SEMIDEFINITE PROGRAMMING APPROXIMATIONS FOR STABLE SETS, COLOURING, AND CUTS IN GRAPHS

MONIQUE LAURENT

Abstract. In the recent years semidefinite programming has become a widely used tool for designing better efficient approximation algorithms for hard combinatorial optimization problems. In this note we review some constructions for nested semidefinite programming relaxations to 0/1 linear programming problems and some of their main properties. In particular we indicate how to get more compact semidefinite programs, thus more amenable to practical computation, using symmetry reduction. We illustrate these techniques on some basic combinatorial problems dealing with stable sets, cuts and colouring in graphs.

1. Introduction

Combinatorial optimization deals typically with the problem of finding an optimum object within a finite (but huge) collection of objects. Think for instance of time-tabling, or scheduling problems, or various problems on graphs, like (1) finding a shortest path between two vertices, or (2) finding a shortest tour traversing all edges, or (3) finding a shortest tour traversing all vertices, or (4) finding a maximum set of pairwise non-adjacent vertices, or (5) colouring the vertices with as few colours as possible in such a way that adjacent vertices receive distinct colours, or (6) partitioning the vertices into two classes so as to maximize the number of edges cut by the partition, etc. As is well known, while efficient (polynomial time) algorithms exist for the first two graph problems (1)-(2), the last four problems (3)-(6) are NP-hard.

A typical approach to attack such hard problems is to represent the objects over which one wishes to optimize by 0/1 vectors and to formulate the problem at hand as a 0/1 linear programming (LP) problem, of the form

$$\max\ c^T x \ \text{ s.t. } Ax \leq b \ \text{ and } x \in \{0,1\}^n, \tag{1}$$

where A is an $m \times n$ matrix, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. Without the integrality constraint on the variable x, (1) would be an LP problem, thus solvable in polynomial time. However, adding the integrality condition makes the problem NP-hard. With P denoting the convex hull of the feasible solutions to (1), the objective is then to find a relaxation of the polytope P which is efficient and tight, meaning that one can compute over it in polynomial time and that the returned solution is close enough to the optimum over P.

Linear programming relaxations of P have been traditionnally considered. However, they are sometimes not powerful enough and semidefinite programming, an extension of linear programming where vector variables are replaced by matrix variables constrained to be positive semidefinite, turns out to be a powerful technique for designing better and efficient approximation algorithms for some problem classes. Semidefinite programming is also widely used in other areas like system and control theory (cf. e.g. [4]), but we focus here on its application to combinatorial optimization.

We now recall a few basic facts about semidefinite programs and refer to [23, 47, 49] and references therein for details. Let $\mathcal{S}_n$ denote the set of $n \times n$ symmetric matrices. A matrix $X \in \mathcal{S}_n$ is positive semidefinite, written as $X \succeq 0$, if $u^T X u \geq 0$ for all $u \in \mathbb{R}^n$ or, equivalently, if there exist vectors $v_1, \dots, v_n \in \mathbb{R}^n$ for which $X = v_i^T v_j$ ($\forall i, j$). The standard form of a semidefinite program (SDP) reads

$$\sup \langle C, X \rangle \quad \text{s.t. } \langle A_j, X \rangle = b_j \ (j = 1, \dots, m) \quad \text{and } X \succeq 0, \tag{2}$$

where $C, A_1, \dots, A_m \in \mathcal{S}_n$, $b \in \mathbb{R}^m$ are given, $X \in \mathcal{S}_n$ is the matrix variable, and $\langle C, X \rangle = \text{Tr}(C^T X)$ is the standard inner product. The program (2) is a convex program and, when C, A_j are diagonal matrices, it reduces to a classic linear program. As one can test in polynomial time whether a given rational matrix is positive semidefinite (e.g. using Gaussian elimination), semidefinite programs can be solved in polynomial time to any fixed precision using the ellpsoid method (cf. [15]). Algorithms based on the ellipsoid method are however not practical since their running time is prohibitively high. Instead, interior-point algorithms are widely used in practice; they return an approximate solution (to any given precision) in polynomially many iterations and their running time is efficient in practice for medium size problems. As most currently available interior-point algorithms are designed to take advantage of block-diagonal matrices, it is computationnally easier to solve an SDP involving many small blocks rather than one large matrix. One way to get such block-diagonal SDP's is by exploiting the symmetry structure of the problem at hand (cf. Section 4); moreover one can also design the SDP relaxation to force such a symmetry structure (cf. Section 3.3).

Contents. We present in Section 2 some basic SDP bounds for stable sets, colouring and maximum cut problems, and discuss their properties. Then we see in Section 3 how to construct hierarchies of SDP relaxations leading to the full representation of the combinatorial problem at hand. In Section 4 we indicate how symmetry reduction can be used to get more compact SDP programs, and Section 5 contains a few pointers to further related results.

2. Basic semidefinite bounds for stable sets, coloring and maximum cuts

We have chosen to illustrate the use of semidefinite programming in combinatorial optimization on the following combinatorial problems: maximum stable sets, minimum graph coloring, and maximum cuts in graphs. For these problems, some milestone results have been obtained in the recent years, that spurred an intense research activity and a wealth of results for other combinatorial optimization problems; we refer e.g. to [12, 23, 34, 38] and references therein for a detailed exposition. In this section we introduce some basic SDP relaxations which have led to interesting results for our three problems.

2.1. Basic SDP bounds for stable sets and colouring via Lovász' theta number.

Let $G = (V, E)$ be a graph; V is the set of vertices (or nodes) (typically $V = [1, n]$) and E is the set of edges, which are pairs of distinct vertices, then said to be *adjacent* in G. A *stable* (or *independent*) set in G is a set of vertices that are pairwise non-adjacent, and a *(vertex) colouring* of G is an assignment of colours to the vertices in such a way that adjacent vertices receive distinct colours, in other words, a partition of V into stable sets. Then the stable set problem asks for the *stability number* $\alpha(G)$, defined as the maximum cardinality of a stable set in G, while the graph colouring problem asks for the *chromatic number* $\chi(G)$, defined as the minimum number of colours in a vertex colouring. Both problems are NP-hard [11]. Note that

$$\chi(G) \geq \omega(G), \tag{3}$$

where $\omega(G)$ is the largest cardinality of a *clique* (i.e., a set of pairwise adjacent vertices) in G. Obviously, $\omega(G) = \alpha(\bar{G})$, where $\bar{G}$ is the complement of G, with the same set V of vertices and two distinct vertices being adjacent in $\bar{G}$ precisely when they are not adjacent in G.

For some graphs the inequality (3) is strict. For instance, it is strict for any circuit C_n of odd length $n \geq 5$, as $\omega(C_n) = 2 < \chi(C_n) = 3$, and for the complement of C_n as well. However there are many interesting classes of graphs for which equality $\omega(G) = \chi(G)$ holds. This is the case e.g. for bipartite graphs, line graphs of bipartite graphs, comparability graphs, chordal graphs, and for their complements as well. In fact the class of graphs for which equality $\omega(G) = \chi(G)$ holds not only for G but also for all its induced subgraphs (i.e. all those graphs that can be obtained by deleting vertices in G) turns out to be very interesting; following Berge, graphs in this class are called *perfect graphs*. Thus C_n and its complement $\bar{C}_n$ are not perfect for odd $n \geq 5$. Berge conjectured in 1962 that a graph is perfect if and only if its complement is perfect, which was proved a few years later by Lovász [35]. Berge also conjectured that a graph is perfect if and only if it does not contain any odd circuit of length at least 5, or its complement, as an induced subgraph; this was proved only recently by Chudnovsky et al. [6] and is known as the *strong perfect graph theorem*. It is intriguing to determine the complexity of computing $\alpha(G)$ and $\chi(G)$ for perfect graphs. As we indicate below this can be done in polynomial time, but to show this one has to use semidefinite programming.

Lovász [37] introduced his celebrated *theta number* $\vartheta(G)$, which serves as bound for both $\alpha(G)$ and $\chi(G)$. The theta number is defined via the semidefinite program

$$\vartheta(G) := \max \ \mathrm{Tr}(JX) \ \text{ s.t. } \ \mathrm{Tr}(X) = 1, \ X_{ij} = 0 \ (ij \in E), \ X \succeq 0, \tag{4}$$

where J denotes the all-ones matrix. A basic property of the theta number is that it satisfies the so-called *sandwich inequality*

$$\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G}), \quad \text{or equivalently,} \quad \omega(G) \leq \vartheta(\bar{G}) \leq \chi(G). \tag{5}$$

Indeed, if $\mathbf{1}_S \in \{0,1\}^V$ is the incidence vector of a stable set S in G, then $X := \mathbf{1}_S(\mathbf{1}_S)^T/|S|$ is feasible for (4) with objective value $|S|$, which gives $\alpha(G) \leq \vartheta(G)$. On the other hand, if X is a feasible solution to (4) and $V = C_1 \cup \ldots \cup C_k$ is a partition into $k := \chi(\bar{G})$ cliques of G, then

$$0 \leq \sum_{h=1}^{k} (k\mathbf{1}_{C_h} - \mathbf{1}_V)^T X (k\mathbf{1}_{C_h} - \mathbf{1}_V) = k^2 \mathrm{Tr}(X) - k\mathbf{1}_V^T X \mathbf{1}_V = k(k - \mathrm{Tr}(JX)),$$

which implies $\mathrm{Tr}(JX) \leq k$ and thus $\vartheta(G) \leq \chi(\bar{G})$. Another basic property of the theta number is

$$\vartheta(G)\bar{\vartheta}(G) \geq |V|, \quad \text{with equality if } G \text{ is vertex-transitive.} \tag{6}$$

Here, $\bar{\vartheta}$ is the 'complementary' graph parameter defined by $\bar{\vartheta}(G) := \vartheta(\bar{G})$ for any graph G. This convention is also used later for other graph parameters.

For perfect graphs, equality holds throughout in (5), which implies $\alpha(G) = \vartheta(G)$ and $\chi(G) = \vartheta(\bar{G})$. As the theta number can be computed in polynomial to any fixed precision, the stability number and the chromatic number can be computed in polynomial time for perfect graphs. Moreover, a maximum stable set and a minimum coloring can also be computed in polynomial time for a perfect graph G (by iterated computations of the theta number of certain induced subgraphs of G). These computations thus rely on using semidefinite programming and as of today no alternative efficient algorithm is known.

Lovász' original motivation for introducing the theta number was to bound the *Shannon capacity* of a graph G, which is defined as

$$\Theta(G) := \lim_{k \to \infty} \alpha(G^k)^{\frac{1}{k}}.$$

Here G^k denotes the product of k copies of G, with vertex set V^k and with two disctinct vertices $(u_1, \ldots, u_k)$ and $(v_1, \ldots, v_k)$ being adjacent in G^k if $u_h = v_h$ or $u_h v_h \in E$ for each position $h = 1, \ldots, k$. If we view V as an alphabet and adjacent vertices $u, v \in V$ as letters that can be confounded, then $\alpha(G^k)$ is the maximum number of words of length k that cannot be confounded, since for any two of them there is a position h where their hth letters are not confoundable. One can verify that $\alpha(G^k) \geq \alpha(G)^k$ and $\vartheta(G^k) \leq \vartheta(G)^k$, which implies

$$\alpha(G) \leq \Theta(G) \leq \vartheta(G).$$

Therefore, when G is perfect, $\Theta(G) = \vartheta(G)$ can thus be computed via semidefinite programming. Lovász could also compute the Shannon capacity of the circuit C_5 using the theta number. He showed that $\Theta(C_5) = \sqrt{5}$, which follows from $\alpha(C_5^2) \geq 5$ (easy to verify) and $\vartheta(C_5) = \sqrt{5}$; the latter follows using (6) since C_5 is vertex-transitive and isomorphic to its complement. The exact value of the Shannon capacity of C_n is not known for odd $n \geq 7$. (Cf. [27] for more details.)

Although it permits to compute the stability and chromatic numbers of perfect graphs, the theta number can be a poor approximation for general graphs. Indeed the gap between $\alpha(G)$ and $\vartheta(G)$ (or $\vartheta(G)$ and $\chi(\bar{G})$) can be as large as $|V|^{1-\epsilon}$ (for any $\epsilon > 0$) [9]. Nevertheless much work has been done on approximating $\alpha(G)$ and $\chi(G)$ with SDP. For example, it is shown in [21] how to colour in randomized polynomial time a k-colourable graph with $O(n^{1-\frac{3}{k+1}}\sqrt{\log n})$ colours [21] using variations of the SDP (4) combined with sophisticated rounding techniques, and the behaviour of the theta number for Erdös-Renyi graphs is analysed in [7].

2.2. **The Goemans-Williamson approximation algorithm for maximum cuts.** Another successful application of semidefinite programming to combinatorial optimization is the celebrated 0.878-approximation algorithm of Goemans and Williamson [14] for the max-cut problem, which we sketch below.

Given a graph $G = (V, E)$ ($|V| = n$) with edge weights $w \in \mathbb{R}_+^E$, the *cut* determined by $S \subseteq V$ is the set of edges having their endnodes in distinct classes of the partition $(S, V \setminus S)$. The *max-cut problem* asks for the maximum weight of a cut in G, denoted as $\mathrm{mc}(G)$. While a minimum weight nonempty cut can be found in polynomial time (using flow algorithms), the max-cut problem is NP-hard [11].

Erdös proposed in 1967 the following simple 1/2-approximation algorithm. Colour the vertices $v_1, \ldots, v_n$ of G with two colours blue and red as follows: First colour v_1 blue. If $v_1, \ldots, v_i$ are coloured, colour v_{i+1} blue if the total weight of the edges joining v_{i+1} to the red vertices in $\{v_1, \ldots, v_i\}$ is more than the total weight of the edges joining v_{i+1} to the blue vertices in this set, and colour v_{i+1} red otherwise. Then the cut given by this partition of V has weight at least $w(E)/2$ and thus at least $\mathrm{mc}(G)/2$. There is an even easier randomized 1/2-approximation algorithm. Namely colour randomly each node blue or red independently, with probability 1/2. The probability that an edge belongs to the cut determined by this partition is 1/2 and thus the expected weight of this cut is $w(E)/2$. Can one do better than 1/2?

It is unfortunately not known how to do better using LP relaxations. Consider, for instance, the classic LP relaxation for max-cut obtained from the triangle inequalities: $x_{ij} - x_{ik} - x_{jk} \leq 0$ and $x_{ij} + x_{ik} + x_{jk} \leq 2$ for $i, j, k \in V$. While this LP relaxation permits to solve max-cut exactly for graphs with no K_5-minor (including planar graphs) [3], it is however a weak relaxation for general graphs, as the integrality gap can be as large as $2 - \epsilon$ ($\forall\ \epsilon > 0$) [41].

However, an improved approximation algorithm was designed by Goemans and Williamson [14] using semidefinite programming. For this, it is convenient to model the max-cut problem using ± 1-valued variables as

$$\mathrm{mc}(G) = \max \sum_{ij \in E} w_{ij}(1 - x_i x_j)/2 \quad \text{s.t.} \quad x \in \{\pm 1\}^n. \tag{7}$$

Observe that the matrix $X := xx^T$ satisfies the constraints: $X \succeq 0$, $X_{ii} = 1$ $(i \in V)$, and $\text{rank}(X) = 1$. If we omit the rank condition, then we find the semidefinite relaxation

$$\text{sdp}(G) := \max \sum_{ij \in E} w_{ij}(1 - X_{ij})/2 \ \text{ s.t. } \ X \succeq 0,\ X_{ii} = 1\ (i \in V), \tag{8}$$

which can be solved in polynomial time (to any fixed precision). Let X be an optimum solution to (8). Goemans and Williamson [14] propose the following randomized rounding procedure for constructing a good cut from X. Say X is the Gram matrix of $v_1, \dots, v_n \in \mathbb{R}^n$, i.e. $X_{ij} = v_i^T v_j$ $\forall i, j \in V$. Consider a random unit vector $r \in \mathbb{R}^n$. The hyperplane with normal r splits V into two sets, depending on the sign of $r^T v_i$ $(i \in V)$, giving $x \in \{\pm 1\}^n$ with $x_i = 1$ iff $r^T v_i \geq 0$. As the probability that an edge ij lies in the cut determined by this partition is equal to $\frac{1}{\pi} \arccos(v_i^T v_j)$, the expected weight of this cut is equal to

$$\sum_{ij \in E} w_{ij} \frac{\arccos(v_i^T v_j)}{\pi} = \sum_{ij \in E} w_{ij} \frac{1 - v_i^T v_j}{2} \, \frac{2}{\pi} \frac{\arccos(v_i^T v_j)}{1 - v_i^T v_j} \geq \alpha_{\text{GW}}\ \text{sdp}(G) \geq\ 0.878567\ \text{mc}(G),$$

after setting $\alpha_{\text{GW}} := \min_{0 < \vartheta \leq \pi} \frac{2}{\pi} \frac{\vartheta}{1 - \cos\vartheta}$ and observing that $\alpha_{\text{GW}} >\ 0.878567$. This implies $\text{mc}(G) \geq \alpha_{\text{GW}}\ \text{sdp}(G)$. This randomized algorithm can be derandomized to yield in polynomial time a deterministic cut achieving the same performance ratio.

Much research has been done trying to improve the Goemans-Williamson approximation algorithm for max-cut and to extend and apply it to other problems. Although improved algorithms could be designed for special graph classes, no better approximation ratio could yet be shown for the general max-cut problem. It has in fact been shown that α_{GW} is the best possible approximation ratio for max-cut that can be achieved in polynomial time (if P $\neq$ NP) under the Unique Games Conjecture [22]. On the negative side, Håstad [20] proved that no polynomial time approximation algorithm exists for max-cut with performance guarantee better than $16/17 \sim 0.94117$ (if P $\neq$ NP). We now indicate a natural extension to ± 1-quadratic programming:

$$\text{m}(A) := \max_{x \in \{\pm 1\}^n} x^T A x, \tag{9}$$

and its SDP relaxation:

$$\text{sdp}(A) := \max_{u_1, \dots, u_n \in S^{n-1}} \sum_{i,j=1}^{n} A_{ij} u_i^T u_j = \max \text{Tr}(AX) \ \text{ s.t. } \ X \succeq 0,\ X_{ii} = 1\ (i \in V), \tag{10}$$

where $A \in \mathbb{R}^{n \times n}$ and S^{n-1} is the unit sphere in $\mathbb{R}^n$. When $A = L_w$, the Laplacian matrix of (G, w) (with $(L_w)_{ii} := \sum_{j | ij \in E} w_{ij}$, $(L_w)_{ij} := -w_{ij}$ if $ij \in E$ and 0 otherwise), we find the programs (7) and (8), in which case $\text{m}(L_w)/\text{sdp}(L_w) \geq 0.8785$. For general $A \succeq 0$, the above analysis can be adapted to show the inequality: $\text{m}(A) \geq \frac{2}{\pi}\ \text{sdp}(A)$ [40]. The constant of Grothendieck is a fundamental tool in functional analysis, which corresponds to the integrality gap of the following analogues of the programs (9)-(10):

$$\text{m}_G(A) := \max_{x \in \{\pm 1\}^n,\ y \in \{\pm 1\}^m} y^T A x = \sum_{i=1}^{n} \sum_{j=1}^{m} A_{ij} x_i y_j,$$

$$\text{sdp}_G(A) := \max_{u_1, \dots, u_n, v_1, \dots, v_m \in S^{n+m-1}} \sum_{i=1}^{n} \sum_{j=1}^{m} A_{ij} u_i^T v_j,$$

where $A \in \mathbb{R}^{m \times n}$. Grothendieck (1953) showed the existence of a smallest constant K_G satisfying $\text{sdp}_G(A) \leq K_G \text{m}_G(A)$ for all $A \in \mathbb{R}^{m \times n}$. It is known that $K_G \geq 1.68$ [42] and $K_G \leq \frac{\pi}{2 \ln(1 + \sqrt{2})} \sim$

1.782 (cf. [1]). This has been used e.g. for deriving an efficient approximation algorithm for the cut-norm of matrices [1], and an extension to quadratic forms on graphs can be found in [2].

3. Hierarchies of semidefinite programming relaxations

We saw above how to define in a natural way a semidefinite relaxation for the maximum stable set problem (via the SDP (4)) and for the max-cut problem (via the SDP (8)). Several procedures have been proposed for constructing nested (increasingly tight) SDP relaxations (cf. [28, 39, 45], also [29, 34] and references therein). We briefly sketch some of these constructions.

3.1. **The N_+-operator of Lovász-Schrijver.** Lovász and Schrijver [39] introduced the N_+-operator which, given a convex set $K \subseteq [0,1]^V$, produces the convex set $N_+(K)$ consisting of the vectors $x \in \mathbb{R}^V$ for which

$$(11) \qquad \exists X \text{ s.t. } x = \mathrm{diag}(X), \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0, \quad \frac{Xe_j}{x_j} \in K \text{ if } x_j \neq 0, \quad \frac{x - Xe_j}{1 - x_j} \in K \text{ if } x_j \neq 1.$$

Let K_I denote the convex hull of the integer points in K, thus the polytope we are interested in. We have:

$$K_I \subseteq N_+(K) \subseteq K.$$

Indeed, if x is an integer point in K, then $X := xx^T$ satisfies the conditions in (11), which shows the left inclusion, and the right inclusion follows easily after noting that $Xe_j = 0$ if $x_j = 0$ and $Xe_j = x$ if $x_j = 1$, and x, X are as in (11).

Defining iteratively $N_+^t(K) := N_+(N_+^{t-1}(K))$ for $t \geq 2$, $N_+^1(K) = N_+(K)$, we obtain a hierarchy of SDP relaxations of K_I, which finds K_I in $n = |V|$ iterations, i.e. $N_+^n(K) = K_I$. An important algorithmic property is that, for any *fixed* t, if one can optimize in polynomial time over K, then the same holds for $N_+^t(K)$.

As in [39], consider for instance the case when $K = \mathrm{FR}(G)$, the fractional stable set polytope of G, defined by the constraints $0 \leq x_i \leq 1$ $(i \in V)$ and $x_i + x_j \leq 1$ $(ij \in E)$. Then K_I is the convex hull of the incidence vectors of all stable sets in G. In fact, $K_I = N_+^t(K)$ for any $t \geq \alpha(G)$. Moreover, for $t = 1$, $N_+(K)$ gives an SDP bound for $\alpha(G)$ which is at least as strong as the theta number, and $N_+(K) = K_I$ when G is a perfect graph. As a clique and a stable set share at most one node, the clique inequalities $\sum_{i \in C} x_i \leq 1$ (C clique in G) are valid for K_I. They are in fact also valid for $N_+(K)$. However, if we omit positive semidefiniteness in (11) (leading to the N-operator), then $t = \omega(G) - 2$ iterations of the N-operator are needed before $N^t(K)$ satisfies all clique inequalities. Thus adding the semidefinite condition definitely helps!

3.2. **A simple recipe for SDP relaxations.** Here is a simple construction for SDP relaxations, leading to refinements of the hierarchy of Lovász and Schrijver. The basic idea is to 'lift' a vector $x \in \{0,1\}^V$ to a higher dimensional vector $y = (x_I := \prod_{i \in I} x_i)_{I \in \mathcal{A}}$ containing the products of the x_i's indexed by a given subset $\mathcal{A}$ of the full power set $\mathcal{P}(V)$. Then the matrix $Y = yy^T$ obviously satisfies the constraints: (i) $Y \succeq 0$, (ii) $Y_{\emptyset,\emptyset} = 1$, and (iii) the (I,J)-entry $Y_{I,J}$ depends only on the union $I \cup J$ (as $Y_{I,J} = x_{I \cup J}$). A matrix indexed by $\mathcal{A}$ and satisfying (iii) is of the form $M_{\mathcal{A}}(z) := (z_{I \cup J})_{I,J \in \mathcal{A}}$ for some $z \in \mathbb{R}^{\mathcal{A} \cup \mathcal{A}}$ (where $\mathcal{A} \cup \mathcal{A} = \{I \cup J \mid I, J \in \mathcal{A}\}$), and is known as a *(combinatorial) moment matrix*. Moreover, if x satisfies a linear constraint $\sum_{i \in V} a_i x_i \leq a_0$, then the lifted vector y satisfies $(a_0 - \sum_{i \in V} a_i x_i) yy^T \succeq 0$, yielding the *localizing constraint* $M_{\mathcal{A}}(az) \succeq 0$, after setting $(az)_I := a_0 z_I - \sum_{i \in V} a_i z_{I \cup \{i\}}$ for all I.

The next lemma implies that, when imposing the 'full' SDP constraints $M_{\mathcal{P}(V)}(z) \succeq 0$ combined with the 'full' localizing constraints, we obtain an exact (SDP or LP) representation of the original 0/1 problem.

Lemma 1. *The following assertions are equivalent for $z \in \mathbb{R}^{\mathcal{P}(V)}$.*

(i) $M_{\mathcal{P}(V)}(z) \succeq 0$.

(ii) $\sum_{S'|S\subseteq S'\subseteq V} (-1)^{|S'\setminus S|} y_{S'} \geq 0$ *for all* $S \subseteq V$.

(iii) $z = \sum_{S\subseteq V} \lambda_S z_S$ *for some* $\lambda_S \geq 0$, *after setting* $z_S := (\prod_{i\in I}(\mathbf{1}_S)_i)_{I\subseteq V} \in \{0,1\}^{\mathcal{P}(V)}$.

If we choose $\mathcal{A} = \mathcal{P}_t(V)$, the collection of all subsets of V with size at most t, then we obtain the SDP hierarchy of Lasserre [28]. If we choose $\mathcal{A} = \mathcal{P}(U)$ for all $U \in \mathcal{P}_t(V)$, then we obtain the hierarchy of Sherali-Adams [45] (thus an LP hierarchy, in view of Lemma 1). In fact, the Lasserre hierarchy refines both the Lovász-Schrijver and the Sherali-Adams hierarchies [29].

Applying this to the fractional stable set polytope $K = \mathrm{FR}(G)$, the Lasserre relaxation of order t gives the upper bound

$$\mathrm{las}^{(t)}(G) := \max_{z\in\mathbb{R}^{\mathcal{P}_{2t}(V)}} \sum_{i\in V} z_i \ \text{ s.t. } z_\emptyset = 1,\ M_{\mathcal{P}_t}(z) \succeq 0,\ z_{ij} = 0\ (ij \in E) \tag{12}$$

for the stability number $\alpha(G)$, while the analogous program

$$\psi^{(t)}_{\mathrm{las}}(G) := \min_{z\in\mathbb{R}^{\mathcal{P}_{2t}(V)}} z_\emptyset \ \text{ s.t. } M_{\mathcal{P}_t(V)}(z) \succeq 0,\ z_i = 1\ (i \in V),\ z_{ij} = 0\ (ij \in E) \tag{13}$$

yields a lower bound for the chromatic number $\chi(G)$. In fact, $\psi^{(t)}_{\mathrm{las}}(G)$ remains bounded above by the *fractional chromatic number*:

$$\chi_f(G) := \min \sum_{S \text{ stable}} \lambda_S \ \text{ s.t. } \sum_{S \text{ stable}} \lambda_S \mathbf{1}_S = \mathbf{1}_V,\ \lambda_S \geq 0. \tag{14}$$

That is, $\psi^{(t)}_{\mathrm{las}} \leq \chi_f \leq \chi$ for all t. Moreover, the pair $(\mathrm{las}^{(t)}, \psi^{(t)}_{\mathrm{las}})$ is equal to $(\vartheta, \bar{\vartheta})$ for $t = 1$, to (α, χ_f) for $t \geq \alpha$, and it satisfies the analogue of the reciprocity relation (6) for any $t \geq 1$ [16, 17].

This construction also applies to max-cut (and to (9)). However, as variable x is now ± 1-valued (instead of $0,1$-valued), one should now require that the (I,J)-th entry in a moment matrix $M_{\mathcal{A}}(z)$ depends only on the symmetric difference $I\Delta J$ (instead of the union $I \cup J$). At order $t = 1$ we find the SDP relaxation (2) (or (10)); geometric properties of these relaxations can be found in [31].

3.3. **A more economical block-diagonal hierarchy.** The SDP (12) involves a matrix of size $O(n^{2t})$, which is thus too costly to compute for large n or t. More economical block-diagonal variations of this hierarchy are proposed in [19], based on the following idea: Instead of the full matrix $M_{\mathcal{P}_t(V)}(z)$, consider several smaller blocks arising from principal submatrices of it. Namely, for any $T \subseteq V$ of size $t-1$, consider the matrices $M(T;z)$ indexed by $\bigcup_{S\subseteq T} \mathcal{A}_S$, where $\mathcal{A}_S := \{S, S\cup\{i\} \mid i \in V\}$. These matrices need only the components of z indexed by $\mathcal{P}_{t+1}(V)$ (as opposed to $\mathcal{P}_{2t}(V)$ in (12)) and each $M(T;z)$ has a very symmetric block-structure. For instance, for $T = \{1,2\}$,

$$M(T;z) = \begin{pmatrix} A_\emptyset(z) & A_1(z) & A_2(z) & A_{12}(z) \\ A_1(z) & A_1(z) & A_{12}(z) & A_{12}(z) \\ A_2(z) & A_{12}(z) & A_2(z) & A_{12}(z) \\ A_{12}(z) & A_{12}(z) & A_{12}(z) & A_{12}(z) \end{pmatrix} \succeq 0 \iff \begin{cases} A_\emptyset(z) - A_1(z) - A_2(z) + A_{12}(z) \succeq 0 \\ A_1(z) - A_{12}(z) \succeq 0 \\ A_2(z) - A_{12}(z) \succeq 0 \\ A_{12}(z) \succeq 0 \end{cases}$$

(where blocks are indexed by $\mathcal{A}_\emptyset, \mathcal{A}_{\{1\}}, \mathcal{A}_{\{2\}}, \mathcal{A}_{\{1,2\}}$). A simple adaptation of Lemma 1 implies that each $M(T;z)$ can be block-diagonalized, namely

$$M(T;z) \succeq 0 \Longleftrightarrow \sum_{S' | S \subseteq S' \subseteq T} (-1)^{|S'\setminus S|} A_{S'}(z) \succeq 0 \quad \forall S \subseteq T. \tag{15}$$

In this way we obtain a *block-diagonal hierarchy*, where the condition '$M_{\mathcal{P}_t(V)}(z)$' is replaced by '$M(T;z) \succeq 0$ for all $T \subseteq V$ with $|T| = t-1$'; the localizing conditions can be analogously replaced by block-diagonal conditions. This block-diagonal hierarchy can be seen as an explicit analogue of the Lovász-Schrijver hierarchy, which is simpler (not recursive) and more economical (as it uses less variables and constraints); cf. [19] for details.

Let $\ell^{(t)}$ denote the block-diagonal variation of the parameter $\mathrm{las}^{(t)}$ from (12), thus defined by

$$\ell^{(t)}(G) = \max \sum_{i\in V} z_i \text{ s.t. } z_\emptyset = 1,\ M(T;z) \succeq 0\ (T \subseteq V,\ |T| = t-1),\ z_{ij} = 0\ (ij \in E). \tag{16}$$

Then, $\ell^{(1)} = \vartheta$, and there is still convergence in $\alpha(G)$ steps to the stability number. Analogously, let $\psi^{(t)}$ denote the block-diagonal variation of $\psi^{(t)}_{\mathrm{las}}$ in (13). In view of (15), $\ell^{(t)}$ (or $\psi^{(t)}$) can be computed via an SDP with $\binom{n}{t-1}2^{t-1}$ matrices of order $n+1$ and $O(n^{t+1})$ variables, as compared to one matrix of order $O(n^t)$ and $O(n^{2t})$ variables for the parameter (12).

The parameter $\ell^{(t)}$ ($t = 2, 3$) has been tested in [19] on the class of Paley graphs P_q; P_q has vertex set the field $\mathbb{F}_q$ on q elements ($q = 1 \mod 4$), with an edge ij iff $i - j$ is a square in $\mathbb{F}_q$. Here are some numerical values: For $q = 401$, $\vartheta \sim 20.02$, $\ell^{(2)} \sim 12.7$, $\ell^{(3)} \sim 10$, $\alpha = 9$; for $q = 809$, $\vartheta \sim 28.4$, $\ell^{(2)} \sim 17.3$, $\ell^{(3)} \sim 13.4$, $\alpha = 11$ (the computation involves symmetry reduction, cf. [16, 19] for details). The bound $\ell^{(2)}$ was also computed for the class of Hamming graphs, enabling to get better upper bounds for the coding problem; details are given in Section 4.

3.4. **SDP hierarchies converging to the chromatic number.** We saw above some construction of SDP bounds for the chromatic number, however these bounds remain below the fractional chromatic number. We now give a 'recipe' for constructing bounds that may go beyond the fractional chromatic number. Following [16, 17], consider the operator Ψ that maps a graph parameter β to the graph parameter Ψ_β defined by

$$\Psi_\beta(G) := \min_{t\in\mathbb{N}} t \text{ s.t. } \beta(G\square K_t) = |V(G)|,$$

where $G_t := G\square K_t$ is the Cartesian product of G and the complete graph K_t. Then, Ψ is monotone non-increasing on the interval $[\frac{|V|}{\chi}, \bar\chi]$ and maps it onto the interval $[\omega, \chi]$. In particular, $\Psi_\alpha = \chi$, which is a well known reduction of graph colouring to the stable set problem. Moreover, Ψ maps the whole interval $[\frac{|V|}{\chi}, \alpha]$ onto the single parameter χ and the whole interval $[\frac{|V|}{\omega}, \bar\chi]$ onto ω; as an application, no parameter in these intervals can be computed in polynomial time unless P=NP.

An important property of Ψ is that it maps any hierarchy converging to α to a hierarchy converging to the *chromatic number* χ. While the block-diagonal hierarchy $\psi^{(t)}$ remains below χ_f, we have $\Psi_{\ell^{(t)}}(G) = \chi(G)$ for $t \geq |V(G)|$. Moreover,

$$\Psi_{\ell^{(t)}} \geq \psi^{(t)} \geq \frac{|V|}{\ell^{(t)}}.$$

We have tested the parameter $\Psi_{\ell^{(t)}}$ on Kneser graphs. For these graphs it is known that the fractional chromatic number is close to the clique number and thus any bound for the chromatic number which remains below the fractional chromatic number is of no use.

The Kneser graph $K(n,r)$ has vertex set the set of words of weight r in $\{0,1\}^n$, with two words adjacent if their Hamming distance is $2r$. It is known that $\alpha = \vartheta = \binom{n-1}{r-1}$, $\omega = \lfloor n/r \rfloor$, $\chi_f = n/r$, and $\chi = n - 2r + 2$ [36]. But numerical tests confirm that the parameter $\Psi_{\ell^{(2)}}$ indeed goes beyond the fractional chromatic number (cf. [16, 18]). Here are some numerical values: For $K(12, 3)$, $\chi_f = 4 < \Psi_{\ell^{(2)}} = 5 < \Psi_{\ell_+^{(2)}} = 6$; for $K(36, 6)$, $\chi_f = 6 < \Psi_{\ell^{(2)}} = 7 < \Psi_{\ell_+^{(2)}} = 9$. (Here $\ell_+^{(2)}$ is the strengthening of $\ell^{(2)}$ obtained by adding non-negativity constraints on the variables.) See the next section for some comments on the computation of $\Psi_{\ell^{(2)}}$ for Kneser graphs.

4. Symmetric reduction and some computational results

We briefly indicate here how symmetry can be used to block-diagonalize SDP's and get more compact programs. (See e.g. [24, 46] for details.) Consider the SDP (2) where the matrices C, A_j are indexed by a set V, $|V| = n$. Assume that (2) is invariant under action of a group $\mathcal{G}$ of permutations of V. That is, if X is feasible for (2) and $\sigma \in \mathcal{G}$, then the matrix $\sigma(X) := (X_{\sigma(i),\sigma(j)})_{i,j\in V}$ is again feasible for (2). As (2) is a convex program, we may restrict w.l.o.g. the matrix variable X to lie in the algebra

$$\mathcal{A}_{\mathcal{G}} := \{X \mid \sigma(X) = X \ \forall \sigma \in \mathcal{G}\}$$

of invariant matrices under action of $\mathcal{G}$. Then, the number of distinct entries in X is at most the dimension d of $\mathcal{A}_{\mathcal{G}}$, which is given by the number of distinct orbits of $V \times V$ under action of $\mathcal{G}$. Thus there is already a gain in the number of variables when the dimension is small. Moreover, one can also reduce the size of the matrices in the SDP. As explained in [25], one can indeed replace the SDP by an equivalent SDP involving $d \times d$ matrices. One can do even better. Indeed the algebra $\mathcal{A}_{\mathcal{G}}$ is a matrix $*$-algebra, which means that it is closed under addition, multiplication, scalar multiplication, and transposition. Hence, by Wedderburn's theorem, it can be *block-diagonalized.* That is, there exists a unitary matrix U and $s, n_1, k_1, \ldots, n_s, k_s \in \mathbb{N}$ such that the set $U^*\mathcal{A}_{\mathcal{G}}U = \{U^*XU \mid X \in \mathcal{A}_{\mathcal{G}}\}$ consists of all the block-diagonal matrices with blocks $A_1 \in \mathbb{C}^{n_1\times n_1}$ (repeated k_1 times), ..., $A_s \in \mathbb{C}^{n_s\times n_s}$ (repeated k_s times), so that $d = \sum_{i=1}^s n_i^2$. In this way one can replace the SDP (2) involving one $n \times n$ matrix X by an equivalent SDP involving s smaller matrices $A_1, \ldots, A_s$. This reduction is particularly interesting when the n_i's are small compared to n.

Let us indicate how this can be applied to compute the block-diagonal bound $\ell^{(2)}$ from (16) for Hamming graphs. Given $\mathcal{D} \subseteq [1, n]$, the Hamming graph $G = H(n, \mathcal{D})$ has vertex set $\{0, 1\}^n$, with an edge between two words when their Hamming distance lies in $\mathcal{D}$. Computing the stability number of Hamming graphs is a basic question in coding theory. Hamming graphs have a rich automorphism group $\mathcal{G}$, arising from all permutations of the n coordinates combined with 'switchings' (replacing all words i by $i \oplus i_0$, for a given word i_0).

Consider first the theta number of $H(n, \mathcal{D})$, given by the SDP (4). Then the matrix X is indexed by $\{0, 1\}^n$ and is invariant under action of $\mathcal{G}$ precisely when the entries $X_{i,j}$ depend only on the Hamming distance between i, j. The corresponding algebra $\mathcal{A}_{\mathcal{G}}$ is known as the Bose-Mesner algebra, which is commutative with dimension $n + 1$. Thus it can be diagonalized, and the theta number of $H(n, \mathcal{D})$ can in fact be computed via an LP of size n, which gives the bound of Delsarte [8].

Consider now the second order bound $\ell^{(2)}(H(n, \mathcal{D}))$ in the block-diagonal hierarchy (16). Note first that in (16) it suffices to consider just *one* of the matrices $M(\{i_0\}; z)$ (for example for the zero word $i_0 = \mathbf{0}$) because the graph $H(n, \mathcal{D})$ is vertex-transitive. Next note that the matrix $M(\{i_0\}; z)$ has the block form

$$\begin{pmatrix} 1 & a^T & b^T \\ a & A & B \\ b & B & B \end{pmatrix} \qquad (17)$$

(blocks being indexed by $\emptyset$, $\{0,1\}^n$, and $\{\{\mathbf{0}, i\} \mid i \in \{0,1\}^n\}$). Here, A can be assumed to be invariant under action of $\mathcal{G}$ and thus lies in the Bose-Mesner algebra. Matrix B can be assumed to be invariant under action of the subgroup $\mathcal{G}_{\mathbf{0}}$ of $\mathcal{G}$ consisting of its elements fixing $\mathbf{0}$, i.e. $B_{i,j}$ depends only on the weights of i, j and $i \oplus j$. The corresponding algebra $\mathcal{A}_{\mathcal{G}_{\mathbf{0}}}$ is known as the Terwilliger algebra of the Hamming scheme, whose dimension is equal to $O(n^3)$. Schrijver [44] has computed the explicit block-diagonalization for the Terwilliger algebra. Thus the condition $B \succeq 0$ (of size 2^n) can be replaced by requiring that several blocks be positive semidefinite, of sizes $n_1, \ldots, n_s$ with $\sum_i n_i^2 = O(n^3)$. Thanks to this symmetry reduction, one is able to compute the parameter $\ell^{(2)}(H(n, \mathcal{D}))$ for n up to 28, which gives upper bounds on the stability number of $H(n, \mathcal{D})$, improving the bound of Delsarte (cf. numerical results in [26, 32, 44]).

The computation of the parameter $\Psi_{\ell^{(2)}}$ for Kneser graphs $G = K(n, r)$ deserves a comment. In order to compute $\Psi_{\ell^{(2)}}(G)$, we need to compute $\ell^{(2)}(G \square K_t)$ for several values of t. Again it suffices to consider in (16) *one* matrix $M(\{u\}; z)$ for just one node u of $G \square K_t$. This matrix has again the shape (17), but A and B are now indexed by $V(G \square K_t)$, of cardinality $t|V|$. However, one can exploit the invariance under the permutations of K_t to replace the condition $M(\{u\}; z) \succeq 0$ by a new SDP condition involving four matrices, of sizes $2|V(G)|+1, 2|V(G)|, |V(G)|, |V(G)|$. Thus, the parameter t which determines the size of $M(\{u\}; z)$ comes now as a numerical parameter within these four matrices. Next one can exploit the invariance under the automorphism group of the Kneser graph to further reduce these four matrices (this involves again the Terwilliger algebra). We refer to [16, 18] for details on the symmetry reduction and for computational results.

5. Final remarks

We have sketched how to use semidefinite programming for designing hierarchies of convex relaxations for 0/1 linear programming problems. Note that an analogous construction applies more generally to polynomial optimization problems, where the objective is a polynomial function and the constraints are polynomial equations or inequalities; cf. e.g. the survey [33] for details.

There are many interesting aspects that have been investigated about these hierarchies of SDP relaxations. A first example is the notion of *rank* of a hierarchy, defined as the smallest number of iterations needed to find the original 0/1 polytope. For instance, the rank of the N_+-operator has been studied e.g. for the max-cut, stable set and traveling salesman problems. For instance, the rank of the N_+ operator for the stable set polytope of G is at most $\alpha(G)$ and it can be equal to $\alpha(G)$ (cf. [13]). As another example, it is shown in [30] that the rank of the Lasserre hierarchy for the cut polytope of K_n is at least $\lceil n/2 \rceil$, and equality is conjectured.

An intriguing question is how to use higher order relaxations to get better approximation algorithms. This is by no way easy. As mentioned earlier, for max-cut, one cannot hope to do better than with the order 1 relaxation if the unique game conjecture holds. The following negative result is shown in [10] dealing with LP hierarchies for max-cut: The integrality gap of the basic LP relaxation by triangle inqualities can be as large as $2 - \epsilon$ ($\forall \epsilon > 0$), and this is still true for the LP relaxation obtained after t iterations of the Sherali-Adams procedure, or after adding all valid inequalities for the cut polytope on at most t nodes, for any fixed t. Moreover, it is shown in [43] that after $\Omega_\epsilon(n)$ iterations of the Lovász-Schrijver N operator, one gets an LP relaxation with the same duality gap $2 - \epsilon$. If one recalls that the basic order 1 SDP relaxation for max-cut permits to achieve the Goemans-Williamson approximation ratio 1/0.878, then this is a good illustration of the strength of the positive semidefinite condition.

On the positive side, the authors of [5] are able to exploit higher order relaxations in the Lasserre hierarchy for the maximum independent set in a 3-uniform hypergraph. Namely, for fixed $\gamma > 0$, if there is an independent set of size at least γn, then they can construct one of size $n^{\Omega(\gamma^2)}$ using the relaxation of order $\Theta(1/\gamma^2)$.

One may also want to exploit the structure of the problem at hand, like sparsity, in order to get more economical relaxations. For instance, when G has small tree-width, one can design sparse relaxations leading to a compact (LP or SDP) representation of the problem (cf. [33, Sec. 8.2], [48]).

References

[1] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. *SIAM Journal on Computing* **35**:787–803, 2006.

[2] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic forms on graphs. *Inventiones Mathematicae* **163**(3):499-522, 2005.

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

[4] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* Volume 15 of Studies in Applied Mathematics, SIAM, 1994.

[5] E. Chlamtac and G. Singh. Improved approximation guarantees through higher levels of SDP hierarchies. APPROX 2008.

[6] M. Chudnovsky, N.Robertson, P.Seymour, and R.Thomas. The strong perfect graph theorem. *Annals of Mathematics* **164**:51–229, 2006.

[7] A. Coja-Oghlan. The Lovász number of random graphs. *Combinatorics, Probability and Computing* **14**:439–465, 2005.

[8] P. Delsarte. *An Algebraic Approach to the Association Schemes of Coding Theory.* [Philips Research Reports Supplements (1973) No. 10] Philips Research Laboratories, Eindhoven, 1973.

[9] U. Feige. Randomized graph products, chromatic numbers, and the Lovász ϑ-function. *Combinatorica* 17:79–90, 1997.

[10] W. Fernandez de la Vega and C. Kenyon-Mathieu. Linear programming relaxations of maxcut. SODA 2007, pages 53–61.

[11] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*, San Francisco, W.H. Freeman & Company, Publishers, 1979.

[12] M.X. Goemans. Semidefinite Programming in Combinatorial Optimization. *Mathematical Programming* **79**:143–161, 1997.

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

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

[15] M. Grötschel, L. Lovász and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, Springer, 1988.

[16] N. Gvozdenović. *Approximating the Stability Number and the Chromatic Number of a Graph via Semidefinite Programming.* Ph.D. thesis, Univ. Amsterdam, 2008.

[17] N. Gvozdenović and M. Laurent. The operator Ψ for the Chromatic Number of a Graph. *SIAM Journal on Optimization* **19**(2):572-591, 2008.

[18] N. Gvozdenović and M. Laurent. Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization. *SIAM Journal on Optimization* **19(2)**:592–615, 2008.

[19] N. Gvozdenović, M. Laurent and F. Vallentin. Block-diagonal semidefinite programming hierarchies for 0/1 programming. *Operations Research Letters* **37**:27–31, 2009.

[20] J. Håstad. Some optimal inapproximability results. In *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing*, ACM, New York, pp. 1–10, 1997.

[21] D. Karger, R. Motwani, and M. Sudan. Approximate graph colouring by semidefinite programming. *Journal of the Association for Computing Machinery* **45**:246–265, 1998.

[22] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM Symposium on the Theory of Computing*, ACM, New York, pp. 767–775, 2002.

[23] E. de Klerk. *Aspects of Semidefinite Programming - Interior Point Algorithms and Selected Applications.* Kluwer, 2002.

[24] E. de Klerk. Exploiting special structure in semidefinite programming: a survey of theory and applications. Preprint, 2008.Available at `http://www.optimization-online.org/DB_HTML/2008/0/2107.html`

[25] E. de Klerk, D. Pasechnik, A. Schrijver. Reduction of symmetric semidefinite programs using the regular *-representation. *Mathematical Programming*, **109**(2-3):613–624, 2007.

[26] E. De Klerk, D.V. Pasechnik. A note on the stability number of an orthogonality graph. *European Journal of Combinatorics* **28**:1971–1979, 2007.

[27] D.E. Knuth. The sanwich theorem. *Electronic Journal of Combinatorics* **1**:1–48, 1994.

[28] 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.

[29] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research* **28(3)**:470–496, 2003.

[30] M. Laurent. Lower bound for the number of iterations in semidefinite relaxations for the cut polytope. *Mathematics of Operations Research,* **28**(4):871–883, 2003.

[31] M. Laurent. Semidefinite relaxations for Max-Cut. In *The Sharpest Cut: The Impact of Manfred Padberg and His Work.* M. Grötschel (ed.), pages 257-290, *MPS-SIAM Series in Optimization* **4**, 2004.

[32] M. Laurent. Strengthened semidefinite programming bounds for codes. *Mathematical Programming* **109(2-3)**:239-261, 2007.

[33] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, Vol. **149** of *The IMA Volumes in Mathematics and its Applications,* M. Putinar and S. Sullivant (eds.), Springer, pages 157-270, 2009.

[34] M. Laurent and F. Rendl. Semidefinite Programming and Integer Programming. In *Handbook on Discrete Optimization*, K. Aardal, G. Nemhauser, R. Weismantel (eds.), pp. 393-514, Elsevier B.V., 2005.

[35] L. Lovász. Normal hypergraphs and the perfect graph conjecture. *Discrete Mathematics* **2**:253–267, 1972.

[36] L. Lovász. Kneser's conjecture, chromatic number and homotopy. *Journal of Combinatorial Theory, Series A* **25**(3):319–324, 1978.

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

[38] L. Lovász. Semidefinite programs and combinatorial optimization. In *Recent Advances in Algorithms and Combinatorics*, B.A. Reed and C.L. Sales (eds.), Springer, 2003.

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

[40] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software* **9**:141–160, 1998.

[41] S. Poljak and Zs. Tuza. On the expected relative error of the polyhedral approximation of the max-cut. *Operations Research Letters* **16**:191–198, 1994.

[42] J.A. Reeds. A new lower bound on the real Grothendieck constant, 1991. Avalaible at `http://www.dtc.umn.edu/~reedsj/bound2.dvi`

[43] G. Schoenebeck, L. Trevisan, and M. Tulsiani. Tight integrality gaps for Lovász-Schrijver LP relaxations of vertex cover and max cut. ECCC, Report 132, 2006.

[44] A. Schrijver. New code upper bounds from the Terwilliger algebra and semidefinite programming. *IEEE Transactions on Information Theory* **51**:2859–2866, 2005.

[45] 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.

[46] F. Vallentin. Symmetry in semidefinite programs. *Linear Algebra and its Applications* **430**:360–369, 2009.

[47] L. Vandenberghe and S. Boyd. Semidefinite Programming. *SIAM Review* **38(1)**:49–95, 1996.

[48] M.J. Wainwright and M.I. Jordan. Treewidth-based conditions for exactness of the Sherali-Adams and Lasserre relaxations. Univ. California, Berkeley, Technical Report 671, 2004.

[49] H. Wolkowicz, R. Saigal, L. Vandeberghe (eds.). *Handbook of Semidefinite Programming*, Kluwer, 2000.

Centrum voor Wiskunde en Informatica (CWI), Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
E-mail address: `monique@cwi.nl`