{e\in M} x_e + D'$, where M is a matching, and C', D' are polynomials whose monomials are all non-matchings. The fact that $\phi(C) = \phi(D) = \sum_\ell \beta_\ell z_\ell$ means that, for every ℓ, we have $\beta_\ell = \sum_{M:|M|=\ell} \alpha_M = \sum_{M:|M|=\ell} \alpha'_M$.

Now, since every matching of size ℓ spans exactly 2ℓ of the $2d$ neighbors of u, the coefficient of z_ℓ in both $\phi(C(1-\sum_v x_{uv}))$ and $\phi(D(1-\sum_v x_{uv}))$ is $\beta_\ell - (2d-2\ell)\beta_{\ell-1}$. The lemma follows. $\square$

Proof of Theorem 3.9: Consider the linear program L'_k defined in Proposition 3.1. From the first set of constraints, if $k \geq d$ then we have $z_j = 0$ for every $j > d$.

We now rewrite the second set of constraints of Proposition 3.1. Consider the basic constraint $C = (1 - \sum_{1\leq \ell \leq 2d} x_{0\ell})$. Take a standard multiplier such that $|I| = j \leq k$:

$$M_j = \prod_{m:0\leq m\leq j-1} x_{2d-2m,2d-2m-1} \cdot \prod_{i:2\leq i\leq |J|+1} (1 - x_{1i}).$$

$$\phi(CM_j) = \phi\left(\prod_{m:0\le m\le j-1} x_{2d-2m,2d-2m-1}\cdot(1-\sum_{1\le\ell\le 2d-2j} x_{0\ell})(1-\sum_{i:2\le i\le |J|+1} x_{1i})\right).$$

$$\phi(CM_j) = z_j - (2d-2j+|J|)z_{j+1} + \sum_{1\le\ell\le 2d-2j}\sum_{2\le i\le |J|+1} \chi(\ell\notin\{1,i\})z_{j+2}.$$

The number of non-zero terms in the double sum is $(2d-2j-1)|J| - |J|\cdot 1 = (2d-2j-2)|J|$, hence we obtain the constraint

$$z_j - (2d-2j)z_{j+1} \ge |J|(z_{j+1} - (2d-2j-2)z_{j+2}).$$

Depending on the sign of the coefficient of $|J|$, the critical constraint as $|J|$ varies is either for $|J|$ minimum, $|J| = 0$, or for $|J|$ maximum. What is the maximum value of $|J|$? Since $|I|+|J|\le k$, we must have $|J|\le k-j$. Since the total number of vertices spanned by the edges of $I\cup J$ is at most $2d$ (all vertices except vertex 0) and I spans exactly $2j$, J must span at most $2d-2j$. Since the set of edges defined by J is a tree, it has at most $2d-2j-1$ edges: we obtain that the maximum value is $|J| = \min(k-j, 2d-2j-1)$. Hence the second set of constraints can be written as:

$$\begin{cases} z_j - (2d-2j)z_{j+1} & \ge 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad (0\le j\le \min(k,d)) \\ z_j - (2d-2j)z_{j+1} & \ge \min(k-j, 2d-2j-1)(z_{j+1} - (2d-2j-2)z_{j+2}) \quad (0\le j\le\min(k,d)) \end{cases}$$

For $j = \min(k,d)$ the two equations coincide, and so, for any j, the first inequality is subsumed by the second inequality. Also note that if $k\le d-1$ then $d\le 2d-k-1$, so every j has $j\le 2d-k-1$, and therefore $k-j\le 2d-2j-1$ so that $\min(k-j,2d-2j-1) = k-j$. Thus the above system is equivalent to:

$$\text{If } k\ge d \text{ then } z_j - (2d-2j)z_{j+1} \ge \min(k-j, 2d-2j-1)(z_{j+1} - (2d-2j-2)z_{j+2}) \quad (0\le j\le d) \tag{4}$$

$$\text{If } k\le d-1 \text{ then } z_j - (2d-2j)z_{j+1} \ge (k-j)(z_{j+1} - (2d-2j-2)z_{j+2}) \quad (0\le j\le k) \tag{5}$$

For the third set of constraints, take a standard multiplier such that $|I| = j\le k+1$.

$$M_j = \prod_{m:0\le m\le j-1} x_{2d-2m,2d-2m-1}\cdot \prod_{i:2\le i\le |J|+1}(1-x_{1i}).$$

$$\phi(M_j) = z_j - |J|z_{j+1}.$$

The critical constraint is for $|J| = 0$ or for $|J|$ maximum. What is the maximum value of $|J|$? Since $|I|+|J|\le k+1$, we must have $|J|\le k+1-j$. Since the total number of vertices spanned by the edges of $I\cup J$ is at most $2d+1$ (all vertices) and I spans exactly $2j$, J must span at most $2d-2j+1$. Since the set of edges defined by J is a tree, it has at most $2d-2j$ edges: we obtain that the maximum value is $|J| = \min(k-j+1, 2d-2j)$. Hence the third set of constraints can be written as:

$$\begin{cases} z_j & \ge 0 \quad (0\le j\le\min(k+1,d)) \\ z_j - \min(k+1-j, 2d-2j)z_{j+1} & \ge 0 \quad (0\le j\le \min(k+1,d)) \end{cases}$$

Again, for $j = \min(k+1,d)$, the two inequalities coincide, and so for any j, the first inequality is implied by the second inequality. Again, if $k\le d-1$ then $k-j\le 2d-2j-1$ for every j and so $k-j+1\le 2d-2j$, so that $\min(k+1-j, 2d-2j) = k+1-j$. Thus this is equivalent to:

$$\text{If } k\ge d \text{ then } z_j - \min(k+1-j, 2d-2j)z_{j+1} \ge 0 \quad (0\le j\le d) \tag{6}$$

$$\text{If } k\le d-1 \text{ then } z_j - (k+1-j)z_{j+1}\ge 0 \quad (0\le j\le k+1) \tag{7}$$

Consider the case $k\ge d$. Then Equation (4) implies that $z_j - (2d-2j)z_{j+1}\ge 0$ for $0\le j\le d$, which implies (6), and so we obtain the claimed linear program.

Now, consider the case $k\le d-1$. Then it is easy to see that Equation (5) implies (7) for $0\le j\le k$ since $2d-2j\ge k+1-j$. For $j = k+1$, (7) is simply $z_{k+1}\ge 0$. Thus we obtain the claimed linear program. □

Proof of Lemma 4.1: Let us write the constraints of L'_k more explicitly. There are two cases, depending on how $k-j$ compares to $2d-2j-1$. If $j \leq 2d-k-1$ then $k-j \leq 2d-2j-1$ and so $\beta_j = k-j$. If $j > 2d-k-1$ then $k-j > 2d-2j-1$ and so $\beta_j = 2d-2j-1$. We can rewrite L'_k as:

$$\max \binom{2d+1}{2} z_1 \;\; \text{s.t.}$$

$$\begin{cases} \left((k+2d) - \frac{2k(d-1)}{z_1/z_2}\right) z_1 & \leq & 1 & (j=0) \\ \frac{z_j}{z_{j+1}} & \geq & (k-j) + 2(d-j) - \frac{(k-j)(2(d-j)-2)}{z_{j+1}/z_{j+2}} & (1 \leq j \leq 2d-k-1) \\ \frac{z_j}{z_{j+1}} & \geq & 4(d-j) - 1 - \frac{(2(d-j)-1)(2(d-j)-2)}{z_{j+1}/z_{j+2}} & (2d-k \leq j \leq d-2) \\ \frac{z_{d-1}}{z_d} & \geq & 3 & (j=d-1) \\ z_d & \geq & 0 & (j=d) \end{cases}$$

Clearly, the optimum is obtained when the ratio z_1/z_2 is minimized. This optimum is obtained when every inequality in the system (other than the last line) is an equality. Solving this system of equalities with unknowns z_j/z_{j+1} for $j \in [2d-k, d-1]$ gives $z_j/z_{j+1} = 2(d-j)+1$: true for $d-1$, and by induction

$$4(d-j) - 1 - \frac{(2(d-j)-1)(2(d-j)-2)}{2(d-j-1)+1} = 4(d-j) - 1 - (2(d-j)-2) = 2(d-j)+1.$$

For $j = 2d-k-1$ we get $\rho_0 = z_j/z_{j+1} = 2(k-d)+3$. Letting $\rho_i = z_j/z_{j+1}$ with $i = 2d-k-1-j$ and substituting for $i = 0, 1, 2, \ldots$ yields the formula of the Lemma. $\square$