Operations Research Letters 37 (2009) 27–31

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Block-diagonal semidefinite programming hierarchies for 0/1 programming

Nebojša Gvozdenović [a], Monique Laurent [b,*], Frank Vallentin [b]

[a] The Faculty of Economics, University of Novi Sad, Segedinski put 9-11, 24000 Subotica, Serbia

[b] Centrum voor Wiskunde en Informatica (CWI), Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 19 December 2007
Accepted 30 October 2008
Available online 7 November 2008

Keywords:
0/1 linear programming
Semidefinite programming
Stable sets
Payley graphs

ABSTRACT

Lovász and Schrijver, and later Lasserre, proposed hierarchies of semidefinite programming relaxations for 0/1 linear programming problems. We revisit these two constructions and propose two new, block-diagonal hierarchies, which are at least as strong as the Lovász–Schrijver hierarchy, but less costly to compute. We report experimental results for the stable set problem of Paley graphs.

1. Introduction

A basic approach in combinatorial optimization consists of formulating the problem at hand as a 0/1 linear programming problem, typically of the form

$$\max \left\{ c^{\mathsf{T}} x \mid Ax \le b,\ x \in \{0,1\}^n \right\},$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$. Then the task is to find an efficiently computable outer approximation of the polytope P, defined as the convex hull of the 0/1 solutions to $Ax \le b$.

On the one hand, extensive research has been done for finding (partial) linear inequality descriptions for many polyhedra arising from specific combinatorial optimization problems. On the other hand, researchers focused on developing general purpose methods for arbitrary 0/1 linear programming problems. Here let us mention the method of Gomory for generating cuts strengthening the initial linear relaxation $\{x \in \mathbb{R}^n \mid Ax \le b\}$ of P and its various extensions for generating strong cutting planes (see e.g. [14,18]), the lift-and-project method [1], the reformulation-linearization technique [17], the matrix-cut method of Lovász and Schrijver [13], and the sums of squares and moment method of Lasserre [8]. Some of these methods are compared in [9]; see also [11]. A common feature of the methods of Lovász–Schrijver and of Lasserre is that they consider hierarchies involving semidefinite relaxations of P: Convex sets Q_t ($t = 1, \dots, n+1$) are constructed which can be described by semidefinite conditions and which form a hierarchy of increasingly stronger relaxations:

$$\{x \in \mathbb{R}^n \mid Ax \le b\} \supseteq Q_1 \supseteq Q_2 \supseteq \cdots \supseteq Q_{n+1} = P.$$

The two hierarchies are related; it is shown in [9] that the hierarchy of Lasserre refines the hierarchy of Lovász–Schrijver.

In this note we revisit these hierarchies and propose two new ones, which differ in the way of encoding the linear constraints defining the starting linear relaxation of P. Moreover one of them (introduced in Section 2.4) can also be defined when the starting relaxation of P is an arbitrary convex body, as is the case for the Lovász–Schrijver construction. The new hierarchies are nested between the Lasserre and Lovász–Schrijver hierarchies, but they are less costly to compute. So they are especially well suited for implementations. For example, at given order t, the new hierarchy from Section 2.4 involves $1/(t+1)!n^{t+1} + O(n^t)$ variables compared to $2^{t-2}n^{t+1} + O(n^t)$ variables for the Lovász–Schrijver hierarchy and to $O(n^{2t})$ variables for the Lasserre hierarchy. The new hierarchies can be seen as a variation of the Lasserre hierarchy, where one replaces a large matrix of order $O(n^t)$ by smaller blocks of order $n+1$ arising by block-diagonalizing suitably defined principal submatrices of the original large matrix. The motivation for considering block matrices is that it is computationally easier to solve a semidefinite program involving many small blocks rather than one large matrix. Most currently available interior-point algorithms for semidefinite programming are indeed designed to exploit block-diagonal matrices. While the hierarchy of Lovász and Schrijver is originally defined recursively, we give an explicit description obtained by "unfolding" the recursion. In this way, the connection to the new hierarchies becomes transparent (see Section 2 for details).

* Corresponding author.
E-mail addresses: nebojsa.gvozdenovic@gmail.com (N. Gvozdenović), monique@cwi.nl, M.Laurent@cwi.nl (M. Laurent), f.vallentin@cwi.nl (F. Vallentin).

0167-6377/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2008.10.003

When applied to the stable set problem, our new construction gives a block-diagonal hierarchy whose first two steps were already used in the literature. The first order relaxation gives the Lovász theta number and the second order one gives parameters considered in [4,10] for the stable set problem and in [6,7] for the coloring problem. In these applications the computational advantage of the new hierarchy was of crucial importance.

Contents of the paper

In Section 2 we first briefly introduce the constructions of Lovász–Schrijver and of Lasserre. Then we give the new construction and show how to derive more compact formulations by block-diagonalization. In Section 3 we apply it to the stable set problem and in Section 4 we present some computational results illustrating the behavior of the new hierarchy for approximating the stability number of Paley graphs.

Notation

Given a finite set V, we denote the collection of all subsets of V by $\mathcal{P}(V)$. Given a non-negative integer r, set $\mathcal{P}_r(V) := \{I \in \mathcal{P}(V) \mid |I| \le r\}$ and $\mathcal{P}_{=r}(V) := \{I \in \mathcal{P}(V) \mid |I| = r\}$. By $\mathbf{0}$ we denote the empty set. Sometimes we identify $\mathcal{P}_{=1}(V)$ with V, i.e., we write i instead of $\{i\}$. Furthermore, we sometimes write ij instead of $\{i, j\}$ and ijk instead of $\{i, j, k\}$, etc. The standard unit vectors in $\mathbb{R}^{\mathcal{P}_1(V)}$ are denoted by $e_{\mathbf{0}}, e_i$ for $i \in V$.

2. Semidefinite programming hierarchies

Suppose we are given a convex cone K contained in the homogenized unit cube $\{x \in \mathbb{R}^{\mathcal{P}_1(V)} \mid 0 \le x_i \le x_{\mathbf{0}}\ (i \in V)\}$. Set

$$P_K := \mathrm{conv}\left\{x \in \{0,1\}^V \mid \begin{pmatrix}1\\x\end{pmatrix} \in K\right\},$$

$$C_K := \mathbb{R}_+\left\{\begin{pmatrix}1\\x\end{pmatrix} \in K \text{ with } x \in \{0,1\}^V\right\}.$$

The general objective is to find the linear inequality description of the polytope P_K or, equivalently, of the cone C_K. In Section 2.1 we recall the construction of Lovász–Schrijver which applies to any convex cone K. While the original construction is recursive we propose an explicit semidefinite programming reformulation. In Section 2.2 we recall the construction of Lasserre which applies to the case when K is represented by polynomial inequalities. Here we focus on polyhedral cones K of the form

$$K = \{x \in \mathbb{R}^{\mathcal{P}_1(V)} \mid a_l^{\mathsf{T}} x \ge 0\ (l = 1, \dots, m)\}, \tag{1}$$

where $a_1, \dots, a_m \in \mathbb{R}^{\mathcal{P}_1(V)}$. In Section 2.3 we introduce our new construction, which can be seen as a variation of the previous methods. We discuss two new hierarchies. The first one applies to polyhedral cones K as in (1) and is more economical than the Lasserre hierarchy while still refining the Lovász–Schrijver hierarchy. The second one applies to any convex cone K and can be seen as a non-recursive analogue of the Lovász–Schrijver hierarchy having a more compact and explicit formulation.

2.1. The Lovász–Schrijver hierarchy

In this section we recall basic facts about the Lovász–Schrijver hierarchy. For proofs and more details we refer to [13]. Set

$$\mathcal{M}_{+,V} := \{Y \in \mathbb{R}^{\mathcal{P}_1(V)\times\mathcal{P}_1(V)} \mid Y \succeq 0,\ Y_{ii} = Y_{\mathbf{0}i}\ (i \in V)\},$$

where "$\succeq 0$" stands for "is positive semidefinite". For a convex cone $K \subseteq \mathbb{R}^{\mathcal{P}_1(V)}$ define

$$\mathcal{M}_+(K) := \{Y \in \mathcal{M}_{+,V} \mid Ye_i \in K,\ Y(e_{\mathbf{0}} - e_i) \in K\ (i \in V)\},$$
$$N_+(K) := \{Ye_{\mathbf{0}} \mid Y \in \mathcal{M}_+(K)\}.$$

The t-th *iterate of the Lovász–Schrijver hierarchy* is $N_+^t(K) := N_+(N_+^{t-1}(K))$ for $t \ge 1$, where $N_+^1(K) := N_+(K)$ and $N_+^0(K) := K$. It lies between K and C_K and $N_+^{t+1}(K) \subseteq N_+^t(K)$. We have $N_+^t(K) = C_K$ for $t = |V|$. Moreover, for any fixed t, if one can optimize over K in polynomial time then the same holds for $N_+^t(K)$.

In the following proposition we "unfold" the recursive definition of $N_+^t(K)$ and give an explicit semidefinite programming formulation. Its proof is straightforward and thus omitted.

Proposition 2.1. *A vector $x \in \mathbb{R}^{\mathcal{P}_1(V)}$ lies in $N_+^t(K)$ if and only if there exist a matrix $Y \in \mathcal{M}_{+,V}$ and matrices $Y_{i_1,\dots,i_s}^{\sigma_1,\dots,\sigma_s} \in \mathcal{M}_{+,V}$, with $s = 1, \dots, t-1$, $i_1, \dots, i_s \in V$ and $\sigma_1, \dots, \sigma_s \in \{\pm 1\}$, satisfying the following conditions:*

(a) $x = Ye_{\mathbf{0}}$.

(b) *For all* $s = 1, \dots, t-1$, $i_1, \dots, i_s \in V$, *and* $\sigma_1, \dots, \sigma_{s-1} \in \{\pm 1\}$:

$$Y_{i_1,\dots,i_{s-1}}^{\sigma_1,\dots,\sigma_{s-1}} e_{i_s} = Y_{i_1,\dots,i_s}^{\sigma_1,\dots,\sigma_{s-1},+1} e_{\mathbf{0}},$$
$$Y_{i_1,\dots,i_{s-1}}^{\sigma_1,\dots,\sigma_{s-1}} (e_{\mathbf{0}} - e_{i_s}) = Y_{i_1,\dots,i_s}^{\sigma_1,\dots,\sigma_{s-1},-1} e_{\mathbf{0}},$$

where $Y_{i_1,\dots,i_{s-1}}^{\sigma_1,\dots,\sigma_{s-1}} = Y$ *for* $s = 1$.

(c) *For all* $i_1, \dots, i_t \in V$ *and* $\sigma_1, \dots, \sigma_{t-1} \in \{\pm 1\}$:

$$Y_{i_1,\dots,i_{t-1}}^{\sigma_1,\dots,\sigma_{t-1}} e_{i_t} \in K,$$
$$Y_{i_1,\dots,i_{t-1}}^{\sigma_1,\dots,\sigma_{t-1}} (e_{\mathbf{0}} - e_{i_t}) \in K,$$

where $Y_{i_1,\dots,i_{t-1}}^{\sigma_1,\dots,\sigma_{t-1}} = Y$ *for* $t = 1$.

The above formulation allows estimating the cost of optimizing over $N_+^t(K)$ in terms of $n = |V|$ and t. Set $h(n, t) := \sum_{s=0}^{t-1}(2n)^s = 2^{t-1}n^{t-1} + O(n^{t-2})$. The formulation involves $\binom{n}{2} h(n, t)$ variables, $h(n, t)$ matrices of order $n+1$, and $(2n)^t$ conditions of type "$x \in K$". Furthermore, it turns out that for $1 \le r \le s$, the i_rth column of the matrix $Y_{i_1,\dots,i_s}^{\sigma_1,\dots,\sigma_s}$ is identically zero whenever $\sigma_r = -1$ and it is equal to the $\mathbf{0}$th column whenever $\sigma_r = 1$. Thus, $Y_{i_1,\dots,i_s}^{\sigma_1,\dots,\sigma_s}$ can be assumed to have order $n - s + 1$.

2.2. The Lasserre hierarchy

In this section we recall some basic facts about the Lasserre construction, applied to the case when K is of the form (1); for more information we refer to [8,9]. The Lasserre hierarchy involves moment matrices: A matrix whose rows and columns are indexed by a subset $\mathcal{A}$ of $\mathcal{P}(V)$ is said to be a *moment matrix* if the (I, J)th entry depends only on the union $I \cup J$ (for all $I, J \in \mathcal{A}$). In this definition one may allow $\mathcal{A}$ to be a multiset, which corresponds to repeated rows and columns in the moment matrix. For a non-negative integer t and a vector $y \in \mathbb{R}^{\mathcal{P}_{2t}(V)}$, define the *moment matrix of y of order t* by $M_t(y) := (y_{I\cup J})_{I,J\in\mathcal{P}_t(V)}$. For a set T and a vector $y \in \mathbb{R}^{\mathcal{P}(T)}$, we write $M_T(y) := (y_{I\cup J})_{I,J\in\mathcal{P}(T)}$; thus $M_n(y) = M_V(y)$ if $|V| = n$.

The following fact, observed in [13,9], explains the relevance of moment matrices to 0/1 polyhedra: For $x \in \mathbb{R}^V$ define $\zeta_x := (\prod_{i\in I} x_i)_{I\in\mathcal{P}(V)}$. Then we have for $y \in \mathbb{R}^{\mathcal{P}(V)}$

$$M_V(y) \succeq 0 \iff y \in \mathbb{R}_+\left\{\zeta_x \mid x \in \{0,1\}^V\right\}$$
$$\iff \forall S \subseteq V : \sum_{S' : S \subseteq S' \subseteq V} (-1)^{|S'\setminus S|} y_{S'} \ge 0. \tag{2}$$

In Lemma 2.2 we give an extension of this result.

Next we explain how to encode the linear constraints $a_l^{\mathsf{T}} x \ge 0$ describing K. Given $y \in \mathbb{R}^{\mathcal{P}_{2t}(V)}$ and $a \in \mathbb{R}^{\mathcal{P}_1(V)}$, define the vector $ay \in \mathbb{R}^{\mathcal{P}_{2t-1}(V)}$ by $(ay)_I := a_{\mathbf{0}} y_I + \sum_{i\in V} a_i y_{I\cup\{i\}}$ for $I \in \mathcal{P}_{2t-1}(V)$. For $t \ge 1$ we define the *t-th iterate of the Lasserre hierarchy* by

$$Q^t(K) := \{x \in \mathbb{R}^{\mathcal{P}_1(V)} \mid \exists y \in \mathbb{R}^{\mathcal{P}_{2t}(V)} : y_{\mathbf{0}} = x_{\mathbf{0}},\ y_i = x_i\ (i \in V),$$
$$M_t(y) \succeq 0,\ M_{t-1}(a_l y) \succeq 0\ (l = 1, \dots, m)\}. \tag{3}$$

It lies between K and C_K and $Q^{t+1}(K) \subseteq Q^t(K)$. The Lasserre hierarchy refines the Lovász–Schrijver hierarchy, since we have

$Q^{t+1}(K) \subseteq N_+(Q^t(K))$ which implies $Q^{t+1}(K) \subseteq N_+^t(K)$ and $Q^{n+1}(K) = C_K$. The formulation (3) involves $\sum_{i=0}^{2t}\binom{n}{i} = O(n^{2t})$ variables, one matrix of order $\sum_{i=0}^{t}\binom{n}{i} = O(n^t)$ and m matrices of order $\sum_{i=0}^{t-1}\binom{n}{i} = O(n^{t-1})$.

2.3. A new block-diagonal hierarchy

One drawback of the Lasserre hierarchy is that the computational cost for optimizing over $Q^t(K)$ is considerably higher than the cost for optimizing over $N_+^{t-1}(K)$. To define a more economical variation of it, which still refines the Lovász–Schrijver hierarchy, we consider a suitable principal submatrix of the full matrix $M_t(y)$.

For a positive integer t and a subset $T \subseteq V$ of cardinality $t-1$, let $M(T;y)$ denote the principal submatrix of $M_t(y)$ whose rows and columns are indexed by

$$\mathcal{A}(T) := \bigcup_{S\subseteq T}\mathcal{A}_S, \quad \text{where } \mathcal{A}_S := \{S\} \cup \{S\cup\{i\} \mid i \in V\}. \tag{4}$$

It will be convenient to consider $\mathcal{A}(T)$ as a multiset: We keep possible repeated occurrences, e.g. S and $S\cup\{i\}$ if $i \in S$. So strictly speaking the matrix $M(T;y)$ is a principal submatrix of $M_t(y)$ only after removing repeated rows and columns. We consider multisets here because it simplifies the notation in Lemma 2.2. Note that $M(\emptyset;y) = M_1(y)$, and observe that we only need to know the components of y indexed by $\mathcal{P}_{t+1}(V)$, instead of $\mathcal{P}_{2t}(V)$ as in the Lasserre hierarchy, in order to define the matrices $M(T;y)$ for all $T \in \mathcal{P}_{=(t-1)}(V)$.

Define the first iterate of the *block-diagonal hierarchy* by $L^1(K) := Q^1(K)$ and, for $t \geq 2$, define its tth iterate by

$$\begin{aligned} L^t(K) := \{x \in \mathbb{R}^{\mathcal{P}_1(V)} \mid \exists y \in \mathbb{R}^{\mathcal{P}_{t+1}(V)} : y_{\mathbf{0}} = x_{\mathbf{0}}, y_i = x_i\ (i \in V),\\ M(T;y) \succeq 0\ (T \in \mathcal{P}_{=(t-1)}(V)),\\ M(T;a_l y) \succeq 0\ (T \in \mathcal{P}_{=(t-2)}(V), l = 1,\ldots,m)\}. \end{aligned}$$

Since we used principal submatrices of the Lasserre hierarchy, we obviously have that the Lasserre hierarchy refines the block-diagonal hierarchy. As we see in Section 2.5 the block-diagonal hierarchy still refines the Lovász–Schrijver hierarchy.

Next we give a more compact formulation for the set $L^t(K)$, based on the fact that the matrix $M(T;y)$ has a special block structure which can be exploited to block-diagonalize it. This property justifies the name "block-diagonal hierarchy".

For a subset S of T, let $A_S(y)$ denote the principal submatrix of $M(T;y)$ indexed by the set $\mathcal{A}_S$, which is defined in (4). It is a $(n+1)\times(n+1)$ matrix lying in $\mathcal{M}_{+,V}$ with entries

$$A_S(y)_{\mathbf{0},\mathbf{0}} = y_S, \qquad A_S(y)_{\mathbf{0},i} = y_{S\cup\{i\}},$$
$$A_S(y)_{i,j} = y_{S\cup\{i,j\}} \quad (i,j \in V).$$

The submatrix $M(T;y)[S,S']$ of $M(T;y)$ with row indices in $\mathcal{A}_S$ and column indices in $\mathcal{A}_{S'}$ depends only on $S \cup S'$: $M(T;y)[S,S'] = A_{S\cup S'}(y)$.

Lemma 2.2. *The matrix $M(T;y)$ is positive semidefinite if and only if for all subsets S of T the matrix*

$$A(S,T)(y) := \sum_{S':S\subseteq S'\subseteq T}(-1)^{|S'\setminus S|}A_{S'}(y) \tag{5}$$

is positive semidefinite.

Proof. The proof is a "block-matrix version" of the one of (2) in [9]. Define the block matrix Z indexed by $\mathcal{A}(T)$, whose (S,S')th block is the identity matrix $\mathbf{I}$ of order $n+1$ if $S \subseteq S'$ and the zero matrix otherwise. Its inverse is the block matrix whose (S,S')th block is $(-1)^{|S'\setminus S|}\mathbf{I}$ if $S \subseteq S'$ and the zero matrix otherwise. Define the block-diagonal matrix D with diagonal blocks $A(S,T)(y)$ for $S \subseteq T$. Direct verification shows that $M(T;y) = ZDZ^{\mathsf{T}}$. Therefore,

$$M(T,y) \succeq 0 \iff D \succeq 0 \iff \forall S \subseteq T : A(S,T)(y) \succeq 0. \quad \square$$

Example 2.3. For $T = \{1,2\}$, $\mathcal{A}(T) = \mathcal{A}_{\mathbf{0}} \cup \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_{12}$ and

$$M(T;y) = \begin{pmatrix} A_{\mathbf{0}} & A_1 & A_2 & A_{12}\\ A_1 & A_1 & A_{12} & A_{12}\\ A_2 & A_{12} & A_2 & A_{12}\\ A_{12} & A_{12} & A_{12} & A_{12}\end{pmatrix} \succeq 0$$

$$\iff \begin{cases} A_{\mathbf{0}} - A_1 - A_2 + A_{12} \succeq 0\\ A_1 - A_{12} \succeq 0\\ A_2 - A_{12} \succeq 0\\ A_{12} \succeq 0\end{cases}$$

where we wrote A_S instead of $A_S(y)$.

Hence, in the formulation of $L^t(K)$, each condition $M(T;y) \succeq 0$, which involves one matrix of order $2^{t-1}(n+1)$, can be replaced by the 2^{t-1} conditions $A(S,T)(y) \succeq 0$, each involving a matrix of order $n+1$. Similarly, the condition $M(T;a_l y) \succeq 0$ can be replaced by the 2^{t-2} conditions $A(S,T)(a_l y) \succeq 0$, each involving a matrix of order $n+1$.

2.4. A variation of the block-diagonal hierarchy

The next lemma deals with other possible ways of encoding the linear conditions defining the set K. It motivates our second variation $\widetilde{L}^t(K)$. It turns out that it has an explicit link to the Lovász–Schrijver hierarchy. A main advantage of $\widetilde{L}^t(K)$ over $L^t(K)$ is that we do not need an explicit linear description of the set K in order to be able to define $\widetilde{L}^t(K)$. Hence $\widetilde{L}^t(K)$ enjoys the same complexity property as $N_+^t(K)$: If one can optimize in polynomial time over K then the same holds for $\widetilde{L}^t(K)$ for any fixed t.

Lemma 2.4. *Let $t \geq 1$, $y \in \mathbb{R}^{\mathcal{P}_{t+1}(V)}$, K be as in (1) and $A(S,T)(y)$ be as in (5). Then, the following two assertions are equivalent:*

(a) *For all $T \in \mathcal{P}_{=(t-1)}(V)$, $S \subseteq T$, $i \in V$:*

$$A(S,T)(y)e_i \in K, \quad A(S,T)(y)(e_{\mathbf{0}} - e_i) \in K.$$

(b) *For all $T \in \mathcal{P}_{=t}(V)$, $l = 1,\ldots,m$:*

$$M_T(a_l y) \succeq 0.$$

Proof. Using the identities

$$a_l^{\mathsf{T}}A_S(y)e_0 = (a_l y)_S, \qquad a_l^{\mathsf{T}}A_S(y)e_i = (a_l y)_{S\cup\{i\}},$$

the conditions $A(S,T)(y)e_i \in K, A(S,T)(y)(e_{\mathbf{0}} - e_i) \in K$ can be rewritten as

$$\sum_{S':S\subseteq S'\subseteq T}(-1)^{|S'\setminus S|}(a_l y)_{S'\cup\{i\}} \geq 0 \quad (l = 1,\ldots,m),$$
$$\sum_{S':S\subseteq S'\subseteq T}(-1)^{|S'\setminus S|}((a_l y)_{S'} - (a_l y)_{S'\cup\{i\}}) \geq 0 \quad (l = 1,\ldots,m).$$

On the other hand, using (2), $M_T(a_l y) \succeq 0$ is equivalent to

$$\sum_{S':S\subseteq S'\subseteq T}(-1)^{|S'\setminus S|}(a_l y)_{S'} \geq 0 \quad (S \subseteq T).$$

From this one can verify the equivalence of (a) and (b). $\square$

Observe that for $t = 1$ property (a) is equivalent to $A_{\mathbf{0}}(y)e_i$, $A_{\mathbf{0}}(y)(e_0 - e_i) \in K$ for all $i \in V$. Combined with the condition $A_{\mathbf{0}}(y) \succeq 0$, this characterizes membership in the set $N_+(K)$.

This motivates replacing in the definition of $L^t(K)$ the condition "$M(T;a_l y) \succeq 0$ for all $T \in \mathcal{P}_{=(t-2)}(V)$" by property (a): For $t \geq 1$ define

$$\begin{aligned} \widetilde{L}^t(K) := \{x \in \mathbb{R}^{\mathcal{P}_1(V)} \mid \exists y \in \mathbb{R}^{\mathcal{P}_{t+1}(V)} : y_{\mathbf{0}} = x_{\mathbf{0}}, y_i = x_i\ (i \in V),\\ M(T;y) \succeq 0\ (T \in \mathcal{P}_{=(t-1)}(V)),\\ A(S,T)(y)e_i \in K,\ A(S,T)(y)(e_{\mathbf{0}} - e_i) \in K\\ (T \in \mathcal{P}_{=(t-1)}(V), S \subseteq T, i \in V)\}. \end{aligned}$$

Table 1
Complexity comparison for $\widetilde{L}^t(K)$ and $N_+^t(K)$.

	$\widetilde{L}^t(K)$	$N_+^t(K)$
# variables	$\sum_{i=0}^{t+1}\binom{n}{i}$ $= \frac{1}{(t+1)!}n^{t+1}+O(n^t)$	$\binom{n}{2}\sum_{i=0}^{t}(2n)^i$ $= 2^{t-2}n^{t+1}+O(n^t)$
# matrices of order $n+1$	$\binom{n}{t-1}2^{t-1}$ $= \frac{2^{t-1}}{(t-1)!}n^{t-1}+O(n^{t-2})$	$\sum_{i=0}^{t}(2n)^i$ $= 2^{t-1}n^{t-1}+O(n^{t-2})$
# conditions "$x \in K$"	$2^t\binom{n}{t-1}$ $= \frac{2^t}{(t-1)!}n^{t-1}+O(n^{t-2})$	$2^t n^t$

2.5. Comparisons

Another advantage is that $\widetilde{L}^t(K)$ can be compared directly to the Lovász–Schrijver hierarchy $N_+^t(K)$. The next proposition shows that our second variation refines the Lovász–Schrijver hierarchy.

Proposition 2.5. *We have $\widetilde{L}^1(K) = N_+(K)$ and $\widetilde{L}^t(K) \subseteq N_+^t(K)$ for $t \geq 2$.*

Proof. As noted above we have $\widetilde{L}^1(K) = N_+(K)$. Now let $t \geq 2$ and $x \in \widetilde{L}^t(K)$. Thus, there is a $y \in \mathbb{R}^{\mathcal{P}_{t+1}(V)}$ which satisfies $y_\mathbf{0} = x_\mathbf{0}$, $y_i = x_i$ $(i \in V)$, and $M(T; y) \succeq 0$ or, equivalently, $A(S,T)(y) \succeq 0$ for all $S \subseteq T \subseteq V$ with $|T| = t-1$. Moreover property (a) of Lemma 2.4 holds. Set $Y := M_1(y)$. Then $x = Ye_\mathbf{0}$ and $Y \in \mathcal{M}_{+,V}$. Given $1 \leq s \leq t-1$, and $i_1, \ldots, i_s \in V$, and $\sigma \in \{\pm 1\}^s$, consider the multisets $T = \{i_1, \ldots, i_s\}$, $S = \{i_r \mid r = 1, \ldots, s, \sigma_r = 1\} \subseteq T$, and define $Y_{i_1,\ldots,i_s}^{\sigma_1,\ldots,\sigma_s} := A(S,T)(y)$. Here we extend the definition of $A(S,T)(y)$ in (5) to the case when S and T are multisets by taking the summation over all multisets S' lying between S and T; moreover, when S' is a multiset with S'' as underlying set, we let $A_{S'}(y) := A_{S''}(y)$. Now one can verify that the conditions from Proposition 2.1 hold, which implies $x \in N_+^t(K)$. $\square$

As one can see from the above proof, the main difference between $\widetilde{L}^t(K)$ and $N_+^t(K)$ is that the matrices $Y_{i_1,\ldots,i_s}^{\sigma_1,\ldots,\sigma_s}$ share many common entries in the definition of $\widetilde{L}^t(K)$. As a consequence, one can describe the set $\widetilde{L}^t(K)$ with less variables compared to $N_+^t(K)$. In Table 1 we compare the complexity of the formulations for $\widetilde{L}^t(K)$ and $N_+^t(K)$. In both cases one has a semidefinite programming formulation involving a number of matrices of size $n+1$ required to be positive semidefinite and a number of conditions of the type "$x \in K$".

Also, as already stated in Section 2.3, the block-diagonal hierarchy refines the Lovász–Schrijver hierarchy. This can be seen by comparing $L^{t+1}(K)$ with the second variation $\widetilde{L}^t(K)$.

Proposition 2.6. *For $t \geq 1$ the inclusion $L^{t+1}(K) \subseteq \widetilde{L}^t(K)$ holds.*

Proof. This follows directly from the definitions, after noting that, for $|T| = t$, the index set of $M_T(y)$ is contained in the index set of $M(T \setminus \{i\}; y)$, where i is any element of T. $\square$

3. Application to the stable set problem

In this section we apply the new hierarchies to the stable set problem. Let $G = (V, E)$ be a graph. A subset $S \subseteq V$ is called a stable set if none of its vertices are adjacent. The incidence vector of S is $\chi^S \in \{0,1\}^V$ with $\chi^S(i) = 1$ iff $i \in S$. The stability number $\alpha(G)$ is the maximum cardinality of a stable set. By $\mathcal{S}_G$ we denote the set of all stable sets of G. Then the stable set polytope is

$$\mathrm{STAB}(G) := \mathrm{conv}\{\chi^S \mid S \in \mathcal{S}_G\},$$

and the corresponding cone is

$$\mathrm{ST}(G) := \mathbb{R}_+ \left\{ \begin{pmatrix} 1 \\ \chi^S \end{pmatrix} \mid S \in \mathcal{S}_G \right\}.$$

A linear relaxation of ST(G) is the fractional stable set cone

$$\mathrm{FR}(G) := \{x \in \mathbb{R}^{\mathcal{P}_1(V)} \mid x_i \geq 0\ (i \in V),\ x_i + x_j \leq x_\mathbf{0}\ (\{i,j\} \in E)\}.$$

A semidefinite relaxation of ST(G) is the theta body

$$\mathrm{TH}(G) := \{Ye_\mathbf{0} \in \mathbb{R}^{\mathcal{P}_1(V)} \mid Y \in \mathcal{M}_{+,V},\ Y_{ij} = 0\ (\{i,j\} \in E)\},$$

which is contained in FR(G). Maximizing the linear function $\sum_{i \in V} x_i$ over the theta body TH(G) intersected with the hyperplane $x_\mathbf{0} = 1$ equals the Lovász theta function $\vartheta(G)$ introduced by Lovász in [12]. For details about these relaxations and the stable set problem we refer e.g. to [11,15].

In [9, Lemma 20] it was shown that when constructing the Lasserre hierarchy for FR(G) one can considerably simplify the formulation. One can replace the condition "$M_{t-1}(a_l y) \succeq 0$", where a_l runs through all linear inequalities defining FR(G), by the simpler equalities $y_{ij} = 0$, where $\{i,j\} \in E$, the so-called edge equalities. We want to apply the same simplification to the definition of $L^t(\mathrm{FR}(G))$ and define another variant $L^t(G)$ of it. However, in contrast to the Lasserre hierarchy, this simplification weakens the block-diagonal hierarchy a little bit since we can only claim the inclusion $L^t(\mathrm{FR}(G)) \subseteq L^t(G)$. Nevertheless the new variant $L^t(G)$ still refines the Lovász–Schrijver hierarchy, as $L^t(G) \subseteq N_+^{t-1}(\mathrm{TH}(G))$ follows from Proposition 3.2 below combined with Proposition 2.5. We define

$$\begin{aligned} L^t(G) := \{x \in \mathbb{R}^{\mathcal{P}_1(V)} \mid \exists y \in \mathbb{R}^{\mathcal{P}_{t+1}(V)} &: y_\mathbf{0} = x_\mathbf{0},\ y_i = x_i\ (i \in V), \\ &M(T; y) \succeq 0\ (T \in \mathcal{P}_{=(t-1)}(V)), \\ &y_{ij} = 0\ (\{i,j\} \in E)\}. \end{aligned}$$

Thus, $L^1(G) = \mathrm{TH}(G)$ and one can easily verify the inclusions $\widetilde{L}^t(\mathrm{FR}(G)) \subseteq L^t(G)$ when $t \geq 1$ and $L^t(\mathrm{FR}(G)) \subseteq L^t(G)$ when $t \geq 2$. Maximizing the objective function $\sum_{i \in V} x_i$ over $L^2(G)$ intersected with the hyperplane $x_\mathbf{0} = 1$ coincides with the parameter $\ell(G)$ considered in [4,6,7,10].

The next lemma says that the edge conditions in the definition of $L^t(G)$ imply that all variables indexed by non-stable sets are identically 0.

Lemma 3.1. *Let $y \in \mathbb{R}^{\mathcal{P}_{t+1}(V)}$ satisfy the conditions in the definition of $L^t(G)$. Then $y_I = 0$ for any subset $I \subseteq V$ with $|I| \leq t+1$ and containing an edge.*

Proof. For $|I| = 2$ the statement is nothing else but the edge equalities. Assume that $|I| \geq 3$, let $i, j \in I$ be adjacent vertices, and let k be another vertex in I. Define $T := I \setminus \{i,j\}$. The matrix $M(T; y)$ is positive semidefinite and the sets $\{i,j\}$ and $T \cup \{k\}$ occur in the index set $\mathcal{A}(T)$. As the (ij, ij)th entry of $M(T; y)$ is $y_{ij} = 0$, we have by the positive semidefiniteness of $M(T; y)$ that its $(ij, T \cup \{k\})$th entry is 0 as well and the statement of the lemma follows. $\square$

Proposition 3.2. *We have the inclusion $L^{t+1}(G) \subseteq \widetilde{L}^t(\mathrm{TH}(G))$ for $t \geq 1$.*

Proof. Assume that $y \in \mathbb{R}^{\mathcal{P}_{t+2}(V)}$ satisfies the conditions of the definition of $L^{t+1}(G)$. In the following we show that the vector consisting of the first $n+1$ coordinates of y belongs to $\widetilde{L}^t(\mathrm{TH}(G))$.

Fix $T \in \mathcal{P}_{t-1}(V)$, $S \subseteq T$ and $k \in V$. We show that $A(S,T)(y)e_k \in \mathrm{TH}(G)$ and $A(S,T)(y)(e_\mathbf{0} - e_k) \in \mathrm{TH}(G)$. For this we construct matrices Y^k and Z^k in $\mathcal{M}_{+,V}$ satisfying $Y_{ij}^k = Z_{ij}^k = 0$ when i and j are adjacent, and satisfying

$$Y^k e_\mathbf{0} = A(S,T)(y)e_k, \qquad Z^k e_\mathbf{0} = A(S,T)(y)(e_\mathbf{0} - e_k).$$

We distinguish between three cases.

(1) $k \in S$: Then $A(S,T)(y)e_\mathbf{0} = A(S,T)(y)e_k$ and define $Y^k := A(S,T)(y)$, $Z^k := 0$.
(2) $k \in T \setminus S$: Then $A(S,T)(y)e_k = 0$ and define $Y^k := 0$, $Z^k := A(S,T)(y)$.
(3) $k \in V \setminus T$: Then we define $Y^k = A(S \cup \{k\}, T \cup \{k\})(y)$ and $Z^k := A(S, T \cup \{k\})(y) = A(S,T)(y) - A(S \cup \{k\}, T \cup \{k\})(y)$.

In all cases we see by Lemmas 2.2 and 3.1 that Y^k, Z^k satisfy the desired conditions. $\square$

Table 2
Optimizing over $L^t(P_q)$ and $N_+(\mathrm{TH}(P_q))$ for Paley graphs.

q	$L^1(P_q)=\mathrm{TH}(P_q)$ $\vartheta(P_q)=\sqrt{q}$	$N_+(\mathrm{TH}(P_q))$	$L^2(P_q)$	$L^3(P_q)$	$\alpha(P_q)$
61	7.810	5.901	5.465	5.035	5
73	8.544	6.377	5.973	5.132	5
89	9.434	7.155	6.304	5.391	5
97	9.849	7.948	7.398	6.596	6
101	10.050	7.290	6.611	5.496	5
109	10.440	8.007	7.366	6.578	6
113	10.630	8.330	7.599	7.009	7
137	11.705	8.829	8.200	7.047	7
149	12.207	9.188	8.231	7.136	7
157	12.530	9.695	8.707	7.485	7
173	13.153	10.316	9.426	8.062	8
181	13.454	10.324	9.112	7.606	7
193	13.892	10.506	9.210	7.651	7
197	14.036	10.652	9.226	8.064	8
229	15.133	11.659	10.290	9.076	9
233	15.264	12.382	10.182	8.245	7
241	15.524	11.595	9.891	8.275	7
257	16.031	11.558	10.247	8.131	7
269	16.401	12.307	10.624	8.778	8
277	16.643	12.469	10.340	8.670	8
281	16.763	11.902	10.605	8.397	7
293	17.117	13.127	10.937	9.183	8
313	17.692	13.128	11.630	9.458	8
317	17.804	13.861	12.377	10.375	9
337	18.358	13.724	11.658	9.464	9
401	20.025	14.927	12.753	10.023	9
509	22.561	16.580	14.307	11.196	9
601	24.515	17.999	16.077	12.484	11
701	26.476	19.332	16.857	12.822	10
809	28.443	20.636	17.371	13.499	11

We summarize the inclusion relations between the various relaxations:

$$\mathrm{ST}(G) \subseteq Q^t(\mathrm{FR}(G)) \subseteq L^t(G) \subseteq \widetilde{L}^{t-1}(\mathrm{TH}(G)) \subseteq N_+^{t-1}(\mathrm{TH}(G)).$$

Moreover, $N_+^{t-1}(\mathrm{TH}(G)) = \mathrm{ST}(G)$ holds for $t \geq \alpha(G)$ (see [5] for a proof).

4. Experimental results

In this section we present some computational results for Paley graphs.

Let $\mathbb{F}_q$ be the finite field with prime power q which is congruent to 1 modulo 4; then -1 is a square in $\mathbb{F}_q$. The *Paley graph* P_q has $\mathbb{F}_q$ as vertex set and two distinct elements $u, v \in \mathbb{F}_q$ are adjacent if $u - v$ is a square in $\mathbb{F}_q$. The Paley graph is isomorphic to its complementary graph, it is a strongly regular graph and its automorphism group acts doubly-transitive on the vertices. It is known [12, Theorem 8] that $\vartheta(G)\vartheta(\overline{G}) = |V(G)|$ when G is a vertex-transitive graph, where $\overline{G}$ denotes the complementary graph of G. Since the Paley graph P_q is vertex-transitive and isomorphic to its complementary graph, we have $\vartheta(P_q) = \vartheta(\overline{P}_q) = \sqrt{q}$ (cf. [2, Theorem 13.14]). J.B. Shearer [16] has computed $\alpha(P_q)$ for all primes $q \leq 7000$. For more information about P_q we refer e.g. to [2, Chapter 13.2].

In order to illustrate the quality of the new relaxations $L^t(P_q)$, we have computed the bounds obtained by maximizing $\sum_{v\in V(P_q)} x_v$ over the sets $L^t(P_q)$ (for $t = 2, 3$) and $N_+(\mathrm{TH}(P_q))$ intersected with $x_{\mathbf{0}} = 1$. The results are given in Table 2. There we consider all primes q congruent to 1 modulo 4 between 61 and 337, as well as a few larger values of q up to 809. We have chosen the Payley graph here because its automorphism group acts doubly-transitive on the vertex set and so our formulation for $L^t(P_q)$ ($t \leq 3$) and $N_+(\mathrm{TH}(P_q))$ considerably simplifies. (See [5, Chapter 6.1] for implementation details.) For instance, optimization over $L^3(P_{809})$ (resp., $L^2(P_{809})$, $N_+(\mathrm{TH}(P_{809}))$) can be formulated via an SDP with 876 (resp., 36, 812) variables and with four matrices with sizes 808, 808, 404 and 202 (resp., two matrices with sizes 809 and 405, three matrices with sizes 810, 810 and 809). For the computations we used the program CSDP [3]. Experiments were conducted on a single machine with an Intel(R) Pentium(R) processor, 3 Ghz and 1GB of RAM. To compute the bounds from Table 2 we needed less than a minute when $q \leq 100$ and, for the largest instance P_{809}, around 45 min for $L^3(P_{809})$, 31 min for $L^2(P_{809})$ and 4.5 h for $N_+(\mathrm{TH}(P_{809}))$. Thus as expected the relaxation $L^2(G)$ gives a sharper bound than $N_+(\mathrm{TH}(G))$, however at a much smaller computational cost.

Finally note that one can strengthen the relaxation $L^t(G)$ by adding the non-negativity constraints $y \geq 0$. However this only gives a marginal improvement for Paley graphs, as the bounds differ only in decimals.

Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research under grant NWO 639.032.203. The third author was partially supported by the Deutsche Forschungsgemeinschaft (DFG) under grant SCHU 1503/4.

References

[1] E. Balas, S. Ceria, G. Cornuéjols, A lift-and-project cutting plane algorithm for mixed 0–1 programs, Mathematical Programming 58 (1993) 295–324.
[2] B. Bollobás, Random Graphs, Cambridge University Press, 2001.
[3] B. Borchers, CSDP, A C library for semidefinite programming, Optimization Methods and Software 11 (1999) 613–623.
[4] E. De Klerk, D.V. Pasechnik, A note on the stability number of an orthogonality graph, European J. Combin. 28 (2007) 1971–1979.
[5] N. Gvozdenović, Approximating the stability and the chromatic numbers of a graph via semidefinite programming, Ph.D. thesis, University of Amsterdam, 2008.
[6] N. Gvozdenović, M. Laurent, The operator Ψ for the chromatic number of a graph, SIAM J. Optim. 19 (2008) 572–591.
[7] N. Gvozdenović, M. Laurent, Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization, SIAM J. Optim. 19 (2008) 592–615.
[8] J.B. Lasserre, An explicit equivalent positive semidefinite program for nonlinear 0 − 1 programs, SIAM J. Optim. 12 (2002) 756–769.
[9] M. Laurent, A comparison of the Sherali-Adams, Lovász–Schrijver and Lasserre relaxations for 0 − 1 programming, Math. Oper. Res. 28 (2003) 470–496.
[10] M. Laurent, Strengthened semidefinite programming bounds for codes, Math. Program. 109 (2007) 239–261.
[11] M. Laurent, F. Rendl, Semidefinite programming and integer programming, in: K. Aardal, G. Nemhauser, R. Weismantel (Eds.), Handbook on Discrete Optimization, Elsevier, 2005, pp. 393–514.
[12] L. Lovász, On the Shannon capacity of a graph, IEEE Trans. Inform. Theory 25 (1979) 1–7.
[13] L. Lovász, A. Schrijver, Cones of matrices and set-functions and 0 − 1 optimization, SIAM J. Optim. 1 (1991) 166–190.
[14] G. Nemhauser, L.A. Wolsey, Integer and Combinatorial Optimization, Wiley, 1999.
[15] A. Schrijver, Combinatorial Optimization — Polyhedra and Efficiency, Springer Verlag, 2003.
[16] J.B. Shearer, Independence numbers of Paley graphs, 1996, available from the web site http://www.research.ibm.com/people/s/shearer/indpal.html.
[17] H.D. Sherali, W.P. Adams, A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems, SIAM J. Discrete Math. 3 (1990) 411–430.
[18] L.A. Wolsey, Integer Programming, Wiley, 1998.