Centrum voor Wiskunde en Informatica

REPORT RAPPORT

A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0 - 1 programming

M. Laurent

Probability, Networks and Algorithms (PNA)

PNA-R0108 July 31, 2001

Report PNA-R0108
ISSN 1386-3711

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

A Comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre Relaxations for $0-1$ Programming

Monique Laurent

CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
monique@cwi.nl

ABSTRACT
Sherali and Adams [SA90], Lovász and Schrijver [LS91] and, recently, Lasserre [Las01b] have proposed lift and project methods for constructing hierarchies of successive linear or semidefinite relaxations of a $0-1$ polytope $P \subseteq \mathbb{R}^n$ converging to P in n steps. Lasserre's approach uses results about representations of positive polynomials as sums of squares and the dual theory of moments. We present the three methods in a common elementary framework and show that the Lasserre construction provides the tightest relaxations of P. As an application this gives a direct simple proof for the convergence of the Lasserre's hierarchy. We describe applications to the stable set polytope and to the cut polytope.

2000 Mathematics Subject Classification: 05C50, 15A57, 52B12, 90C05, 90C22, 90C27.
Keywords and Phrases: linear relaxation, semidefinite relaxation, lift and project, cut polytope, stable set polytope.
Note: The research of this paper was carried out under the PNA1.1-project of CWI.

1 Introduction

Given a set $F \subseteq \{0,1\}^n$, we are interested in finding the linear inequality description for the polytope $P := \text{conv}(F)$. A first (often easy) step is to find a *linear programming formulation* for P; that is, to find a linear system $Ax \le b$ for which the polytope

$$K := \{x \in \mathbb{R}^n \mid Ax \le b\}$$

satisfies $K \cap \{0,1\}^n = F$.

If all vertices of K are integral then $K = \text{conv}(F)$ and we are done. Otherwise we have to find 'cutting planes' permitting to strenghten the relaxation K and to cut off its fractional vertices. Such cutting planes can be found by exploiting the combinatorial structure of the problem at hand; extensive research has been done for finding (partial) linear descriptions for many polyhedra arising from specific combinatorial optimization problems. Next to that, research has also focused on developing general purpose methods applying to arbitrary 0-1 problems or, more generally, integer programming problems.

One of the first such methods, which applies more generally to integral polyhedra, is the method of Gomory for generating cuts tightening the linear relaxation K. Given a linear inequality $\sum_i a_i x_i \le \alpha$ valid for K where all the coefficients a_i are integers, the inequality $\sum_i a_i x_i \le \lfloor \alpha \rfloor$ (known as a

Gomory-Chvátal cut) is still valid for conv(F) but may eliminate some part of K. If we apply this transformation to any inequality $\sum_i a_i x_i \leq \alpha$ which can be obtained by taking linear combinations of the inequalities defining K with suitable nonnegative multipliers ensuring that the a_i's are integral, then we obtain a polytope K' satisfying

$$\mathrm{conv}(F) \subseteq K' \subseteq K.$$

Set $K^{(1)} := K'$ and define recursively $K^{(t+1)} := (K^{(t)})'$. Chvátal [C73] proved that $K^{(t)} = \mathrm{conv}(F)$ for some t; the smallest t for which this is true is the *Chvátal rank* of the polytope K. The Chvátal rank may be very large as it depends in general not only on the dimension n but also on the coefficients of the inequalities involved. However, when K is assumed to be contained in the cube $[0,1]^n$ then its Chvátal rank is bounded by $O(n^2 \log n)$ [ES99]. Even if we can optimize a linear objective function over K in polynomial time optimizing a linear objective function over the first Chvátal closure K' is a co-NP-hard problem in general [E99].

Another popular method is to try to represent P as the projection of another polytope Q lying in a higher (but preferably still polynomial) dimensional space. The idea behind being that the projection of a polytope Q may have more facets than Q itself. Hence it could be that even if P has an exponential number of facets, such Q exists having only a polynomial number of facets and lying in a space whose dimension is polynomial in the original dimension of P (such Q is sometimes called a *compact representation* of P). If this is the case then we have a proof that any linear optimization problem over P can be solved in polynomial time.

Several methods have been developed for constructing projection representations for general 0-1 polyhedra; in particular, by Balas, Ceria and Cornuéjols [BCC93], by Sherali and Adams [SA90], by Lovász and Schrijver [LS91] and, recently, by Lasserre [Las00, Las01b]. A common feature of these methods is the construction of a hierarchy $K \supseteq K^1 \supseteq K^2 \supseteq \ldots \supseteq P$ of relaxations of P which finds the exact convex hull in n steps; that is, $K^n = P$. These relaxations are linear or semidefinite (in the case of Lovász-Schrijver and Lasserre). Moreover, under some assumptions over K, one can optimize in polynomial time a linear objective over an iterate K^t for any fixed t.

The following inclusions are known among these various hierarchies: the Sherali-Adams iterate is contained in the Lovász-Schrijver iterate which in turn is contained in the Balas-Ceria-Cornuéjols iterate. The latter inclusion is an easy verification and the former was mentioned in [LS91] as an application of somewhat complicated algebraic manipulations; we present in Section 4 a simple direct proof for this inclusion.

The construction of Lasserre is motivated by results about representations of nonnegative polynomials as sums of squares and the dual theory of moments and his proof that the $0-1$ polytope P is found after n steps relies on a nontrivial result of Curto and Fialkow [CF00] about truncated moment sequences. In fact, the Sherali-Adams series of relaxations can also be formulated within this framework of moment matrices. The fact of formulating both Lasserre and Sherali-Adams constructions in a common setting permits a better understanding of how they relate; both constructions apply in fact to the case when K is a semi-algebraic set contained in the cube $[0,1]^n$. Moreover, the *same* argument can be used for showing that the $0-1$ polytope P is found after n steps in both constructions. This argument concerns an elementary property of the zeta matrix of the lattice $\mathcal{P}(V)$ of subsets of the set $V = [1,n]$, presented in Section 3.1. We show in Section 4 that the Lasserre hierarchy is a refinement of both the Sherali-Adams and the Lovász-Schrijver hierarchies. We give in Section 5 two examples showing that n steps are sometimes needed for finding P using the Sherali-Adams construction and we illustrate in Section 6 how the various methods apply to the stable set polytope and to the cut polytope of a graph. Section 7 contains some background information about the moment problem and the representation of positive polynomials as sums of squares, useful for understanding Lasserre's approach. In particular, we show that our presentation of Lasserre's method in Section 3 (in terms of moment matrices indexed by the semigroup $\mathcal{P}(V)$) is equivalent to the original presentation of Lasserre (in terms of moment matrices indexed by the semigroup $\mathbb{Z}_+^n$).

2 The Lovász-Schrijver hierarchy

Let K be a convex body contained in the cube $[0,1]^n$ and let

$$P := \text{conv}(K \cap \{0,1\}^n)$$

be the $0-1$ polytope to be described. For convenience, define

$$\tilde{K} := \{\lambda(1,x) \mid x \in K\}, \tag{1}$$

the *homogenization* of K; $\tilde{K}$ is a cone in $\mathbb{R}^{n+1}$ (the additional coordinate is indexed by 0) and $K = \{x \in \mathbb{R}^n \mid (1,x) \in \tilde{K}\}$. Let $M(K)$ denote the set of symmetric matrices $Y = (y_{ij})_{i,j=0}^n$ satisfying

$$y_{j,j} = y_{0,j} \text{ for } j = 1, \ldots, n, \tag{2}$$

$$Ye_j,\ Y(e_0 - e_j) \in \tilde{K} \text{ for } j = 1, \ldots, n \tag{3}$$

and set

$$N(K) := \{x \in \mathbb{R}^n \mid (1,x) = Ye_0 \text{ for some } Y \in M(K)\},$$

where $e_0, e_1, \ldots, e_n$ denote the standard unit vectors in $\mathbb{R}^{n+1}$. Then,

$$P \subseteq N(K) \subseteq K.$$

The inclusion $P \subseteq N(K)$ follows from the fact that, for $x \in K \cap \{0,1\}^n$, the matrix $Y := (1,x)(1,x)^T$ belongs to $M(K)$ and the inclusion $N(K) \subseteq K$ from property (3). Define iteratively $N^1(K) := N(K)$ and, for $t \geq 2$, $N^t(K) := N(N^{t-1}(K))$. Then,

$$P \subseteq N^n(K) \subseteq \ldots \subseteq N^{t+1}(K) \subseteq N^t(K) \subseteq \ldots \subseteq N(K) \subseteq K.$$

Lovász and Schrijver [LS91] show that $N^n(K) = P$. (Their proof assumes that K is a polytope but remains valid for any convex body K.)

Stronger relaxations are obtained by adding positive semidefiniteness. Set

$$M_+(K) := \{Y \in M(K) \mid Y \succeq 0\} \text{ and } N_+(K) = \{x \in \mathbb{R}^n \mid (1,x) = Ye_0 \text{ for some } Y \in M_+(K)\}.$$

Then,

$$P \subseteq N_+(K) \subseteq K;$$

the inclusion $P \subseteq N_+(K)$ following from the fact that, for $x \in K \cap \{0,1\}^n$, the matrix $Y := (1,x)(1,x)^T$ is positive semidefinite and, thus, belongs to $M_+(K)$. Define iteratively $N_+^1(K) := N_+(K)$ and $N_+^t(K) := N_+(N_+^{t-1}(K))$ for $t \geq 2$. Then,

$$P \subseteq N_+^t(K) \subseteq N^t(K) \text{ for } t \geq 1.$$

3 The Sherali-Adams and Lasserre hierarchies

The Sherali-Adams and Lasserre constructions apply to semi-algebraic sets contained in the cube $[0,1]^n$. Let

$$K := \{x \in [0,1]^n \mid g_\ell(x) \geq 0 \ \text{ for } \ell = 1, \ldots, m\} \tag{4}$$

where $g_1, \ldots, g_m$ are polynomials in $x_1, \ldots, x_n$ and let $P := \text{conv}(K \cap \{0,1\}^n)$ be the $0-1$ polytope to be described. As $x_i^2 = x_i$ $(i = 1, \ldots, n)$ for any $x \in \{0,1\}^n$, we can assume that each variable occurs in every g_ℓ with a degree ≤ 1 and, thus, $g_\ell(x)$ can be written as

$$\sum_{I \subseteq V} g_\ell(I) \prod_{i \in I} x_i.$$

We use the same symbol g_ℓ for denoting the vector in $\mathbb{R}^{\mathcal{P}(V)}$ with components $g_\ell(I)$ $(I \subseteq V)$. We first describe the two constructions in the common setting of moment matrices. We need some definitions.

Given $V := [1, n]$, $\mathcal{P}(V)$ denotes the collection of all subsets of V and, for $1 \leq t \leq n$, $\mathcal{P}_t(V)$ denotes the collection of subsets of cardinality $\leq t$. The components of a vector $y \in \mathbb{R}^{\mathcal{P}(V)}$ are denoted as y_I or $y(I)$; we also set $y_0 = y_\emptyset$, $y_i = y_{\{i\}}$ and $y_{ij} = y_{\{i,j\}}$. Given $y \in \mathbb{R}^{\mathcal{P}(V)}$, an integer $1 \leq t \leq n$, and a subset $U \subseteq V$, define the matrices

$$M_t(y) := (y(I \cup J))_{|I|,|J| \leq t}, \ \ M_U(y) := (y(I \cup J))_{I,J \subseteq U}. \tag{5}$$

Thus, $M_V(y) = M_n(y)$; this matrix is known as the *moment matrix* of y (cf. Section 7.2 for background information). For $x, y \in \mathbb{R}^{\mathcal{P}(V)}$, $x * y$ denotes the vector of $\mathbb{R}^{\mathcal{P}(V)}$ with entries

$$x * y(I) := \sum_{K \subseteq V} x_K y_{I \cup K}. \tag{6}$$

One can easily verify the following commutation rule which will be used later in Section 4. For $x, y, z \in \mathbb{R}^{\mathcal{P}(V)}$,

$$x * (y * z) = y * (x * z). \tag{7}$$

The Sherali-Adams and Lasserre relaxations are both based on the following observation.

Lemma 1. *Given $x \in K \cap \{0,1\}^n$, the vector $y \in \mathbb{R}^{\mathcal{P}(V)}$ with entries $y(I) := \prod_{i \in I} x_i$ $(I \subseteq V)$ satisfies*

$$M_V(y) \succeq 0, \ \ M_V(g_\ell * y) \succeq 0 \ \ \textit{for } \ell = 1, \ldots, m. \tag{8}$$

Proof. Indeed, $M_V(y) = yy^T$ and $M_V(g_\ell * y) = g_\ell(x) \ yy^T$, since $y(I \cup J) = y(I) \cdot y(J)$ for all $I, J \subseteq V$. ∎

One can relax the condition (8) and require positive semidefiniteness of certain principal submatrices of the moment matrices $M_V(y)$ and $M_V(g_\ell * y)$. Namely, Lasserre requires that

$$M_{t+1}(y) \succeq 0, \ \ M_{t-v_\ell+1}(g_\ell * y) \succeq 0 \text{ for } \ell = 1, \ldots, m \tag{9}$$

(for an integer $t \geq v_\ell - 1$, where $v_\ell := \lceil \frac{w_\ell}{2} \rceil$, w_ℓ being the degree of g_ℓ) while Sherali and Adams require that

$$M_W(y) \succeq 0,\ M_U(g_\ell * y) \succeq 0 \text{ for } \ell = 1, \ldots, m \text{ and } U, W \subseteq V \text{ with } |U| = t, |W| = \min(t+1, n) \tag{10}$$

(for an integer $t = 1, \ldots, n$). The corresponding relaxations of P are obtained by projecting the variable y onto the subspace $\mathbb{R}^n$ indexed by the singletons in $\mathcal{P}(V)$. Sherali and Adams and Lasserre show that P is found after n steps in the two constructions. These two results are a direct consequence of Corollary 3 below asserting that the cone in $\mathbb{R}^{\mathcal{P}(V)}$ consisting of the vectors y satisfying (8) is generated by $0-1$ vectors.

The Sherali-Adams relaxations turn out to be linear relaxations since the condition (10) can be reformulated as a linear system in y (cf. Lemma 2 below). We present in Section 3.2 the original definition of the Sherali-Adams relaxations given in [SA90] and its equivalence with the above definition.

3.1 Preliminary results

Let Z denote the square $0-1$ matrix indexed by $\mathcal{P}(V)$ with entry $Z_{I,J} = 1$ if and only if $I \subseteq J$. Its inverse Z^{-1} has entries

$$Z^{-1}_{I,J} = (-1)^{|J \setminus I|} \text{ if } I \subseteq J,\ 0 \text{ otherwise.} \tag{11}$$

The matrix Z is known as the *zeta matrix* of the lattice $\mathcal{P}(V)$ and its inverse Z^{-1} as the *Möbius matrix* of $\mathcal{P}(V)$ (cf. [Wi68]). Let ζ^J denote the J-th column of Z; it has entries $\zeta^J(I) = \prod_{i \in I} x_i$ ($I \subseteq V$), setting $x := \chi^J$. Given a subset $\mathcal{J} \subseteq \mathcal{P}(V)$, let $\mathcal{C}_\mathcal{J}$ denote the cone generated by the columns ζ^J of Z for $J \in \mathcal{J}$. Hence, $\mathcal{C}_\mathcal{J}$ is a simplicial cone in $\mathbb{R}^{\mathcal{P}(V)}$ and

$$\mathcal{C}_\mathcal{J} = \{y \in \mathbb{R}^{\mathcal{P}(V)} \mid Z^{-1}y \geq 0,\ (Z^{-1}y)_J = 0\ \forall J \notin \mathcal{J}\}. \tag{12}$$

The next lemma is based on ideas from section 3.a in [LS91].

Lemma 2. *Let $g, y \in \mathbb{R}^{\mathcal{P}(V)}$. Then,*

(i) $M_V(g * y) \succeq 0 \iff (Z^{-1}y)_H \cdot g^T \zeta^H \geq 0$ *for all* $H \subseteq V$.

(ii) $M_V(y) \succeq 0 \iff Z^{-1}y \geq 0 \iff \sum_{H \supseteq I} (-1)^{|H \setminus I|} y(H) \geq 0$ *for all* $I \subseteq V$.

Proof. (i) Let $u \in \mathbb{R}^{\mathcal{P}(V)}$ with entries $u_H := (Z^{-1}y)_H \cdot g^T \zeta^H$ for $H \subseteq V$ and let D_u denote the diagonal matrix indexed by $\mathcal{P}(V)$ with diagonal entries u_H ($H \subseteq V$). We show that $ZD_uZ^T = M_V(g * y)$. For this note that, for $H \subseteq V$,

$$u_H = (Z^{-1}y)_H \cdot g^T \zeta^H = \left(\sum_{R \supseteq H} (-1)^{|R \setminus H|} y_R \right) \cdot \left(\sum_{K \subseteq H} g_K \right) = \sum_{K \subseteq H \subseteq R} (-1)^{|R \setminus H|} y_R g_K.$$

Therefore, given $I, J \subseteq V$, the (I, J)-th entry of ZD_uZ^T is equal to

$$\sum_H Z_{IH} Z_{JH} u_H = \sum_{H \supseteq I \cup J} u_H = \sum_{K,R} y_R g_K \left(\sum_{I \cup J \cup K \subseteq H \subseteq R} (-1)^{|R \setminus H|} \right) = \sum_K g_K y_{I \cup J \cup K} = g * y(I \cup J)$$

and thus to $M_V(g*y)_{IJ}$, using the fact that $\sum_{I\cup J\cup K\subseteq H\subseteq R}(-1)^{|R\setminus H|}=1$ if $R=I\cup J\cup K$ and 0 otherwise. Assertion (i) now follows from the fact that $u\geq 0$ is equivalent to $W\succeq 0$.

The first equivalence in (ii) follows directly from (i) applied to g with all zero components except $g_\emptyset=1$ and the second equivalence follows from the description of Z^{-1} in (11). ∎

Let $g_\ell(x)$ ($\ell=1,\ldots,m$) be polynomials in which every variable occurs with degree ≤ 1 and set

$$\mathcal{J}:=\{J\subseteq V \mid g_\ell^T\zeta^J\geq 0 \ \text{ for all } \ell=1,\ldots,m\}=\{J\subseteq V\mid g_\ell(\chi^J)\geq 0 \ \text{ for all } \ell=1,\ldots,m\}. \tag{13}$$

In the case $\mathcal{J}=\mathcal{P}(V)$, the next result is given in [LS91] and [SA90].

Corollary 3. *$\mathcal{C}_\mathcal{J}=\{y\in\mathbb{R}^{\mathcal{P}(V)}\mid M_V(y)\succeq 0$ and $M_V(g_\ell*y)\succeq 0$ for all $\ell=1,\ldots,m\}$.*

Proof. Let $y\in\mathbb{R}^{\mathcal{P}(V)}$. By definition, $y\in\mathcal{C}_\mathcal{J}$ if and only if $Z^{-1}y\geq 0$ and $(Z^{-1}y)_J=0$ for $J\notin\mathcal{J}$. This is equivalent to $Z^{-1}y\geq 0$ and $(Z^{-1}y)_J\cdot g_\ell^T\zeta^J\geq 0$ for all $\ell=1,\ldots,m$ and $J\subseteq V$. Therefore, by Lemma 2, $y\in\mathcal{C}_\mathcal{J}$ if and only if $M_V(y)\succeq 0$ and $M_V(g_\ell*y)\succeq 0$ for $\ell=1,\ldots,m$. ∎

We see in Lemma 5 below how positive semidefiniteness of the moment matrices of $g*y$, when $g(x)$ belongs to the polynomials x_i, $1-x_i$ ($i=1,\ldots,n$), can be reformulated in terms of positive semidefiniteness of the moment matrix of y. This result tells us how to handle the bound inequalities $0\leq x_i\leq 1$ and will be used in Section 4. The proof uses the following fact.

Lemma 4. *Let X be a symmetric matrix with block decomposition $X=\begin{pmatrix}A & B\\ B & B\end{pmatrix}$ where A,B have the same order p. Then, $X\succeq 0 \iff B\succeq 0$ and $A-B\succeq 0$.*

Proof. Use the fact that, for $x,y\in\mathbb{R}^p$, $(x,y)^TX(x,y)=x^T(A-B)x+(x+y)^TB(x+y)$. ∎

Lemma 5. *Let $y\in\mathbb{R}^{\mathcal{P}(V)}$. Then,*

(i) *$M_U(e_i*y)$, $M_U((e_\emptyset-e_i)*y)\succeq 0$ for all $U\subseteq V$ with $|U|=t$ and $i=1,\ldots,n \iff M_W(y)\succeq 0$ for all $W\subseteq V$ with $|W|=\min(n,t+1)$.*

(ii) *$M_t(y)\succeq 0 \implies M_{t-1}(e_i*y)$, $M_{t-1}((e_\emptyset-e_i)*y)\succeq 0$ for all $i=1,\ldots,n$.*

Proof. (i) Let $W:=U\cup\{i\}$ where $|U|=t$ and $i\notin U$. Then, the matrix $M_W(y)$ has the block decomposition

$$M_W(y)=\begin{array}{c c} & \begin{array}{cc}\mathcal{P}(U) & \mathcal{P}(W)\setminus\mathcal{P}(U)\end{array}\\ \begin{array}{c}\mathcal{P}(U)\\ \mathcal{P}(W)\setminus\mathcal{P}(U)\end{array} & \begin{pmatrix} A & B\\ B & B\end{pmatrix}\end{array},$$

$M_U(e_i*y)=B$, $M_U(y)=A$ and $M((e_\emptyset-e_i)*y)=A-B$. From this follows the 'only if part' of (i) and the 'if part' in the case when $i\notin U$. For the 'if part' in the case when $i\in U$, note that

$$M_U(y)=\begin{pmatrix}A & B\\ B & B\end{pmatrix},\ M_U(e_i*y)=\begin{pmatrix}B & B\\ B & B\end{pmatrix},\ M_U((e_\emptyset-e_i)*y)=\begin{pmatrix}A-B & 0\\ 0 & 0\end{pmatrix},$$

with respect to the partition of $\mathcal{P}(U)$ into $\mathcal{P}(U \setminus \{i\})$ and its complement and use Lemma 4.
(ii) Set $\mathcal{P}_1 := \{I \in \mathcal{P}_{t-1}(V) \mid i \notin I\}$, $\mathcal{P}'_1 := \{I \in \mathcal{P}_{t-1}(V) \mid i \in I\}$, and $\mathcal{P}_2 := \{I \cup \{i\} \mid I \in \mathcal{P}_1\}$. Then, the principal submatrix of $M_t(y)$ indexed by $\mathcal{P}_1 \cup \mathcal{P}'_1 \cup \mathcal{P}_2$ has the block configuration:

$$\begin{array}{c} \\ \mathcal{P}_1 \\ \mathcal{P}'_1 \\ \mathcal{P}_2 \end{array} \begin{array}{c} \begin{array}{ccc} \mathcal{P}_1 & \mathcal{P}'_1 & \mathcal{P}_2 \end{array} \\ \begin{pmatrix} A & D & B \\ D & C & D \\ B & D & B \end{pmatrix} \end{array};$$

$$M_{t-1}(e_i * y) = \begin{array}{c} \\ \mathcal{P}_1 \\ \mathcal{P}'_1 \end{array} \begin{array}{c} \begin{array}{cc} \mathcal{P}_1 & \mathcal{P}'_1 \end{array} \\ \begin{pmatrix} B & D \\ D & C \end{pmatrix} \end{array}, \; M_{t-1}(y) = \begin{array}{c} \\ \mathcal{P}_1 \\ \mathcal{P}'_1 \end{array} \begin{array}{c} \begin{array}{cc} \mathcal{P}_1 & \mathcal{P}'_1 \end{array} \\ \begin{pmatrix} A & D \\ D & C \end{pmatrix} \end{array}, \; M_{t-1}((e_\emptyset - e_i) * y) = \begin{array}{c} \\ \mathcal{P}_1 \\ \mathcal{P}'_1 \end{array} \begin{array}{c} \begin{array}{cc} \mathcal{P}_1 & \mathcal{P}'_1 \end{array} \\ \begin{pmatrix} A - B & 0 \\ 0 & 0 \end{pmatrix} \end{array}.$$

Therefore, using Lemma 4, we find that $M_t(y) \succeq 0$ implies that $M_{t-1}(e_i * y), M_{t-1}((e_\emptyset - e_i) * y) \succeq 0$. ∎

3.2 The Sherali-Adams hierarchy

Let K be a semi-algebraic set as in (4), where the g_ℓ's are polynomials in which every variable occurs with degree at most 1 and let $P = \text{conv}(K \cap \{0,1\}^n)$ be the polytope to be described. Let $\mathcal{J}$ be as in relation (13). Let w_ℓ denote the degree of the polynomial g_ℓ, $v_\ell := \lceil \frac{w_\ell}{2} \rceil$, $w := \max w_\ell$, and $v := \max v_\ell$.

We now introduce the Sherali-Adams relaxations[1] as *linear* relaxations and then observe that they can be reformulated as the semidefinite programs (10). Let $t \in \{1, \ldots, n\}$. Multiply each inequality $g_\ell(x) \geq 0$ ($\ell = 1, \ldots, m$) by each product

$$f(I, J) := \prod_{i \in I} x_i \cdot \prod_{j \in J} (1 - x_j) \tag{14}$$

where I, J are disjoint subsets of $V = [1, n]$ such that $|I \cup J| = t$. Then, we obtain a set of inequalities that are still valid for P; add to this set all the inequalities $f(I, J) \geq 0$ where I, J are disjoint subsets with $|I \cup J| = \min(t+1, n)$. Replace each square x_i^2 by x_i and linearize the product $\prod_{i \in I} x_i$ by a new variable y_I for $I \subseteq V$ (thus setting $y_i = x_i$ for $i \in V$); this defines a set $R_t(K)$ in the space $\mathbb{R}^{\mathcal{P}_{t+w}(V)}$. As

$$\prod_{i \in I} x_i \cdot \prod_{j \in J} (1 - x_j) = \sum_{I \subseteq H \subseteq I \cup J} (-1)^{|H \setminus I|} \prod_{h \in H} x_h$$

the quantity obtained by linearizing $g_\ell(x) \prod_{i \in I} x_i \cdot \prod_{j \in J} (1 - x_j)$ reads

$$\Big(\sum_{K \subseteq V} g_\ell(K) \prod_{k \in K} x_k\Big) \cdot \Big(\sum_{I \subseteq H \subseteq I \cup J} (-1)^{|H \setminus I|} \prod_{h \in H} x_h\Big) = \sum_{I \subseteq H \subseteq I \cup J} (-1)^{|H \setminus I|} g_\ell * y(H).$$

[1] In their paper [SA90], the authors consider semi-algebraic sets of a special form, but the treatment extends to arbitrary semi-algebraic sets.

Therefore, $R_t(K)$ consists of the vectors $y \in \mathbb{R}^{\mathcal{P}_{t+w}(V)}$ satisfying the inequalities:

$$\sum_{I \subseteq H \subseteq U} (-1)^{|H \setminus I|} g_\ell * y(H) \geq 0 \;\text{ for all } \ell = 1, \ldots, m \;\text{ and all } I \subseteq U \subseteq V \text{ with } |U| = t, \tag{15}$$

$$\sum_{I \subseteq H \subseteq W} (-1)^{|H \setminus I|} y(H) \geq 0 \;\text{ for all } I \subseteq W \subseteq V \text{ with } |W| = \min(t+1, n). \tag{16}$$

In fact, the inequalities (15) (resp. (16)) remain valid for $R_t(K)$ for any U with $|U| \leq t$ (resp. any W with $|W| \leq \min(t+1, n)$); this follows from the fact that $f(I, J) = f(I \cup \{k\}, J) + f(I, J \cup \{k\})$ for any element $k \in V \setminus I \cup J$ and any disjoint $I, J \subseteq V$. By Lemma 2, $R_t(K)$ can be reformulated[2] as

$$R_t(K) = \{y \in \mathbb{R}^{\mathcal{P}_{t+w}(V)} \mid \begin{array}{l} M_U(g_\ell * y) \succeq 0 \;\text{ for all } U \subseteq V \text{ with } |U| = t \text{ and } \ell = 1, \ldots, m \\ M_W(y) \succeq 0 \text{ for all } W \subseteq V \text{ with } |U| = \min(t+1, n)\}. \end{array} \tag{17}$$

In view of Corollary 3 we find that

$$R_n(K) = \mathcal{C}_{\mathcal{J}}.$$

Let $S_t(K)$ denote the projection of $R_t(K) \cap \{y \mid y_\emptyset = 1\}$ on the subspace $\mathbb{R}^n$ indexed by the singletons. By the above, we have that

$$P = S_n(K) \subseteq \ldots \subseteq S_{t+1}(K) \subseteq S_t(K) \subseteq \ldots \subseteq S_1(K).$$

In general, the set $S_1(K)$ is *not* contained in K; this is due to the fact that $S_1(K)$ is convex while K need not be convex. (As an example, consider $K = \{x \in [0,1]^2 \mid x_1 + x_2 - x_1 x_2 \geq 1\}$ which is the union of two intervals, $K = \{x \in [0,1]^2 \mid x_1 = 1 \text{ or } x_2 = 1\}$, while $P = \{x \in [0,1]^2 \mid x_1 + x_2 \geq 1\}$.) In the linear case, i.e., when all polynomials g_ℓ have degree 1, then K is convex and $S_1(K) \subseteq K$.

Matrix reformulation. Let $\mathcal{K}$ denote the *linearization* of K defined by

$$\mathcal{K} := \{y \in \mathbb{R}^{\mathcal{P}_w(V)} \mid g_\ell^T y \geq 0 \;\text{ for } \ell = 1, \ldots, m, \; 0 \leq y_i \leq y_\emptyset \;\text{ for } i = 1, \ldots, n\}. \tag{18}$$

Given $y \in \mathbb{R}^{\mathcal{P}_{t+w}(V)}$, consider the matrix Y whose rows and columns are indexed, respectively, by $\mathcal{P}_w(V)$ and $\mathcal{P}_t(V)$ and with entries $Y(K, H) := y(K \cup H)$ for $K \in \mathcal{P}_w(V)$ and $H \in \mathcal{P}_t(V)$. Denote by e_H ($H \in \mathcal{P}_t(V)$) the elementary unit vectors in $\mathbb{R}^{\mathcal{P}_t(V)}$; then $Y e_H$ is the column of Y indexed by H. Then,

$$y \in R_t(K) \Longleftrightarrow Y \left(\sum_{I \subseteq H \subseteq U} (-1)^{|H \setminus I|} e_H \right) \in \mathcal{K} \;\text{ for } I \subseteq U \subseteq V \text{ with } |U| = t. \tag{19}$$

3.3 The Lasserre hierarchy

For $t \geq v - 1$, where $v = \max_{1 \leq \ell \leq m} v_\ell$, define

$$P_t(K) := \{y \in \mathbb{R}^{\mathcal{P}_{2t+2}(V)} \mid M_{t+1}(y) \succeq 0, \; M_{t+1-v_\ell}(g_\ell * y) \succeq 0 \text{ for } \ell = 1, \ldots, m\} \tag{20}$$

[2]Remark that we would obtain the same set $R_t(K)$ if instead of including the relations $f(I, J) \geq 0$ for disjoint I, J with $|I \cup J| = \min(t+1, n)$, we would include the relations $f(I, J)x_i \geq 0$, $f(I, J)(1 - x_i) \geq 0$ for disjoint I, J with $|I \cup J| = t$ and $i = 1, \ldots, n$; this follows from Lemma 5.

and define $Q_t(K)$ as the projection of $P_t(K) \cap \{y \mid y_\emptyset = 1\}$ onto $\mathbb{R}^n$. Therefore,

$$P \subseteq Q_{n+v-1}(K) \subseteq \ldots \subseteq Q_v(K) \subseteq Q_{v-1}(K).$$

In the case when $K = [0,1]^n$ (i.e., there is no additionnal polynomial constraint $g_\ell(x) \geq 0$), we let $v := 0$; the first relaxation $Q_{-1}(K)$ is trivial and can thus be ignored. Lasserre [Las01b] shows that $P = Q_{n+v-1}(K)$. This result is, in fact, a direct consequence of Corollary 3, since $P_{n+v-1}(K) = \mathcal{C}_{\mathcal{J}}$. More strongly, it follows from the fact that the Lasserre hierarchy refines the Sherali-Adams hierarchy.

Proposition 6. *For any $t = 1, \ldots, n$, $Q_{t+v-1}(K) \subseteq S_t(K)$ when $v \geq 1$ and $Q_t(K) \subseteq S_t(K)$ when $K = [0,1]^n$ (i.e., $v = 0$).*

Proof. Suppose that $v \geq 1$; we show that $Q_{t+v-1}(K) \subseteq S_t(K)$. Let $y \in P_{t+v-1}(K)$; that is, $y \in \mathbb{R}^{\mathcal{P}_{2t+2v}(V)}$ satisfies $M_{t+v}(y) \succeq 0$ and $M_{t+v-v_\ell}(g_\ell * y) \succeq 0$ for $\ell = 1, \ldots, m$. We verify that the restriction of y to $\mathcal{P}_{t+w}$ belongs to $R_t(K)$. Indeed, given $U, W \subseteq V$ with $|U| = t$ and $|W| = \min(t+1, n)$, $M_W(y) \succeq 0$ since it is a principal submatrix of $M_{t+v}(y)$ (as $v \geq 1$) and $M_U(g_\ell * y) \succeq 0$ since it is a principal submatrix of $M_{t+v-v_\ell}(g_\ell * y)$. The same argument shows the inclusion $Q_t(K) \subseteq S_t(K)$ when $K = [0,1]^n$. ∎

The construction of Lasserre is originally presented in terms of moment matrices indexed by integer sequences (rather than subsets of V) and the proof of convergence uses results about moment sequences and the representation of positive polynomials as sums of squares. We review Lasserre's approach in Section 7 and show that it is equivalent to the above presentation.

4 Comparing the Lasserre, Sherali-Adams and Lovász-Schrijver Relaxations

We assume here that K is polytope; that is, K is defined by (4) where all the polynomials g_ℓ have degree 1 (thus, $v = 1$, or $v = 0$ if $K = [0,1]^n$). As is well known, the first steps of the Sherali-Adams and Lovász-Schrijver hierarchies are then identical; that is, $S_1(K) = N(K)$. (To see it, compare (3) and (19).) It follows from results in [LS91] that $S_t(K) \subseteq N^t(K)$; that is, the Sherali-Adams hierarchy refines the Lovász-Schrijver hierarchy. The above inclusion also follows from Theorem 7 which has a simple direct proof.

Theorem 7. *If K is a polytope, then $S_t(K) \subseteq N(S_{t-1}(K))$ for all $t = 1, \ldots, n$ (setting $S_0(K) := K$).*

Proof. Let $t \geq 2$ and let $(y_1, \ldots, y_n)^T \in S_t(K)$; that is, $(y_1, \ldots, y_n)^T$ is the projection of some $y \in R_t(K)$ with $y_\emptyset = 1$. We show that the matrix $Y := M_1(y) = (y_{I \cup J})_{|I|,|J| \leq 1}$ belongs to $M(S_{t-1}(K))$; that is, Ye_k, $Y(e_\emptyset - e_k)$ belong to $\widetilde{S_{t-1}(K)}$, the homogenization of $S_{t-1}(K)$, for all $k = 1, \ldots, n$. As Ye_k (resp. $Y(e_\emptyset - e_k)$) is the projection on $\mathbb{R}^{\mathcal{P}_1(V)}$ of the vector $e_k * y$ (resp. $(e_\emptyset - e_k) * y$), it suffices to show that $e_k * y$ and $(e_\emptyset - e_k) * y$ belong to $R_{t-1}(K)$, i.e., that $M_W(e_k * y)$, $M_W((e_\emptyset - e_k) * y)$, $M_U(g_\ell * (e_k * y))$, $M_U(g_\ell * [(e_\emptyset - e_k) * y]) \succeq 0$ for all $\ell = 1, \ldots, m$, $U, W \subseteq V$ with $|U| = t-1$, $|W| = t$. This follows directly from the assumption that $y \in R_t(y)$ together with Lemma 5 and the commutation rule (7). ∎

Corollary 8. *$S_t(K) \subseteq N^t(K)$ for all $t = 1, \ldots, n$.*

Proof. Directly from Theorem 7 using induction on t. ∎

By Proposition 6, for any $t = 1, \ldots, n$, we have the inclusions:

$$Q_t(K) \subseteq S_t(K) \subseteq N^t(K).$$

In fact, one can show that the Lasserre hierarchy also refines the Lovász-Schrijver hierarchy obtained using the N_+ operator.

Observe that $M(K)$ can be alternatively viewed as the set of matrices $Y := M_1(y)$ where $y \in \mathbb{R}^{\mathcal{P}_2(V)}$ for which $Ye_k,\ Y(e_0 - e_k) \in \tilde{K}$, i.e., $g_\ell^T Y e_k,\ g_\ell^T Y(e_\emptyset - e_k) \geq 0$ for all $\ell = 1, \ldots, m$ and $k = 1, \ldots, n$. As $g_\ell^T Y e_0 = g_\ell * y(\emptyset)$, $g_\ell^T Y e_k = g_\ell * y(k)$, the latter holds if and only if the principal submatrix of $M_1(g_\ell * y)$ indexed by $\emptyset$ and $\{k\}$ is positive semidefinite. In comparison, membership in $Q_0(K)$ requires only that $g_\ell * y(\emptyset) \geq 0$ for all ℓ, while membership in $Q_1(K)$ requires that $M_1(g_\ell * y) \succeq 0$ for all ℓ. Therefore, we have the following inclusions:

$$Q_1(K) \subseteq N_+(K) \subseteq Q_0(K). \tag{21}$$

Theorem 9. *If K is a polytope, then $Q_t(K) \subseteq N_+(Q_{t-1}(K))$ for all $t = 1, \ldots, n$.*

Proof. Let $(y_1, \ldots, y_n)^T \in Q_t(K)$; that is, $(y_1, \ldots, y_n)^T$ is the projection of some $y \in P_t(K)$ with $y_\emptyset = 1$. Set $Y := M_1(y)$. We show that $Ye_k,\ Y(e_\emptyset - e_k) \in \widetilde{Q_{t-1}(K)}$, the homogenization of $Q_{t-1}(K)$, for $k = 1, \ldots, n$. As Ye_k (resp. $Y(e_\emptyset - e_k)$) is the projection on $\mathbb{R}^{\mathcal{P}_1(V)}$ of $e_k * y$ (resp. $(e_\emptyset - e_k) * y$), it suffices to show that $e_k * y$ and $(e_\emptyset - e_k) * y$ belong to $P_{t-1}(K)$, i.e., that $M_t(e_k * y)$, $M_t((e_\emptyset - e_k) * y)$, $M_{t-1}(g_\ell * (e_k * y))$, $M_{t-1}(g_\ell * [(e_\emptyset - e_k) * y]) \succeq 0$ for all $\ell = 1, \ldots, m$. This follows directly from the assumption that $y \in P_t(K)$ together with Lemma 5 and (7). ∎

Corollary 10. *If K is a polytope, then $Q_t(K) \subseteq N_+^t(K)$ for all $t = 1, \ldots, n$.*

Proof. Directly from Theorem 9 and (21) using induction on t. ∎

An algorithmic comparison. Summarizing, we have:

$$Q_t(K) \subseteq S_t(K) \cap N_+^t(K)$$

for any $t = 1, \ldots, n$. Therefore, the Lasserre set $Q_t(K)$ provides the sharpest relaxation of P. From an algorithmic point of view, it is however less well behaved than the Sherali-Adams and Lovász-Schrijver relaxations.

Given a convex body $B \subseteq \mathbb{R}^n$, the separation problem for B is the problem of determining whether a given vector $y \in \mathbb{R}^n$ belongs to B and, if not, of finding a hyperplane separating y from B; the weak separation problem is the analogue problem where one allows for numerical errors. An important consequence of the ellipsoid method is that, if one can solve the weak separation problem for B in polynomial time, then one can optimize any linear objective function over B in polynomial time (with an arbitrary precision) and vice versa (assuming some technical information about B like the knowledge of a ball contained in B and of a ball containing B); see [GLS88] for details.

If one can solve the weak separation problem for K in polynomial time, then the same holds for $M(K)$ and $M_+(K)$ and, thus, for the projections $N(K)$ and $N_+(K)$. Therefore, one can optimize a linear objective function in polynomial time over the relaxations $N^t(K)$, $N_+^t(K)$, $S_t(K)$ for any *fixed* t; this is observed in [LS91] for the LS sets and the same argument works for the SA sets in view of the matrix reformulation of the SA method. The assumption made over K is trivially satisfied if m is polynomial in n but it may sometimes be satisfied even if m is exponential in n. On the other hand, in order to claim that one can optimize over $Q_t(K)$ in polynomial time, one needs to assume that m is polynomial in n, since the system defining $Q_t(K)$ involves m LMI's associated to the inequalities of the linear system defining K.

5 The rank of the Sherali-Adams Procedure

We present here two examples of a polytope $K \subseteq [0,1]^n$ for which n iterations of the Sherali-Adams procedure are needed for finding the integer polytope $P = \mathrm{conv}(K \cap \{0,1\}^n)$.

Example 1. Let

$$K := \{x \in [0,1]^n \mid \sum_{r\in R}(1-x_r) + \sum_{r\in V\setminus R} x_r \geq \frac{1}{2} \ \text{ for all } R \subseteq [1,n]\}; \tag{22}$$

then $P = \emptyset$. We show in Proposition 11 below that $S_{n-1}(K) \neq \emptyset$, which implies that $P \neq S_{n-1}(K)$. The polytope K has been used earlier to show that n iterations are needed for the following procedures: taking Chvátal cuts [CCH89], the N_+ operator [GT00], the N_+ operator combined with taking Chvátal cuts [CD01], and the N_+ operator combined with taking Gomory mixed integer cuts (equivalent to disjunctive cuts) [CL01]. The following (easy to verify) identities will be used in the proof:

$$\sum_{K\subseteq A} \frac{(-1)^{|K|}}{2^{|K|}} = \frac{1}{2^{|A|}}, \quad \sum_{K\subseteq A} |K| \frac{(-1)^{|K|}}{2^{|K|}} = -\frac{|A|}{2^{|A|}} \tag{23}$$

for any set A. (For the second one, use the fact that $k\binom{n}{k} = n\binom{n-1}{k-1}$.)

Proposition 11. *Let $y \in \mathbb{R}^{\mathcal{P}(V)}$ with entries $y_I := \frac{1}{2^{|I|}}$ $(I \subseteq V)$. Then, $y \in R_{n-1}(K)$ where K is defined by* (22).

Proof. Let $v_R \in \mathbb{R}^{\mathcal{P}(V)}$ be the vector of coefficients of an inequality defining K, with all components zero except $v_R(\emptyset) = -\frac{1}{2} + |R|$, $v_R(r) = -1$ if $r \in R$, $v_R(r) = 1$ if $r \in V \setminus R$, where R is a given subset of V. Then, for $H \subseteq V$,

$$\begin{aligned} v_R * y(H) &= (|R| - \tfrac{1}{2})y(H) - \textstyle\sum_{r\in R} y(H \cup \{r\}) + \sum_{r\in V\setminus R} y(H\cup\{r\}) \\ &= \tfrac{1}{2^{|H|}}(|R| - \tfrac{1}{2} + |H \setminus R| - |R\cap H|) + \tfrac{1}{2^{|H|+1}}(|V \setminus (H\cup R)| - |R \setminus H|) \\ &= \tfrac{1}{2^{|H|+1}}(n - 1 + |H| - 2|R \cap H|). \end{aligned}$$

Given a subset $U \subseteq V$ with $|U| = n-1$ and $I \subseteq U$, we have:

$$\begin{aligned}\varphi := \sum_{I \subseteq H \subseteq U} (-1)^{|H \setminus I|} (v_R * y)(H) &= \frac{n-1}{2^{|I|+1}} \sum_{K \subseteq U \setminus I} \frac{(-1)^{|K|}}{2^{|K|}} \\ + \tfrac{1}{2^{|I|+1}} \sum_{K \subseteq U \setminus I} \frac{(-1)^{|K|}}{2^{|K|}} (|I| + |K|) &- \frac{1}{2^{|I|+1}} \sum_{K \subseteq U \setminus I} \frac{(-1)^{|K|}}{2^{|K|}} (|R \cap I| + |R \cap K|).\end{aligned}$$

Using (23), one can verify that the second term in the above expression of φ is equal to $\frac{1}{2^n}(2|I| - n + 1)$ while the third term is equal to $\frac{1}{2^{n-1}}(2|R \cap I| - |R \cap U|)$. Therefore,

$$\varphi = \frac{1}{2^{n-1}} (|I| + |R \cap U| - 2|R \cap I|) \geq 0$$

since $I \subseteq U$. By Lemma 2 (ii), this shows that $M_U(v_R * y) \succeq 0$.

Finally, $M_V(y) \succeq 0$, since $\sum_{I \subseteq H} (-1)^{|H \setminus I|} y_H = \frac{1}{2^n} \geq 0$. ∎

Example 2. Consider the polytope

$$K := \{x \in [0,1]^n \mid \sum_{i=1}^{n} x_i \geq \frac{1}{2}\}, \tag{24}$$

then $P = \{x \in [0,1]^n \mid \sum_{i=1}^{n} x_i \geq 1\}$. This example was considered by Cook and Dash [CD01] as an example where the Lovász-Schrijver rank is n. The next result shows that the Sherali-Adams rank is also equal to n.

Proposition 12. *Let $y \in \mathcal{P}(V)$ with zero entries except $y_\emptyset := 1$ and $y_i := \frac{1}{n+1}$ ($i \in V$). Then, $y \in R_{n-1}(K)$ where K is defined by* (24)*. Therefore, $P \subset S_{n-1}(K)$.*

Proof. One can easily verify (using Lemma 2 (ii)) that $M_V(y) \succeq 0$ and $M_U(g * y) \succeq 0$ for $U \subseteq V$ with $|U| = n-1$, where $g(x)$ is the polynomial $-\frac{1}{2} + \sum_{i=1}^{n} x_i$. ∎

It would be interesting to determine the Lasserre rank of the polytope K in the above two examples. In the second example, when K is defined by (24), we verified that the Lasserre rank is equal to n when $n = 2$; indeed, the minimum value of $x_1 + x_2$ for $x \in Q_1(K)$ is equal to $\frac{25}{26} < 1$. It is not clear how to construct a point $x \in Q_{n-1}(K)$ with $\sum_i x_i < 1$ for general $n \geq 2$.

On the other hand, when K is given by (22), we verified that the Lasserre rank of K is equal to 1 when $n = 2$. Again it would be interesting to see what is the exact rank for higher values of n (we believe that $n-1$ is the correct value).

6 Two Applications

We describe here how the various lift and project methods apply to two concrete examples, namely, to the stable set polytope and to the cut polytope of a graph. They are the two most extensively studied examples with respect to this class of methods; the original paper by Lovász and Schrijver

[LS91] studies the stable set problem while the paper [La00] studies the case of max-cut. Moreover, these two examples have been the objects of milestone results in the field of semidefinite optimization.

Indeed, the idea of constructing semidefinite relaxations for a combinatorial problem goes back to the seminal work of Lovász [Lo79] who introduced the semidefinite bound $\vartheta(G)$ for the stability number of a graph G, obtained by optimizing over the semidefinite relaxation $\mathrm{TH}(G)$ of the stable set polytope $\mathrm{ST}(G)$ of G. An important result is that $\mathrm{TH}(G) = \mathrm{ST}(G)$ precisely when G is a perfect graph, in which case one can solve the maximum stable set problem in polynomial time (with an arbitrary precision) using semidefinite programming; this is still the only polynomial time algorithm known up to today (cf [GLS88]).

This idea of approximating combinatorial problems using semidefinite relaxations was used later again successfully by Goemans and Williamson [GW95] who, using a basic semidefinite relaxation of the cut polytope, could prove a good approximation algorithm for the max-cut problem. Since then, semidefinite relaxations have been widely used (in conjonction with clever rounding schemes) for constructing good approximation algorithms for a large number of combinatorial problems. It is therefore of interest to construct new stronger semidefinite relaxations for the stable set and cut problems, as they could potentially be used for designing better approximation algorithms.

6.1 Application to the stable set polytope

Given a graph $G = (V = [1, n], E)$, let $\mathrm{ST}(G)$ denote the stable set polytope of G, let

$$\mathrm{FR}(G) := \{x \in \mathbb{R}^n_+ \mid x_i + x_j \leq 1 \ \forall ij \in E\}$$

be its basic linear relaxation defined by nonnegativity and the edge inequalities, and let

$$\begin{array}{ll}\mathrm{TH}(G) := \{x \in \mathbb{R}^n \mid & (1\ x) = Ye_0 \text{ for some positive semidefinite matrix } Y = (Y_{ij})_{i,j=0}^n \\ & \text{satisfying } Y_{ii} = Y_{0i}\ (i \in V),\ Y_{ij} = 0\ (ij \in E)\}\end{array} \tag{25}$$

be the basic semidefinite relaxation of $\mathrm{ST}(G)$. Let us compare how the various lift and project methods apply to the pair $P := \mathrm{ST}(G)$, $K := \mathrm{FR}(G)$.

Define the *N-rank* (resp. *N_+-rank*) of $\mathrm{FR}(G)$ as the smallest integer t for which $N^t(\mathrm{FR}(G)) = \mathrm{ST}(G)$ (resp. $N^t_+(\mathrm{FR}(G)) = \mathrm{ST}(G)$); define similarly the *SA-rank* and the *Lasserre rank* of $\mathrm{FR}(G)$.

The relaxations $N(\mathrm{FR}(G))$ and $N_+(\mathrm{FR}(G))$ are studied in detail in [LS91]. In particular, the following results are shown there. The polytope $N(\mathrm{FR}(G))$ is defined by the nonnegativity and edge constraints together with the odd hole inequalities: $\sum_{i \in V(C)} x_i \leq \frac{|C|-1}{2}$ for C odd hole in G. If G has n nodes and stability number $\alpha(G)$, then its N-rank t satisfies:

$$\frac{n}{\alpha(G)} - 2 \leq t \leq n - \alpha(G) - 1; \tag{26}$$

the N-rank t of an inequality $a^T x \leq \beta$ valid for $\mathrm{ST}(G)$ (with integer coefficients and distinct from the nonnegativity constraints) satisfies:

$$\frac{1}{\beta}\left(\sum_{i \in V} a_i - 2\beta\right) \leq t \leq \sum_{i \in V} a_i - 2\beta. \tag{27}$$

The lower bounds follow from the fact that

$$\frac{1}{t+2}(1, \ldots, 1)^T \in N^t(\mathrm{FR}(G)) \tag{28}$$

for any $t \geq 0$. The N_+ operator yields a much stronger relaxation, as clique inequalities, odd wheel and odd antihole inequalities are valid for $N_+(\mathrm{FR}(G))$ (while the N-rank of a clique inequality based on a clique of size k is $k-2$). Thus, perfect graphs have N_+-rank 1. Moreover,

$$N_+(\mathrm{FR}(G)) \subseteq \mathrm{TH}(G)$$

for any graph G and the N_+ rank t of G satisfies:

$$t \leq \alpha(G). \tag{29}$$

The Sherali-Adams hierarchy does not seem to yield a significant improvement with respect to the sequence $N^t(\mathrm{FR}(G))$. Indeed, the vector $\frac{1}{t+2}(1,\ldots,1)^T \in \mathbb{R}^n$ considered in (28) belongs also to $S_t(\mathrm{FR}(G))$. (Because the vector $y \in \mathbb{R}^{\mathcal{P}_{t+1}(V)}$ defined by $y_\emptyset := 1$, $y_I := \frac{1}{t+2}$ if $|I| = 1$, and $y_I := 0$ if $|I| \geq 2$ belongs to $R_t(\mathrm{FR}(G))$.) Therefore, the lower bounds from (26) and (27) remain valid for the SA-rank of $\mathrm{FR}(G)$.

On the other hand, the Lasserre hierarchy does improve on the sequence $N_+^t(\mathrm{FR}(G))$ as we now see. We begin with giving a more compact formulation for the relaxation $Q_t(\mathrm{FR}(G))$. For an edge $ab \in E$, let $g_{ab}(x) := 1 - x_a - x_b$ be the polynomial corresponding to the edge inequality $x_a + x_b \leq 1$. We show that the positive semidefinite constraint $M_t(g_{ab} * y) \succeq 0$ can be replaced by the linear equation: $y_{ab} = 0$.

Lemma 13. *Let $t \geq 1$ and $y \in \mathbb{R}^{\mathcal{P}_{2t+2}(V)}$. The following assertions are equivalent.*

(i) $y \in P_t(\mathrm{FR}(G))$

(ii) $M_{t+1}(y) \succeq 0$ *and* $y_{ab} = 0$ *for any edge* $ab \in E$.

(iii) $M_{t+1}(y) \succeq 0$ *and* $y_I = 0$ *for any* $I \in \mathcal{P}_{2t+2}(V)$ *which is not stable.*

Proof. Note first that the condition $M_{t+1}(y) \succeq 0$ implies that $y_I \geq 0$ for all $I \in \mathcal{P}_{t+1}(V)$.
(i) $\Longrightarrow$ (ii) The (a,a)-th entry of $M_t(g_{ab} * y)$ is equal to $g_{ab} * y(a) = -y_{ab}$ and is nonnegative, which implies that $y_{ab} = 0$.
(ii) $\Longrightarrow$ (iii) Suppose I contains the edge ab. If $|I| \leq t+1$, then the (ab, I)-th entry of $M_{t+1}(y)$ is equal to 0 since the (ab, ab)-th entry is 0, which implies that $y_I = 0$. Otherwise, write $I = I_1 \cup I_2$ where $I_1, I_2 \in \mathcal{P}_{t+1}(V)$ with $\{a,b\} \subseteq I_1$; by the above the (I_1, I_1)-th entry of $M_{t+1}(y)$ is 0 and, thus, its (I_1, I_2)-th entry too is 0, implying $y_I = 0$.
(iii) $\Longrightarrow$ (i) We show that $M_t(g_{ab} * y) \succeq 0$. Set $\mathcal{P}_0 := \mathcal{P}_t(V \setminus \{a,b\})$ and $\mathcal{P}_c := \{I \cup \{c\} \mid I \in \mathcal{P}_0\}$ for $c = a$ or b. Then, the principal submatrix X of $M_t(y)$ indexed by $\mathcal{P}_0 \cup \mathcal{P}_a \cup \mathcal{P}_b$ has the form: $\begin{array}{c c} & \begin{array}{ccc} \mathcal{P}_0 & \mathcal{P}_a & \mathcal{P}_b \end{array} \\ \begin{array}{c} \mathcal{P}_0 \\ \mathcal{P}_a \\ \mathcal{P}_b \end{array} & \begin{pmatrix} C & A & B \\ A & A & 0 \\ B & 0 & B \end{pmatrix} \end{array}$ and $X \succeq 0$ implies that $C - A - B \succeq 0$. (To see it, note that $(-x,x,x)^T X(-x,x,x) = x^T(C-A-B)x$ for all $x \in \mathbb{R}^p$, $p := |\mathcal{P}_0|$.) The result now follows since, with respect to the partition of $\mathcal{P}_t(V)$ into $\mathcal{P}_0$ and its complement $\mathcal{P}_0'$, the matrix $M_t(g_{ab} * y)$ has the

form: $\begin{array}{c c} & \begin{array}{cc} \mathcal{P}_0 & \mathcal{P}_0' \end{array} \\ \begin{array}{c} \mathcal{P}_0 \\ \mathcal{P}_0' \end{array} & \begin{pmatrix} C - A - B & 0 \\ 0 & 0 \end{pmatrix} \end{array}.$ ∎

In view of Corollary 10 and (29), it follows that $Q_{\alpha(G)}(\mathrm{FR}(G)) = \mathrm{ST}(G)$. In fact, the Lasserre hierarchy already finds $\mathrm{ST}(G)$ at step $\alpha(G) - 1$.

Proposition 14. $\mathrm{ST}(G) = Q_{\alpha(G)-1}(\mathrm{FR}(G))$ *for a graph G with stability number $\alpha(G) \geq 2$.*

Proof. We show that $Q_{\alpha-1}(\mathrm{FR}(G)) \subseteq Q_n(\mathrm{FR}(G))$, where $\alpha := \alpha(G)$. Let $y \in P_{\alpha-1}(\mathrm{FR}(G))$; define $z \in \mathbb{R}^{\mathcal{P}(V)}$ by $z_I := y_I$ if $|I| \leq 2\alpha$ and $z_I := 0$ otherwise. Thus, $z_{ab} = 0$ for all edges $ab \in E$. By Lemma 13, it suffices to verify that $M_V(z) \succeq 0$, which holds since, with respect to the partition of $\mathcal{P}(V)$ into $\mathcal{P}_\alpha(V)$ and its complement, $M_V(z)$ has the form $\begin{pmatrix} M_\alpha(y) & 0 \\ 0 & 0 \end{pmatrix}$. ∎

Let G be the line graph of K_n with n odd; then, $\mathrm{ST}(G)$ is the matching polytope of K_n. Stephen and Tunçel [ST99] show that $\alpha(G) = \frac{n-1}{2}$ iterations of the N_+ operator are needed for finding $\mathrm{ST}(G)$. Thus, this gives an instance of a graph G for which $\mathrm{ST}(G) = Q_{\alpha-1}(\mathrm{FR}(G))$ is strictly contained in $N_+^{\alpha-1}(\mathrm{FR}(G))$.

We conclude with a comparison with the basic semidefinite relaxation $\mathrm{TH}(G)$. By the definition (25), $\mathrm{TH}(G)$ can be seen as the projection on $\mathbb{R}^n$ of the set of vectors $y \in \mathbb{R}^{\mathcal{P}_2(V)}$ satisfying $y_\emptyset = 1$ and

$$M_1(y) \succeq 0, \ y_{ab} = 0 \ (ab \in E).$$

Therefore, we have the following chain of inclusions:

$$Q_1(\mathrm{FR}(G)) \subseteq N_+(\mathrm{FR}(G)) \subseteq \mathrm{TH}(G) \subseteq Q_0(\mathrm{FR}(G))$$

and, in view of Lemma 13, the Lasserre relaxations $Q_t(\mathrm{FR}(G))$ $(t \geq 1)$ are natural refinements of the basic SDP relaxation $\mathrm{TH}(G)$.

6.2 Application to the max-cut problem

Given a graph $G = (V = [1, n], E)$, the max-cut problem asks for a partition $(S, V \setminus S)$ minimizing the total cardinality (or weight) of the edges ij cut by the partition (i.e., such that $|S \cap \{i, j\}| = 1$). Hence it can be formulated as an unconstrained quadratic ± 1-problem:

$$\max(x^T Ax \mid x \in \{\pm 1\}^n), \tag{30}$$

where A is a (suitably defined) symmetric matrix, but the treatment below remains valid for A arbitrary.

Since we are now working with ± 1 variables in place of $0-1$ variables, we should modify some definitions. In particular, when defining the moment matrices in (5), one should consider the semigroup $\mathcal{P}(V)$ with the symmetric difference Δ as semigroup operation in place of the union; thus the (I, J)-th entry of the moment matrix is $y(I \Delta J)$. Moreover, (6) becomes: $x{*}y(I) = \sum_K x_I y_{I\Delta J}$, the zeta-matrix Z has entry $Z_{IJ} = (-1)^{|I \Delta J|}$, the inequalities (15) defining $R_t(K)$ become: $\sum_{H \subseteq U} (-1)^{|H \cap I|} g_\ell{*}y(H) \geq 0$.

There are two possible strategies in order to formulate relaxations for the problem (30).

First strategy. The first possible strategy is to formulate (30) as a linear problem

$$\max(\langle A, X\rangle \mid X \in \mathrm{CUT}(K_n))$$

over the *cut polytope*

$$\mathrm{CUT}(K_n) := \mathrm{conv}(xx^T \mid x \in \{\pm 1\}^n)$$

(which is in fact a $\binom{n}{2}$-dimensional polytope) and to apply the various lift and project methods to some linear programming formulation of $\mathrm{CUT}(K_n)$. As linear programming formulation for $\mathrm{CUT}(K_n)$, one can take the *metric polytope* $\mathrm{MET}(K_n)$ consisting of the symmetric matrices X with diagonal entries 1 satisfying the *triangle inequalities*:

$$X_{ij} + X_{ik} + X_{jk} \geq -1,\ X_{ij} - X_{ik} - X_{jk} \geq -1$$

for all distinct $i, j, k \in V$.

One can also consider linear relaxations of the cut polytope $\mathrm{CUT}(G)$ of an arbitrary graph G. Given a graph $G = (V, E)$, let $\mathrm{CUT}(G)$ and $\mathrm{MET}(G)$ denote the projections of $\mathrm{CUT}(K_n)$ and $\mathrm{MET}(K_n)$, respectively, on the subspace $\mathbb{R}^E$ indexed by the edge set of G. Then, $\mathrm{CUT}(G) \subseteq \mathrm{MET}(G)$ with equality if and only if G has no K_5-minor [BM86].

When applying the Lovász-Schrijver construction to $K := \mathrm{MET}(G)$, one finds the relaxation $N(\mathrm{MET}(G))$ of $\mathrm{CUT}(G)$. Another possibility is to first apply the LS construction to $K := \mathrm{MET}(K_n)$ and then project back on the edge space $\mathbb{R}^E$, thus yielding the relaxation $N(G) := \pi_E(N(\mathrm{MET}(K_n)))$ of $\mathrm{CUT}(G)$ (with π_E denoting the projection from the space indexed by the edge set of K_n to the space indexed by the edge set of G). One has:

$$N(G) \subseteq N(\mathrm{MET}(G))$$

but it is not known whether equality holds in general.

The following results about the relaxations $N(G)$ and $N(\mathrm{MET}(G))$ are shown in [La00]. Equality: $N^t(\mathrm{MET}(G)) = \mathrm{CUT}(G)$ holds if G has t edges whose contraction produces a graph with no K_5-minor. In particular, $N^{n-\alpha(G)-3}(G) = \mathrm{CUT}(G)$; moreover, $N^{n-\alpha(G)-3}(\mathrm{MET}(G)) = \mathrm{CUT}(G)$ if G has a maximum stable set whose complement induces a graph with at most three connected components. In particular, $N^{n-4}(K_n) = \mathrm{CUT}(K_n)$ for $n \geq 4$. The value $n-4$ is known to be the correct value for the N-rank of $\mathrm{MET}(K_n)$ when $n \leq 7$ and is conjectured to be the correct value for any n. Although the inclusion $N_+(\mathrm{MET}(G)) \subseteq N(\mathrm{MET}(G))$ is strict in general (e.g., for $G = K_n$ and $n \geq 6$), no example is known of a graph for which the number of iterations needed for finding $\mathrm{CUT}(G)$ is smaller when using the N_+ operator instead of the N operator.

Second strategy. Another possible strategy is to apply the various constructions to the cube $K := C_n = [-1, 1]^n$ and to take projections on the space $\mathbb{R}^{E_n}$ indexed by the set E_n of pairs ij of points of V (instead of projections on the space $\mathbb{R}^V$ indexed by the singletons). Thus we now consider the Sherali-Adams set $R_t(C_n)$ and the Lasserre set $P_t(C_n)$ and their respective projections $\hat{S}_t(C_n)$ and $\hat{Q}_t(C_n)$ on $\mathbb{R}^{E_n}$. (The 'hat' symbol is meant to remind that the projection is taken over the set of pairs.) As no polynomial constraint is present in the definition of K, we have that

$$\hat{S}_{n-1}(C_n) = \hat{Q}_{n-1}(C_n) = \mathrm{CUT}(K_n).$$

By the definition, the relaxation $\hat{S}_t(C_n)$ consists of the vectors $y \in \mathbb{R}^{E_n}$ whose restriction on a subset of $t+1$ points belongs to $\mathrm{CUT}(K_{t+1})$; in other words, $\hat{S}_t(C_n)$ is the polytope in $\mathbb{R}^{E_n}$ determined by all the valid inequalities for $\mathrm{CUT}(K_n)$ on at most $t+1$ points.

For $t \geq 0$, the t-th Lasserre relaxation of the max-cut problem reads:

$$\max(\sum_{i,j \in V} a_{ij} y_{ij} \mid M_{t+1}(y) = (y(I \Delta J))_{I,J \in \mathcal{P}_{t+1}(V)} \succeq 0,\ y_\emptyset = 1). \tag{31}$$

Let $\tilde{M}_{t+1}(y)$ denote the principal submatrix of $M_{t+1}(y)$ whose rows and columns are indexed by the sets $I \in \mathcal{P}_{t+1}(V)$ having even (resp. odd) cardinality if $t+1$ is even (resp. odd). The program (31) can be reformulated as the smaller program:

$$\max(\sum_{i,j \in V} a_{ij} y_{ij} \mid \tilde{M}_{t+1}(y) \succeq 0,\ y_\emptyset = 1). \tag{32}$$

Indeed, $M_{t+1}(y) = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix}$, where A is the submatrix of $M_{t+1}(y)$ indexed by all even sets and B the submatrix indexed by all odd sets. As the objective function does not involve any variable y_I with $|I|$ odd, we can assume that $C = 0$. Moreover, A is a submatrix of B if $t+1$ odd and vice-versa if $t+1$ is even. (To see it, use the fact that $I \Delta J = (I \Delta \{1\}) \Delta (J \Delta \{1\})$.)

Therefore, we find again the following facts observed by Lasserre [Las00]. For $t = 0$, the feasible set of the program (32) is the basic semidefinite relaxation $\mathcal{E}_n$ consisting of the semidefinite matrices of order n with diagonal entries 1. For $t = 1$, the feasible set of the program (32) is the set $\mathcal{F}'_n$ consisting of the positive semidefinite matrices Z indexed by $E_n \cup \{\emptyset\}$ satisfying

$$Z_{ij,ik} = Z_{\emptyset,jk} \text{ and } Z_{ij,rs} = Z_{ir,js} = Z_{is,jr}$$

for all distinct $i, j, k, r, s \in V$. If we remove in the definition of $\mathcal{F}'_n$ the condition $Z_{ij,rs} = Z_{ir,js} = Z_{is,jr}$, we obtain the larger matrix set $\mathcal{F}_n$ underlying the relaxation (SDP3) defined by Anjos and Wolkowicz [AW01]. Setting

$$F_n := \{x \in \mathbb{R}^{E_n} \mid (1\ x) = Ze_0 \text{ for some } Z \in \mathcal{F}_n\}$$

we have:

$$\mathrm{CUT}(K_n) \subseteq \hat{Q}_1(C_n) \subseteq F_n \subseteq \mathcal{E}_n \cap \mathrm{MET}(K_n).$$

The right most inclusion is shown in [AW01]; both left and right most inclusions are strict for $n = 5$.

It is shown in [La00] that $M_+(\mathrm{MET}(K_n)) \subseteq \mathcal{F}_n$ (M_+ being the Lovász-Schrijver matrix operator introduced in Section 2) and $M'_+(\mathrm{MET}(K_n)) \subseteq \mathcal{F}'_n$ (M'_+ being a strenghtening of M_+ considered in [La00]). Therefore, applying the operator M'_+ yields a relaxation $N'_+(\mathrm{MET}(K_n))$ which is contained in the Lasserre relaxation $\hat{Q}_1(C_n)$. The inclusion $N'_+(\mathrm{MET}(K_n)) \subseteq \hat{Q}_1(C_n)$ is strict for $n = 5$, since $N'_+(\mathrm{MET}(K_5)) = \mathrm{CUT}(K_5)$.

7 Lasserre's Approach Revisited

In this section we revisit the hierarchy of relaxations of Lasserre introduced in Section 3 from the algebraic point of view of representing nonnegative polynomials as sums of squares and the dual theory of moments. This approach applies to general (not necessarily $0-1$) polynomial programming. The idea of approximating polynomial programming problems using sums of squares of polynomials has been used in several other works, in particular, by Shor [Sh87, Sh98], Nesterov [Ne00], Parillo [Pa00], De Klerk and Pasechnik [KP01]. We begin with introducing the main ideas on the unconstrained problem of minimizing a polynomial function over $\mathbb{R}^n$, considered in [Las01a].

7.1 A gentle introduction

Suppose we want to solve the problem:

$$p^* := \min g(x) \text{ subject to } x \in \mathbb{R}^n, \tag{33}$$

where $g(x)$ is a polynomial of even degree $2d$ which can be assumed to satisfy $g(0) = 0$. It is easy to see that (33) can be reformulated as

$$p^* = \min_{\mu} \int g(x) d\mu(x) \tag{34}$$

where the minimum is taken over all probability measures μ on $\mathbb{R}^n$. Write the polynomial $g(x)$ as sum of monomials: $g(x) = \sum_{\alpha \in \mathcal{S}_{2d}} g_\alpha x^\alpha$, where $x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$ and, for an integer m, $\mathcal{S}_m$ denotes the set of $\alpha \in \mathbb{Z}_+^n$ with $|\alpha| := \sum_{i=1}^n \alpha_i \leq m$. Then, $\int g(x) d\mu(x) = \sum_\alpha g_\alpha \int x^\alpha d\mu(x)$. If we define a sequence $y = (y_\alpha)_{\alpha \in \mathcal{S}_{2d}}$ to be a *moment sequence* when

$$y_\alpha = \int x^\alpha d\mu(x) \tag{35}$$

(for all $\alpha \in \mathcal{S}_{2d}$) for some nonnegative measure μ on $\mathbb{R}^n$, then (34) can be rewritten as

$$p^* = \min \sum_\alpha g_\alpha y_\alpha \quad \text{s.t. } y \text{ is a moment sequence and } y_0 = 1. \tag{36}$$

Lower bounds for p^* can be obtained by relaxing the condition that y be a moment sequence. A necessary condition for y to be a moment sequence is that its *moment matrix* $M_d^{\mathbb{Z}}(y) := (y_{\alpha+\beta})_{\alpha,\beta \in \mathcal{S}_d}$ be positive semidefinite. Write

$$M_d^{\mathbb{Z}}(y) = \sum_{\gamma \in \mathcal{S}_{2d}} y_\gamma B_\gamma \tag{37}$$

where $B_\gamma := \sum_{\alpha,\beta \in \mathcal{S}_d | \alpha+\beta=\gamma} E_{\alpha,\beta}$, with $E_{\alpha,\beta}$ denoting the elementary matrices (with all zero entries except ones at the positions (α, β) and (β, α)). Therefore, one has the following lower bound for p^*:

$$p^* \geq \begin{array}{ll} \min & g^T y \\ \text{s.t.} & M_d^{\mathbb{Z}}(y) \succeq 0 \\ & y_0 = 1 \end{array} = \begin{array}{ll} \min & g^T y \\ \text{s.t.} & B_0 + \sum_{\gamma \in \mathcal{S}_{2d} \setminus \{0\}} B_\gamma y_\gamma \succeq 0 \end{array} \tag{38}$$

One can also proceed in the following dual manner for computing p^*. Rewrite (33) as

$$p^* = \max \ \lambda \ \text{ subject to } g(x) - \lambda \geq 0 \ \forall x \in \mathbb{R}^n. \tag{39}$$

Lower bounds for p^* are now obtained by considering sufficient conditions for $g(x) - \lambda$ to be nonnegative over $\mathbb{R}^n$. An obvious sufficient condition being that $g(x) - \lambda$ be a sum of squares (SOS, for short). Testing whether a polynomial $p(x)$ is a SOS amounts to deciding feasibility of a semidefinite program (see, e.g., [Ne00],[Pa00]). Say, $p(x)$ has degree $2d$ and let $z := (x^\alpha)_{\alpha\in\mathcal{S}_d}$ be the vector consisting of all monomials of degree $\le d$. Then, one can easily verify that $p(x)$ is a SOS if and only if $p(x) = z^T X z$ (identical polynomials) for some positive semidefinite matrix X. As

$$z^T X z = \sum_{\alpha,\beta\in\mathcal{S}_d} X_{\alpha,\beta} x^{\alpha+\beta} = \sum_{\gamma\in\mathcal{S}_{2d}} x^\gamma \left(\sum_{\substack{\alpha,\beta\in\mathcal{S}_d \\ \alpha+\beta=\gamma}} X_{\alpha,\beta} \right) = \sum_{\gamma\in\mathcal{S}_{2d}} x^\gamma \langle B_\gamma, X\rangle,$$

it follows that $p(x)$ is a SOS if and only if the following SDP program:

$$X \succeq 0,\ \langle B_\gamma, X\rangle = p_\gamma\ (\gamma\in\mathcal{S}_{2d}) \tag{40}$$

is feasible, where X is of order $\binom{n+d}{d}$ and with $\binom{n+2d}{2d}$ equations (thus, polynomially solvable for fixed n or d). Based on this we can formulate the following lower bound for p^*:

$$\begin{array}{llllll} p^* \ge & \max & \lambda & = & \max & -\langle B_0, X\rangle \\ & \text{s.t.} & g(x) - \lambda \ \text{ is SOS} & & \text{s.t.} & \langle B_\gamma, X\rangle = g_\gamma\ (\gamma \in \mathcal{S}_{2d}\setminus\{0\}). \end{array} \tag{41}$$

The SDP programs (38) and (41) are, in fact, dual of each other and there is no duality gap if (41) is feasible.

The lower bound from (41) is equal to p^* if $g(x) - p^*$ is a SOS; this holds for $n = 1$ but not in general if $n \ge 2$. In general one can estimate p^* asymptotically by a sequence of SDP's analogue to (41) if one assumes that an upper bound R is known a priori on the norm of a global minimizer x of $g(x)$, in which case (33) is equal to

$$\min g(x) \text{ subject to } R - \sum_{i=1}^n x_i^2 \ge 0.$$

Using a result of Putinar (cf. Theorem 15 below), it follows that, for any $\epsilon > 0$, $g(x) - p^* + \epsilon$ can be decomposed as $p(x) + q(x)\left(R - \sum_i x_i^2\right)$ for some polynomials $p(x)$ and $q(x)$ that are SOS. Testing for the existence of such decomposition can be expressed as a SDP program analogue to (41). Details are given in Section 7.3 where the general problem of minimizing a polynomial function over a semi-algebraic set is considered. Section 7.2 contains preliminaries over moment sequences and polynomials.

7.2 The moment problem and sums of squares of polynomials

The moment problem. Let $(S, +)$ be a commutative semigroup and let S^* denote the set of nonzero mappings $f : S \longrightarrow \mathbb{R}$ that are multiplicative, i.e., $f(\alpha+\beta) = f(\alpha)f(\beta)$ for all $\alpha, \beta \in S$. Given a sequence $y = (y_\alpha)_{\alpha\in S}$ indexed by S, its *moment matrix* $M(y)$ is the square matrix indexed by S whose (α,β)-th entry is $y_{\alpha+\beta}$ for $\alpha,\beta\in S$.

When S is the semigroup $\mathcal{P}(V)$ with the union as semigroup operation, we find the moment matrix $M_V(y)$ already introduced earlier in (5). When S is the semigroup $(\mathbb{Z}_+^n, +)$, we use the notation $M^{\mathbb{Z}}(y)$ for the moment matrix of $y \in \mathbb{R}^{\mathbb{Z}_+^n}$ and $M_t^{\mathbb{Z}}(y)$ for its principal submatrix indexed by all sequences $\alpha\in\mathbb{Z}_+^n$ with $|\alpha| \le t$ (considered above).

Following [BCJ79, BCR84], a sequence $y \in \mathbb{R}^S$ is said to be *positive semidefinite* if every finite principal submatrix of its moment matrix $M(y)$ is positive semidefinite and, given a subset $F \subseteq S^*$, y

is called a F-*moment sequence* if there exists a positive Radon measure μ on S^* supported by F such that

$$y_\alpha = \int_{S^*} f_\alpha d\mu(f) \text{ for all } \alpha \in S. \tag{42}$$

Given two sequences $x, y \in \mathbb{R}^S$, definition (6) extends as

$$(x * y)_\alpha := \sum_{\gamma \in S} x_\gamma y_{\alpha+\gamma} \quad \text{for } \alpha \in S.$$

The *moment problem* is the problem of characterizing moment sequences. It has been much studied in the literature especially for the semigroup $S = \mathbb{Z}_+^n$, in which case $S^* = \mathbb{R}^n$ and the moment condition (42) reads as relation (35); see [Fu83, BCR84] for a survey.

Obviously, every F-moment sequence should be positive semidefinite. Much research has been done for characterizing moment sequences for various closed sets F. For instance, for $n = 1$ and $F = \mathbb{R}$, every positive semidefinite sequence is a moment sequence, a result of Hamburger in 1920. For $n = 1$ and $F = \mathbb{R}_+$, a sequence $y = (y_i)_{i\geq 0}$ is a F-moment sequence if and only if both y and $e_1 * y = (y_{i+1})_{i\geq 0}$ are positive semidefinite, a result shown by Stieltjes in 1894. When F is a compact semi-algebraic set in $\mathbb{R}^n$, i.e.,

$$F = \{x \in \mathbb{R}^n \mid g_\ell(x) \geq 0 \text{ for } \ell = 1, \ldots, m\} \tag{43}$$

where g_ℓ are polynomials, Schmüdgen [Sc91] shows that y is a F-moment sequence if and only if y and $g * y$ are positive semidefinite for any product $g = g_{i_1} \ldots g_{i_k}$ of distinct polynomials among g_ℓ ($\ell = 1, \ldots, m$).

Reformulating Corollary 3 as a moment result in a semigroup. In fact, the result from Corollary 3 can also be viewed as a result about moments, if we consider sequences indexed by the semigroup $S := \mathcal{P}(V)$ (with the union as semigroup operation). Then, $S^* = \{\zeta^S \mid S \in \mathcal{P}(V)\}$. Hence, a sequence $y \in \mathbb{R}^{\mathcal{P}(V)}$ is a moment sequence if and only if $y \in \mathcal{C}_{\mathcal{P}(V)}$ which, by Corollary 3, is equivalent to y being a positive definite sequence. (Noting that $\mathcal{P}(V)$ is an idempotent semigroup, the result from Corollary 3 in the unconstrained case when $\mathcal{J} = \mathcal{P}(V)$ also follows from Proposition 4.17 in [BCR84].)

Let $F = \{x \in \{0,1\}^n \mid g_\ell(x) \geq 0 \ \forall l = 1, \ldots, m\}$, where the g_ℓ's are polynomials in which each variable occurs with degree ≤ 1, and let $\mathcal{J}$ be defined as in (13). Then, y is a F-moment sequence (meaning that the measure μ is nonzero only at ζ^S with $\chi^S \in F$, i.e., $S \in \mathcal{J}$) if and only if $y \in \mathcal{C}_\mathcal{J}$ which, by Corollary 3, is equivalent to the sequences y and $g_\ell * y$ ($\ell = 1, \ldots, m$) being positive semidefinite. Therefore, this gives a 'discrete' analogue of the above mentioned result of Schmüdgen.

Representations of nonnegative polynomials as sums of squares. Let $\mathcal{P}_+(F)$ denote the set of polynomials $p(x) = \sum_\alpha p_\alpha x^\alpha$ that are nonnegative on F; that is, $p(x) \geq 0$ for all $x \in F$. One of the basic results about moments, due to Haviland (1935), is that, given a closed subset F in $\mathbb{R}^n$, $y = (y_\alpha)_{\alpha\in\mathbb{Z}_+^n}$ is a F-moment sequence if and only if $y^T p \geq 0$ for any $p = (p_\alpha)_{\alpha\in\mathbb{Z}_+^n}$ in $\mathcal{P}_+(F)$.

Since a linear functional f on the set $\mathbb{R}[x_1, \ldots, x_n]$ of polynomials is completely determined by the sequence $(f(x^\alpha))_{\alpha\in\mathbb{Z}_+^n}$, the above result says that the set of F-moment sequences can be identified with the set of linear functionals that are nonnegative on $\mathcal{P}_+(F)$.

Let Σ^2 denote the convex cone generated by all squares of polynomials in $\mathbb{R}[x_1, \ldots, x_n]$. One can easily verify that a linear functional f on $\mathbb{R}[x_1, \ldots, x_n]$ is nonnegative on Σ^2 if and only if the sequence

$(f(x^\alpha))_{\alpha\in\mathbb{Z}_+^n}$ is positive semidefinite. The obvious inclusion

$$\Sigma^2 \subseteq \mathcal{P}_+(F)$$

corresponds by duality to the fact that every F-moment sequence is positive semidefinite. For $n = 1$, $F = \mathbb{R}$, it is well known that every nonnegative polynomial on $\mathbb{R}$ can be represented as the sum of squares of two polynomials, which gives again the result of Hamburger. For $n \geq 2$, not every nongegative polynomial can be expressed as a sum of squares of polynomials; this problem of representing polynomials as sums of squares goes back to Hilbert's 17th problem.

Let us reformulate the result of Schmüdgen in terms of polynomials. Let F be as in (43) and let $\Sigma^2(g_1,\ldots,g_m) := \sum_{I\subseteq[1,m]} (\prod_{i\in I} g_i)\Sigma^2$ denote the set of all polynomials of the form $\sum_{I\subseteq[1,m]} p_I \cdot \prod_{i\in I} g_i$, where all p_I belong to Σ^2. One can easily verify that a linear functional f on $\mathbb{R}[x_1,\ldots,x_n]$ is nonnegative on $\Sigma^2(g_1,\ldots,g_m)$ if and only if the associated sequence $y := (f(x^\alpha))_\alpha$ is positive semidefinite as well as the sequences $(\prod_{i\in I} g_i) * y$ for all $I \subseteq [1,m]$. Thus what Schmüdgen shows is that both sets $\mathcal{P}_+(F)$ and $\Sigma^2(g_1,\ldots,g_m)$ have the same sets of nonnegative linear functionals. From this follows that every polynomial p which is *positive* on F belongs to $\Sigma^2(g_1,\ldots,g_m)$. Putinar [Pu93] shows the following stronger result.

Theorem 15. *Let F be a compact semi-algebraic set as in* (43)*. Assume that there exists a polynomial $u \in \Sigma^2 + g_1\Sigma^2 + \ldots + g_m\Sigma^2$ for which the set $\{x \in \mathbb{R}^n \mid u(x) \geq 0\}$ is compact. If p is a polynomial positive on F, then $p \in \Sigma^2 + g_1\Sigma^2 + \ldots g_m\Sigma^2$.*

As we see below, this result plays a central role for evaluating asymptotically polynomial programs.

7.3 Lasserre's lift and project method for polynomial programs

Successive relaxations for polynomial programs. Let F be a semi-algebraic set as in (43). Assume that the assumptions from Theorem 15 hold: F is compact and the set $\{x \in \mathbb{R}^n \mid u(x) \geq 0\}$ is compact for some polynomial $u \in \Sigma_1^2(g) := \Sigma^2 + g_1\Sigma^2 + \ldots g_m\Sigma^2$. Suppose we want to solve the problem

$$p^* := \min \ g_0(x) \text{ subject to } x \in F \tag{44}$$

where g_0 is a polynomial of degree w_0 which can be assumed to satisfy $g_0(0) = 0$. Let w_ℓ denote the degree of g_ℓ and $v_\ell := \lceil \frac{w_\ell}{2} \rceil$, $v := \max_{\ell=1,\ldots,m} v_\ell$.

Lasserre [Las01a] constructs successive relaxations for problem (44) that converge asymptotically to its optimum solution. His construction is based on the following observation: For $x \in \mathbb{R}^n$, let $y^x \in \mathbb{R}^{\mathbb{Z}_+^n}$ with α-th entry x^α. Then, $M^{\mathbb{Z}}(y^x) = y^x(y^x)^T \succeq 0$ and $M^{\mathbb{Z}}(g * y^x) = g(x) \cdot y^x(y^x)^T \succeq 0$ if $g(x) \geq 0$. This leads to the following semidefinite relaxation of problem (44) for any $t \geq \max(v_0 - 1, v - 1)$:

$$\begin{array}{rll} p_t^* := & \min & \sum_\alpha (g_0)_\alpha y_\alpha \\ & \text{s.t.} & M_{t+1}^{\mathbb{Z}}(y) \succeq 0 \\ & & M_{t-v_\ell+1}^{\mathbb{Z}}(g_\ell * y) \succeq 0 \ (\ell = 1,\ldots,m) \\ & & y_0 = 1 \end{array} \tag{45}$$

The dual SDP program of (45) reads:

$$\begin{array}{rll} \rho_t^* := & \max & -X(0,0) - \sum_{\ell=1}^m g_\ell(0) Z_\ell(0,0) \\ & \text{s.t.} & \langle X, B_\gamma \rangle + \sum_{\ell=1}^m \langle Z_\ell, C_\gamma^\ell \rangle = (g_0)_\gamma \ (\gamma \neq 0) \\ & & X, Z_\ell \succeq 0 \ (\ell = 1, \ldots, m), \end{array} \tag{46}$$

where $M_{t+1}^{\mathbb{Z}}(y) = \sum_\gamma y_\gamma B_\gamma$ (as in (37), with $d = t+1$) and $M_{t-v_\ell+1}^{\mathbb{Z}}(g_\ell * y) = \sum_\gamma y_\gamma C_\gamma^\ell$, with $C_\gamma^\ell = \sum_{\substack{\alpha,\beta \in \mathcal{S}_{t-v_\ell+1}, \delta \\ \alpha+\beta+\delta=\gamma}} (g_\ell)_\delta E_{\alpha,\beta}$. We have:

$$\rho_t^* \leq p_t^* \leq p^*.$$

For $x \in F$, the sequence y^x is obviously an F-moment sequence (of the Dirac measure at x) and, thus, the primal program (45) states necessary conditions for y to be a moment sequence. The dual program (46) is related to representations of positive polynomials on F. Namely, if X, Z_ℓ are feasible for (46) with objective value ρ, then one can verify that the polynomial $g_0(x) - \rho$ belongs to the set $\Sigma_1^2(g) = \Sigma^2 + \sum_{\ell=1}^m g_\ell \Sigma^2$. For this, write

$$X = \sum_{j=1}^{r_0} q_j q_j^T, \ Z_\ell = \sum_{j=1}^{r_\ell} q_{\ell j} q_{\ell j}^T$$

for some vectors $q_j, q_{\ell j}$. Then, the polynomial $g_0(x) - \rho$ is equal to

$$\begin{aligned} &\sum_{\gamma \neq 0} (g_0)_\gamma x^\gamma + X(0,0) + \sum_\ell g_\ell(0) Z_\ell(0,0) = \langle X, \sum_\gamma x^\gamma B_\gamma \rangle + \sum_\ell \langle Z_\ell, \sum_\gamma x^\gamma C_\gamma^\ell \rangle \\ &= \langle X, M_{t+1}^{\mathbb{Z}}(y^x) \rangle + \sum_\ell \langle Z_\ell, M_{t-v_\ell+1}^{\mathbb{Z}}(g_\ell * y^x) \rangle = \sum_{j=1}^{r_0} (q_j(x))^2 + \sum_\ell g_\ell(x) \cdot \left(\sum_{j=1}^{r_\ell} (q_{\ell j}(x))^2 \right), \end{aligned}$$

using the facts that $\langle X, M_{t+1}^{\mathbb{Z}}(y^x) \rangle = \sum_j q_j^T M^{\mathbb{Z}}(y^x) q_j = \sum_j \sum_{\alpha,\beta} q_j(\alpha) q_j(\beta) x^{\alpha+\beta} = \sum_j (q_j(x))^2$ and $\langle Z_\ell, M_{t-v_\ell+1}^{\mathbb{Z}}(g_\ell * y^x) \rangle = \sum_j g_\ell(x)(q_{\ell j}(x))^2$. In particular, the polynomial $g_0(x) - \rho_t^*$ belongs to $\Sigma_1^2(g)$.

Conversely, given any $\epsilon > 0$, the polynomial $g_0(x) - p^* + \epsilon$ is positive on F which, by Theorem 15, implies that it belongs to $\Sigma_1^2(g)$. The above arguments can be reversed to construct from a decomposition of $g_0(x) - p^* + \epsilon$ in $\Sigma_1^2(g)$ a feasible solution X, Z_ℓ to (46) for some t with objective value $p^* - \epsilon$, which shows that $\rho_t^* \geq p^* - \epsilon$.

Therefore, for any $\epsilon > 0$, there exists t for which $p^* - \epsilon \leq \rho_t^* \leq p_t^* \leq p^*$. This shows that $\lim_{t \longrightarrow \infty} p_t^* = p^*$ and

$$p^* = \rho_t^* \text{ for some } t \Longleftrightarrow g_0(x) - p^* \in \Sigma^2 + \sum_{\ell=1}^m g_\ell \Sigma^2.$$

Moreover,

$$\operatorname{conv}(F) = \bigcap_{t \geq v-1} \mathcal{Q}_t(F)$$

where $\mathcal{Q}_t(F)$ is defined as the projection of the feasible set of the program (45) intersected with the hyperlane $y_0 = 1$, on the subspace $\mathbb{R}^n$ indexed by the sequences $\alpha \in \mathbb{Z}_+^n$ with $|\alpha| = 1$.

Relation with the previously defined Lasserre relaxations for $0-1$ programs. Consider now the case when F is the set of $0-1$ solutions of a polynomial system; that is,

$$F = \{x \in \mathbb{R}^n \mid g_\ell(x) \geq 0 \ (\ell = 1, \ldots, m), \ h_i(x) = 0 \ (i = 1, \ldots, n)\} \tag{47}$$

setting $h_i(x) := x_i - x_i^2$ for $i = 1, \ldots, n$. Then, one can assume without loss of generality that each g_ℓ has degree at most 1 in every variable and the assumptions from Theorem 15 hold (with $u(x) := \sum_{i=1}^n h_i(x)$). Using a result of Curto and Fialkow [CF00] about rank extensions of moment matrices, Lasserre [Las01b] shows finite convergence of the successive relaxations $\mathcal{Q}_t(F)$ to $\text{conv}(F)$; namely,

$$\mathcal{Q}_{n+v-1}(F) = \text{conv}(F). \tag{48}$$

The set

$$K := \{x \in [0,1]^n \mid g_\ell(x) \geq 0 \ (\ell = 1, \ldots, m)\} \tag{49}$$

is a natural relaxation of F. As we see in Proposition 16 below, the relaxation $\mathcal{Q}_t(F)$ coincides with the relaxation $Q_t(K)$ introduced earlier in Section 3.3. Proposition 16 shows in fact the following results: Our presentation in Section 3.3 of the Lasserre relaxations in terms of moment matrices indexed by subsets is equivalent to the original definition of Lasserre (in terms of moment matrices indexed by integer sequences); as an application, this gives an elementary proof for the convergence result from relation (48).

Proposition 16. *Let F and K be defined by* (47), (49) *respectively. Then, $\mathcal{Q}_t(F) = Q_t(K)$ for any $t \geq v-1$ and $\mathcal{Q}_t(F) = \mathcal{Q}_{n+v-1}(F)$ for any $t \geq n+v-1$.*

Proof. For $\alpha \in \mathbb{Z}_+^n$, define $\overline{\alpha} \in \{0,1\}^n$ by $\overline{\alpha}_i := 1$ if and only if $\alpha_i \geq 2$. Then, the condition $M_t^{\mathbb{Z}}(h_i * y) = 0$ means that

$$y_\alpha = y_{\overline{\alpha}} \tag{50}$$

for any α with $|\alpha| \leq 2t$. From this follows that the α-th column of the moment matrix $M^{\mathbb{Z}}(y)$ is identical to its $\overline{\alpha}$-th column; similarly for the matrices $M^{\mathbb{Z}}(g_\ell * y)$. A first consequence is that, for $t \geq n$,

$$M_t^{\mathbb{Z}}(y) \succeq 0 \Longleftrightarrow M_n^{\mathbb{Z}}(y) \succeq 0, \quad \text{and} \quad M_t^{\mathbb{Z}}(g_\ell * y) \succeq 0 \Longleftrightarrow M_n^{\mathbb{Z}}(g_\ell * y) \succeq 0;$$

this shows equality $\mathcal{Q}_t(F) = \mathcal{Q}_{n+v-1}(F)$ for $t \geq n+v-1$. Let $z \in \mathbb{R}^{\mathcal{P}(V)}$ with I-th entry $z_I := y_\alpha$ with $\alpha_i = 1$ if $i \in I$ and $\alpha_i = 0$ otherwise. Then, $M_t(z)$ is a principal submatrix of $M_t^{\mathbb{Z}}(y)$ and another consequence of (50) is that

$$M_t(z) \succeq 0 \Longleftrightarrow M_t^{\mathbb{Z}}(y) \succeq 0;$$

similarly, $M_t(g_\ell * z) \succeq 0 \Longleftrightarrow M_t^{\mathbb{Z}}(g_\ell * y) \succeq 0$. This shows equality $Q_t(K) = \mathcal{Q}_t(F)$ for $t \geq v-1$. ∎

The quadratic case. Consider finally the case when F is a semi-algebraic set defined by a set of quadratic constraints; that is, each g_ℓ is of the form $g_\ell(x) = x^T Q_\ell x + 2q_\ell^T x + \gamma_\ell$ (Q_ℓ symmetric $n \times n$ matrix, $q_\ell \in \mathbb{R}^n$, $\gamma_\ell \in \mathbb{R}$). For $\ell = 1, \ldots, m$, set $P_\ell := \begin{pmatrix} \gamma_\ell & q_\ell^T \\ q_\ell & Q_\ell \end{pmatrix}$. Then, $g_\ell(x) = \langle P_\ell, \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} \rangle$. Therefore, the following set $\hat{F}$ is a natural semidefinite relaxation of F:

$$\hat{F} := \{x \in \mathbb{R}^n \mid (1 \ x) = Ye_0 \text{ for some } Y \succeq 0 \text{ with } \langle P_\ell, Y \rangle \geq 0 \text{ for } \ell = 1, \ldots, m\} \tag{51}$$

(considered, e.g., in [FK97]). In fact, the set $\hat{F}$ coincides with the first Lasserre relaxation $\mathcal{Q}_0(F)$.

Proposition 17. *$\mathcal{Q}_0(F) = \hat{F}$.*

Proof. By definition, $x \in \mathbb{R}^n$ belongs to $\mathcal{Q}_0(F)$ if there exists $y = (y_\alpha)_{|\alpha|\leq 2}$ satisfying $y_0 = 1$, $y_{e_i} = x_i$ ($i = 1,\ldots,n$) ($e_1,\ldots,e_n$ denoting the standard unit vectors in $\mathbb{R}^n$), $M_1^{\mathbb{Z}}(y) \succeq 0$ and $g_\ell * y(0) \geq 0$ ($\ell = 1,\ldots,m$). The equality $\mathcal{Q}_0(F) = \hat{F}$ follows from the following fact: Given a symmetric matrix $Y = (Y_{ij})_{i,j=0}^n$, define $y = (y_\alpha)_{|\alpha|\leq 2}$ by $y_0 := Y_{00}$, $y_{e_i} := Y_{0i}$, $y_{e_i+e_j} := Y_{ij}$ ($i,j = 1,\ldots,n$); then, $M_1(y) = Y$ and $g_\ell * y(0) = \langle P_\ell, Y\rangle$. ∎

References

[AW01] M.F. Anjos and H. Wolkowicz. Strenghtened semidefinite relaxations via a second lifting for the max-cut problem. To appear in the Special Issue of Discrete Applied Mathematics devoted to Foundations of Heuristics in Combinatorial Optimization, 2001.

[BCC93] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming*, 58:295–324, 1993.

[BM86] F. Barahona and A.R. Mahjoub. On the cut polytope. *Mathematical Programming*, 36:157–173, 1986.

[BCJ79] C. Berg, J.P.R. Christensen and C.U. Jensen. A remark on the multidimensional moment problem. *Mathematische Annalen*, 243:163–169, 1979.

[BCR84] C. Berg, J.P.R. Christensen and P. Ressel. *Harmonic Analysis and Semigroups - Theory of Positive Definite and Related Functions.* Springer-Verlag, New York, 1984.

[C73] V. Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete Mathematics*, 4:305–337, 1973.

[CCH89] V. Chvátal, W. Cook, and M. Hartman. On cutting-plane proofs in combinatorial optimization. *Linear Algebra and its Applications*, 114/115:455–499, 1989.

[CD01] W. Cook and S. Dash. On the matrix-cut rank of polyhedra. *Mathematics of Operations Research*, 26:19–30, 2001.

[CL01] G. Cornuéjols and Y. Li. On the rank of mixed 0-1 polyhedra. In K. Aardal and A.M.H. Gerards, eds., IPCO 2001, *Lecture Notes in Computer Science*, 2081:71–77, 2001.

[CF00] R.E. Curto and L.A. Fialkow. The truncated complex K-moment problem. *Transactions of the American Mathematical Society*, 352:2825–2855, 2000.

[DL97] M. Deza and M. Laurent. *Geometry of Cuts and Metrics.* Springer Verlag, Berlin, 1997.

[E99] F. Eisenbrand. On the membership problem for the elementary closure of a polyhedron. *Combinatorica*, 19:299-300, 1999.

[ES99] F. Eisenbrand and A.S. Schulz. Bounds on the Chvátal rank of polytopes in the 0/1 cube. In G. Cornuéjols, R.E. Burkard and G.J. Woeginger, eds., IPCO 1999, *Lecture Notes in Computer Science*, 1610:137–150, 1999.

[Fu83] B. Fuglede. The multidimensional moment problem. *Expositiones Mathematicae*, 1:47–65, 1983.

[FK97] T. Fujie and M. Kojima. Semidefinite programming relaxation for nonconvex quadratic programs. *Journal of Global Optimization*, 10:367–380, 1997.

[GT00] M.X. Goemans and L. Tunçel. When does the positive semidefiniteness constraint help in lifting procedures. *Mathematics of Operations Research*, to appear.

[GW95] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cuts and satisfiability problems using semidefinite programming. *Journal of the Association for Computing Machinery*, 42:1115–1145, 1995.

[GLS88] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization.* Springer Verlag, Berlin, New York, 1988.

[Ha35] E.K. Haviland. On the momentum problem for distributions in more than one dimension. I. II. *American Journal of Mathematics*, 57:562–568, 1935 and 58:164–168, 1936.

[KP01] E. de Klerk and D.V. Pasechnik. Approximating the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, to appear.

[La00] M. Laurent. Tighter linear and semidefinite relaxations for max-cut based on the Lovász-Schrijver lift-and-project procedure. *SIAM Journal on Optimization*, to appear.

[Las00] J.B. Lasserre. Optimality conditions and LMI relaxations for $0-1$ programs. Technical Report N. 00099, 2000.

[Las01a] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11:796–817, 2001.

[Las01b] J.B. Lasserre. An explicit exact SDP relaxation for nonlinear $0-1$ programs. In K. Aardal and A.M.H. Gerards, eds., *Lecture Notes in Computer Science*, 2081:293–303, 2001.

[Lo79] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, IT-25:1–7, 1979.

[LS91] L. Lovász and A. Schrijver. Cones of matrices and set-functions and $0-1$ optimization. *SIAM Journal on Optimization*, 1:166–190, 1991.

[Ne00] Y. Nesterov. Squared functional systems and optimization problems. In J.B.G. Frenk, C. Roos, T. Terlaky, and S. Zhang, eds., *High Performance Optimization*, pages 405–440. Kluwer Academic Publishers, 2000.

[Pa00] P.A. Parillo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, May 2000.

[Pu93] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42:969–984, 1993.

[Sc91] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289:203–206, 1991.

[SA90] 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*, 3:411–430, 1990.

[Sh87] N.Z. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Kibernetika*, 5:102–106, 1987.

[Sh98] N.Z. Shor. *Nondifferentiable Optimization and Polynomial Problems.* Kluwer Academic Publishers, 1998.

[ST99] T. Stephen and L. Tunçel. On a representation of the matching polytope via semidefinite liftings. *Mathematics of Operations Research*, 24:1–7, 1999.

[Wi68] H.S. Wilf. Hadamard determinants, Möbius functions, and the chromatic number of a graph. *Bulletin of the American Mathmatical Society*, 74:960–964, 1968.