Math. Program., Ser. A (2010) 121:249–268
DOI 10.1007/s10107-008-0233-x

FULL LENGTH PAPER

Copositive programming motivated bounds on the stability and the chromatic numbers

Igor Dukanovic · Franz Rendl

Received: 29 May 2006 / Accepted: 16 April 2008 / Published online: 4 June 2008

Abstract The Lovász theta number of a graph G can be viewed as a semidefinite programming relaxation of the stability number of G. It has recently been shown that a copositive strengthening of this semidefinite program in fact equals the stability number of G. We introduce a related strengthening of the Lovász theta number toward the chromatic number of G, which is shown to be equal to the fractional chromatic number of G. Solving copositive programs is NP-hard. This motivates the study of tractable approximations of the copositive cone. We investigate the Parrilo hierarchy to approximate this cone and provide computational simplifications for the approximation of the chromatic number of vertex transitive graphs. We provide some computational results indicating that the Lovász theta number can be strengthened significantly toward the fractional chromatic number of G on some Hamming graphs.

Keywords Copositive programming · Semidefinite programming · (fractional) Chromatic number · Lovász theta number

Mathematics Subject Classification (2000) 90C27 · 90C22 · 90C06

Partial support by the EU project Algorithmic Discrete Optimization (ADONET), MRTN-CT-2003-504438, is gratefully acknowledged.

I. Dukanovic (✉)
Univerza v Mariboru, EPF, Razlagova 14, 2000 Maribor, Slovenia
e-mail: igor.mat@uni-mb.si

F. Rendl
Universität Klagenfurt, Institut für Mathematik, Universitätsstraße 65-67, 9020 Klagenfurt, Austria
e-mail: franz.rendl@uni-klu.ac.at

1 Introduction

The Lovász theta number $\vartheta(G)$ has been introduced as an upper bound on the Shannon capacity of a graph G [22]. It can be formulated as the optimal value of a semidefinite program (SDP), and therefore computed to an arbitrary fixed precision in polynomial time (see the survey by Todd [33]). The theta number and various strengthenings toward the stability number $\alpha(G)$ have been studied intensively in the last years (see the survey [21]). Recently a strengthening $\vartheta^{\mathcal{C}}(G)$ of $\vartheta(G)$ toward the stability number has been introduced by de Klerk and Pasechnik in [6]. The strengthening is obtained by going from semidefinite matrices to the cone of completely positive matrices. De Klerk and Pasechnik in fact show that this approximation is tight, $\vartheta^{\mathcal{C}}(G) = \alpha(G)$.
Lovász showed that the theta number is also a lower bound on the chromatic number [22]. Based on the semidefinite program defining the theta number, Karger, Motwani and Sudan suggested a graph-coloring heuristic which was a major advance at that time for the worst case analysis [17]. Their breakthrough result has been slightly improved since, see, e.g., [16]. Inspired by the definition of $\vartheta^{\mathcal{C}}(G)$ [6] we introduce a related copositive program $\Theta^{\mathcal{C}}(G)$ which strengthens the Lovász theta number toward the chromatic number.[1] While the copositive programming relaxation of the stability number is exact, we will see that the number $\Theta^{\mathcal{C}}(G)$ equals the fractional chromatic number, which is a lower bound on the chromatic number. An exact copositive programming formulation of the chromatic number has recently been proposed by Gvozdenović and Laurent [15]. It has recently been shown [28] that the quadratic assignment problem also has an exact copositive programming formulation. In fact, quite a general class of quadratic programs with linear as well as binary constraints allows a copositive representation [3].
Copositive programs cannot be solved efficiently. Just testing whether a matrix is copositive is co-NP-complete [25]. So a chain of inner approximations of the copositive cone $\mathcal{P} \subseteq \mathcal{K}_0 \subseteq \mathcal{K}_1 \subseteq \cdots \subseteq \mathcal{C}$ has been suggested in [26]. Optimizing over the semidefinite cone $\mathcal{P}$ and the cone $\mathcal{K}_0$ are well-understood (see [10]). On the other hand, optimizing over the next weakest cone $\mathcal{K}_1$ is a large SDP with $O(n^3)$ variables, $n = |V(G)|$. We show that this relaxation can be restated as a program with just $O(n^2)$ variables in the case of a vertex transitive graph. We also show that bounds on the stability and the chromatic number obtained by optimizing over any cone $\mathcal{K}$ are related. In fact, in the case of vertex transitive graphs a simple change of variables transforms one SDP into the other.
We give preliminary numerical experience with the cone $\mathcal{K}_1$ on vertex transitive Hamming graphs. On some instances it shows a significant strengthening over the Lovász theta number.

[1] A remark on notation: actually, a lower bound on the chromatic number $\chi(G)$ is the theta number of the complement graph $\vartheta(\bar{G})$. In the literature on the stability number this bound is often denoted as $\bar{\vartheta}(G)$. However, to clarify the exposition on the lower bounds of the chromatic number, and to emphasize the symmetry in approaches to bound the stability number and the (fractional) chromatic number, we use the notation $\Theta(G) := \bar{\vartheta}(G)$. So we denote by $\Theta^{\mathcal{C}}(G)$ the corresponding copositive programming bound which would by analogy to the existing literature be denoted by $\bar{\vartheta}^{\mathcal{C}}(G)$.

2 Notation

A graph with vertex set $V := V(G) := \{1, \ldots, n\}$ and edge set E will be denoted by $G(V, E)$, or shortly G. We assume G to be simple (loopless, without multiple edges) and undirected. The *complement* graph $\bar{G}$ is the simple graph on the same vertex set V and the edge set $\bar{E} := \{ij \notin E : i \neq j\}$. A (proper) *s-coloring* of a graph $G(V, E)$ is a mapping $c : V \to \{1, \ldots, s\}$ such that $ij \in E \Rightarrow c(i) \neq c(j)$. *The chromatic number* $\chi(G)$ is the smallest number of colors needed to (properly) color the graph G. *The stability number* $\alpha(G)$ is the size of the largest stable set in the graph G. (A subset S of the vertices of G is stable, if no edge of G has both endpoints in S.)
The n-dimensional vector of all ones is denoted by e. The matrix $J = ee^T$ is the $n \times n$ matrix of all ones. In fact, rows and columns of any matrix considered in this paper are induced by the n vertices of the graph G, and are therefore $n \times n$ real matrices. An edge ij induces the symmetric matrix $E_{ij} = e_i e_j^T + e_j e_i^T$ where e_i is the ith column of the identity matrix I. All our matrices belong to the universal space of the symmetric matrices denoted by $\mathcal{S}^n$ with the trace inner product

$$\langle X, Y \rangle := \text{tr } XY = \sum_{i,j=1}^{n} x_{ij} y_{ij}.$$

The operator diag: $\mathcal{S}^n \to \mathbb{R}^n$ maps a matrix into its diagonal. A set $\mathcal{K}$ is a *cone*, if $K \in \mathcal{K}, \lambda \geq 0 \Rightarrow \lambda K \in \mathcal{K}$. The cone K is *convex*, if it additionally satisfies $K, K' \in \mathcal{K} \Rightarrow K + K' \in \mathcal{K}$. We will consider only convex cones, such as $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \geq 0\}$ and

$$\mathcal{N} := \{K \in \mathcal{S}^n : K \geq 0\}, \tag{1}$$

the cone of elementwise nonnegative matrices. Positive semidefinite matrices also form a convex cone, denoted $\mathcal{P} := \{K \in \mathcal{S}^n : K \succeq 0\}$. Its interior is $\{K \in \mathcal{S}^n : K \succ 0\}$ where $A \succeq B$ ($A \succ B$) means that $A - B$ is a positive semidefinite (respectively positive definite) matrix. The *dual* of a cone $\mathcal{K}$ is defined by

$$\mathcal{K}^* := \{D \in \mathcal{S}^n : \langle K, D \rangle \geq 0 \ \forall K \in \mathcal{K}\}. \tag{2}$$

The cones $\mathcal{N}$ and $\mathcal{P}$ are self-dual, $\mathcal{N}^* = \mathcal{N}$ and $\mathcal{P}^* = \mathcal{P}$. The following well-known characterization of positive semidefinite matrices

$$\mathcal{P} = \{K \in \mathcal{S}^n : x^T K x \geq 0 \ \forall x \in \mathbb{R}^n\} \tag{3}$$

shows that $\mathcal{P}$ is a subcone of the cone of copositive matrices defined by

$$\mathcal{C} := \{K \in \mathcal{S}^n : x^T K x \geq 0 \ \forall x \in \mathbb{R}^n_+\}. \tag{4}$$

Its dual is the cone of completely positive matrices given by $\mathcal{C}^* = \{x_1 x_1^T + \cdots + x_k x_k^T : x_i \in \mathbb{R}^n_+ \ \ i = 1, \ldots, k, \ \text{ and } k \in \mathbb{N}\}$. Now $\mathcal{P} \subseteq \mathcal{C}$ implies $\mathcal{C}^* \subseteq \mathcal{P}^*$.

Optimization over $\mathcal{N}$ corresponds to linear programming and optimizing over $\mathcal{P}$ to semidefinite programming. We refer to both optimizing over the cone $\mathcal{C}$ and its dual cone $\mathcal{C}^*$ as copositive programming. This is justified by the fact that in our instances there is no duality gap (see Remark 7), so primal-dual techniques apply.

3 The stability number and its semidefinite programming relaxation

We briefly outline the technique for developing semidefinite and copositive programming relaxations of the stability number. The *characteristic vector* of a set $S \subseteq V$, $\chi_S \in \{0, 1\}^n$ is defined by

$$(\chi_S)_v := \begin{cases} 1 & v \in S, \\ 0 & v \notin S. \end{cases}$$

Any nonempty set S induces the matrix

$$X_S := \frac{1}{\chi_S^T \chi_S} \chi_S \chi_S^T. \tag{5}$$

This matrix satisfies the following conditions: $X_S \in \mathcal{C}^* \subseteq \mathcal{P}$, tr $(X_S) = 1$ and $|S| = \langle X_S, J \rangle$.

Let S be a stable set. Then clearly at most one endpoint of any edge can be in S, i.e., $(\chi_S)_i(\chi_S)_j = 0$ whenever $ij \in E$. In terms of the matrix (5) this is equivalent to $(X_S)_{ij} = 0$ for any edge ij. To describe this sparsity pattern of the matrix X_S we introduce the following linear operator $A_G : \mathcal{S}^n \to \mathbb{R}^E$

$$(A_G X)_{ij} := \langle X, E_{ij} \rangle = x_{ij} + x_{ji} = 2x_{ij} \qquad \forall ij \in E. \tag{6}$$

So $A_G X_S = 0$. The Lovász theta number introduced in [22]

$$\vartheta(G) := \max\{\langle X, J \rangle : \text{ tr } X = 1, \ A_G X = 0, X \in \mathcal{P}\}, \tag{7}$$

and the optimum of the following copositive program introduced by de Klerk and Pasechnik in [6]

$$\vartheta^{\mathcal{C}}(G) := \max\{\langle X, J \rangle : \text{ tr } X = 1, \ A_G X = 0, \ X \in \mathcal{C}^*\} \tag{8}$$

are upper bounds on $\alpha(G)$ since the matrix X_S induced by any stable set is feasible for (7) and (8).

Since $\mathcal{C}^* \subseteq \mathcal{P}$

$$\alpha(G) \leq \vartheta^{\mathcal{C}}(G) \leq \vartheta(G). \tag{9}$$

The first inequality is tight [6].

Theorem 1 (de Klerk–Pasechnik) $\alpha(G) = \vartheta^{\mathcal{C}}(G)$. □

Since it is NP-hard to compute the stability number (see, e.g., [31]), computing $\vartheta^{\mathcal{C}}(G)$ is NP-hard, too. On the other hand, the semidefinite program (8) is solvable in polynomial time.

4 Graph coloring

In this section we recall the well-known linear programming (LP) and semidefinite programming (SDP) relaxations for graph coloring. Then, inspired by Theorem 1, we introduce a new copositive programming relaxation $\Theta^{\mathcal{C}}(G)$ complementing $\vartheta^{\mathcal{C}}(G)$, and prove $\Theta^{\mathcal{C}}(G) = \chi_f(G)$, the fractional chromatic number.

An s-coloring c partitions the vertex set V into s stable sets (*color classes*) defined by $S_i := c^{-1}(i)$. In fact, any partition of the vertex set into stable sets $S_1, \dots, S_s$ defines an s-coloring by $c(v) = i$ whenever $v \in S_i$. A partition of V into s sets is characterized by $\chi_{S_1} + \cdots + \chi_{S_s} = e$.

Let us denote the family of all stable sets by $\mathcal{S}$, i.e., $\mathcal{S} := \{S \subseteq V : S \text{ stable}\}$. A subfamily $\mathcal{T} \subseteq \mathcal{S}$ partitions V, if $\sum_{S \in \mathcal{T}} \chi_S = e$. The chromatic number $\chi(G)$ is the minimal cardinality of such a partitioning into stable sets, i.e.,

$$\chi(G) = \min\left\{\sum_{S\in\mathcal{S}} \lambda_S : \lambda_S \in \{0,1\}, \sum_{S\in\mathcal{S}} \lambda_S \chi_S = e\right\}. \tag{10}$$

It is well-known that computing the chromatic number is NP-hard, see, e.g., [31]. The *fractional chromatic number* $\chi_f(G)$ defined by the linear program

$$\chi_f(G) := \min\left\{\sum_{S\in\mathcal{S}} \lambda_S : \lambda_S \geq 0, \sum_{S\in\mathcal{S}} \lambda_S \chi_S = e\right\} \tag{11}$$

is obviously a lower bound on $\chi(G)$. Unfortunately, this number is also NP-hard to compute, (see, e.g., the recent reformulation of the stability number $\alpha(G)$ as (fractional) chromatic number of a certain graph [15]).

Let $S_1, \dots, S_{\chi(G)}$ be a partition of V into $\chi(G)$ stable sets, and let us define the corresponding *coloring matrix* $C \in \mathcal{S}^n$ by

$$C := \sum_{i=1}^{\chi(G)} \chi_{S_i} \chi_{S_i}^T. \tag{12}$$

This coloring matrix in turn uniquely determines a partition of the vertex set V, i.e., if we denote by $S_{[u]} \in \{S_1, \dots, S_{\chi(G)}\}$ the stable set which contains vertex u,

$$c_{uv} = e_u^T C e_v = \sum_{i=1}^{\chi(G)} (e_u^T \chi_{S_i})(\chi_{S_i}^T e_v) = \begin{cases} 1 & S_{[u]} = S_{[v]}, \\ 0 & S_{[u]} \neq S_{[v]}. \end{cases} \tag{13}$$

Since the endpoints of an edge uv cannot be in the same color class, $c_{uv} = 0$. By applying operator (6) this fact can be written as

$$A_G C = 0. \tag{14}$$

On the other hand $c_{uu} = 1$ for any $u \in V$, i.e.,

$$\operatorname{diag} C = e. \tag{15}$$

Furthermore $C \in \mathcal{C}^* \subseteq \mathcal{P}$. These facts motivate all our semidefinite and copositive programming relaxations of the chromatic number. So the following Lemma 2 which relates vectors appearing in the definition (11) to matrices which generalize (12) is the key to linking linear to semidefinite and copositive programming relaxations.

Lemma 2 *Let $S_i \subseteq V$ and $\lambda_i \geq 0$ for $i = 1, \ldots, k$. Define the matrix $X_\lambda := \sum_{i=1}^k \lambda_i \chi_{S_i} \chi_{S_i}^T$ and the vector $x_\lambda := \sum_{i=1}^k \lambda_i \chi_{S_i}$. Then*

$$M := \left(\sum_{i=1}^k \lambda_i\right) X_\lambda - x_\lambda x_\lambda^T \succeq 0.$$

Proof We have to show that $x^T M x \geq 0$ for any vector $x \in \mathbb{R}^n$. Note that

$$M = \sum_{i,j=1}^k \lambda_i \lambda_j \chi_{S_j} \chi_{S_j}^T - \sum_{i,j=1}^k \lambda_i \lambda_j \chi_{S_i} \chi_{S_j}^T.$$

Denote $a_i := \chi_{S_i}^T x$. Then $x^T M x = \sum_{i,j=1}^k \lambda_i \lambda_j (a_j^2 - a_i a_j) = \sum_{i=1}^k \sum_{j>i} \lambda_i \lambda_j (a_j^2 - 2a_j a_i + a_i^2) \geq 0$. □

The coloring matrix (12) can be written as

$$C = \sum_{i=1}^{\chi(G)} \lambda_i \chi_{S_i} \chi_{S_i}^T,$$

where each $\lambda_i = 1$. Since the sum of characteristic vectors of a partition $x_\lambda := \sum_{i=1}^{\chi(G)} \lambda_i \chi_{S_i} = e$, by Lemma 2 we get

$$\left(\sum_{i=1}^k \lambda_i\right) C - ee^T = \chi(G) C - J \succeq 0.$$

Now $C \succeq 0$ implies

$$t \geq \chi(G) \Rightarrow tC - J \succeq 0.$$

Also notice that $t \geq \chi(G) > 0$ implies $tC \in \mathcal{C}^*$.
The adjoint operator of (6) is given by $A_G^T y = \sum_{ij \in E(G)} y_{ij} E_{ij}$. So the conditions (14) and (15) can also be written as

$$tC = tI + A_{\bar{G}}^T y, \tag{16}$$

where $y \in \mathbb{R}^{\bar{E}}$. Therefore the Lovász theta number

$$\Theta(G) := \min\{t : tI + A_{\bar{G}}^T y - J \succeq 0\}, \tag{17}$$

and the optimum of the copositive program

$$\Theta^{\mathcal{C}}(G) := \min\{t : tI + A_{\bar{G}}^T y - J \succeq 0, tI + A_{\bar{G}}^T y \in \mathcal{C}^*\} \tag{18}$$

are lower bounds on $\chi(G)$. Since the feasible set of (18) is contained in the one of (17)

$$\Theta(G) \leq \Theta^{\mathcal{C}}(G) \leq \chi(G). \tag{19}$$

The rest of this section is devoted to proving $\Theta^{\mathcal{C}}(G) = \chi_f(G)$ defined by (11). To our knowledge, the relaxation (18) has not been investigated before.

Lemma 3 $\Theta^{\mathcal{C}}(G) \leq \chi_f(G)$.

Proof Let $\{S_1, \ldots, S_k\}$ be stable sets and $\lambda_i \geq 0$ such that $\sum_{i=1}^k \lambda_i \chi_{S_i} = e$ is an optimal solution of (11). Then $t := \chi_f(G) = \sum_{i=1}^k \lambda_i \geq 0$. Since $\chi_S \in \{0, 1\}^n$ is a vector of zeros and ones, diag $(\chi_S \chi_S^T) = \chi_S$, and the diagonal of the completely positive matrix

$$C_\lambda := \sum_{i=1}^k \lambda_i \chi_{S_i} \chi_{S_i}^T, \quad \text{diag } (C_\lambda) = \sum_{i=1}^k \lambda_i \chi_{S_i} = e. \tag{20}$$

This resembles (15). Likewise for $ab \in E$, $e_a^T C_\lambda e_b = \sum_{i=1}^k \lambda_i (e_a^T \chi_{S_i})(\chi_{S_j}^T e_b) = \sum 0 = 0$ as in (14). Again like (14) and (15) together produce (16), we get $tC_\lambda = tI + A_{\bar{G}}^T y$. If $tI + A_{\bar{G}}^T y - J \succeq 0$, (t, y) is feasible for (18) and $\Theta^{\mathcal{C}}(G) \leq t = \chi_f(G)$. But indeed by applying Lemma 2 on C_λ where by (20) $\sum \lambda_i \chi_{S_i} = e$

$$\left(\sum_{i=1}^k \lambda_i\right) C_\lambda - ee^T \succeq 0,$$

which is equivalent to $tC_\lambda - J = tI + A_{\bar{G}}^T y - J \succeq 0$. □

Lemma 4 [12] $\chi_f(G) \leq \Theta^{\mathcal{C}}(G)$.

Proof The dual of the linear program (11) is

$$\chi_f(G) = \max\{e^T w : \chi_S^T w \le 1 \ \forall S \in \mathcal{S}\}. \tag{21}$$

Let w be an optimal solution of (21). Notice that $w \ge 0$ as otherwise w_+ defined by

$$(w_+)_i = \begin{cases} w_i & w_i \ge 0, \\ 0 & \text{otherwise} \end{cases}$$

would also be feasible as $\chi_S^T w_+ = \chi_{S\cap\{i:w_i\ge 0\}}^T w \le 1$ by (21), contradicting $\chi_f(G) = e^T w < e^T w_+$.
Let (t, y) be an optimal solution of (18). As $U := tI + A_{\bar{G}}^T y \in \mathcal{C}^*$ there are vectors $x_k \ge 0$ such that $U = \sum_k x_k x_k^T$. For any edge ij, $0 = u_{ij} = \sum_k (x_k)_i (x_k)_j$ must be a sum of zeros. Therefore the support of each vector x_k

$$S_k := \{i : (x_k)_i > 0\}$$

is a stable set. Moreover $t = u_{ii} = \sum_k (x_k)_i^2$ for any $i \in V$. As $Z := tI + A_{\bar{G}}^T y - J = U - J = (\sum_k x_k x_k^T) - ee^T \succeq 0$ by (18), we deduce $w^T Z w \ge 0$ which is equivalent to

$$(e^T w)^2 \le \sum_k (w^T x_k)^2. \tag{22}$$

After writing $w_i = \sqrt{w_i}\sqrt{w_i}$, each summand in (22) satisfies

$$\begin{aligned}\left(w^T x_k\right)^2 &= \left(\sum_{i\in S_k} w_i (x_k)_i\right)^2 \le \left(\sum_{i\in S_k} w_i (x_k)_i^2\right) \sum_{i\in S_k} w_i \le \left(\sum_{i\in S_k} w_i (x_k)_i^2\right) \\ &= \left(\sum_i w_i (x_k)_i^2\right)\end{aligned}$$

by the Cauchy–Schwartz inequality, since $\sum_{i\in S_k} w_i = \chi_{S_k}^T w \le 1$ by (21) as S_k is stable.
Therefore, their sum $\sum_k (w^T x_k)^2 \le \sum_k \sum_i w_i (x_k)_i^2 = \sum_i \left(\sum_k (x_k)_i^2\right) w_i = \sum_i t w_i = te^T w$. Combining with (22) this gives $(e^T w)^2 \le te^T w$. Since $\chi_f(G) = e^T w > 0$, we get $\chi_f(G) \le t$. □

Corollary 5 $\Theta^{\mathcal{C}}(G) = \chi_f(G)$.

Remark 6 In [3], Burer suggests a copositive programming reformulation of a large class of quadratic programs including the linear formulation (11) of the fractional chromatic number $\chi_f(G)$. However his formulation for this particular problem is intractable as it would require as many variables as there are stable sets in the graph G.

5 Relations between strengthenings of ϑ and Θ

Looking at the chains of relaxations (9) and (19) we conclude that having a tractable cone $\mathcal{K}$ "between" $\mathcal{P}$ and $\mathcal{C}$, $\mathcal{P} \subseteq \mathcal{K} \subseteq \mathcal{C}$, yields a tractable relaxation for the respective problems. Motivated by [6] we define

$$\vartheta^{\mathcal{K}}(G) := \max\{\langle X, J\rangle : \text{tr } X = 1, A_G X = 0, X \in \mathcal{K}^*\}, \tag{23}$$

and the corresponding bound on the chromatic number

$$\Theta^{\mathcal{K}}(G) := \min\{t : tI + A_{\bar{G}}^T y - J \in \mathcal{P}, tI + A_{\bar{G}}^T y \in \mathcal{K}^*\}. \tag{24}$$

Notice that it generalizes (8) and (18). Meanwhile $\vartheta^{\mathcal{P}}(G) = \vartheta(G)$, and $\Theta^{\mathcal{P}}(G) = \Theta(G)$ since $J \in \mathcal{P}$ implies that the constraint $tI + A_{\bar{G}}^T y \in \mathcal{P}$ is redundant in (17). Now $\mathcal{C}^* \subseteq \mathcal{K}^* \subseteq \mathcal{P}$ implies

$$\alpha(G) = \vartheta^{\mathcal{C}}(G) \le \vartheta^{\mathcal{K}}(G) \le \vartheta(G), \quad \text{and} \quad \Theta(G) \le \Theta^{\mathcal{K}}(G) \le \Theta^{\mathcal{C}}(G) = \chi_f(G).$$

Remark 7 The dual of (23)

$$\inf\{t : tI + A_G^T y - J \in \mathcal{K}\} = \vartheta^{\mathcal{K}}(G), \tag{25}$$

(i.e., there is no duality gap) because $X = \frac{1}{n} I \in \mathcal{C}^* \subseteq \mathcal{K}^*$ while $t = 2n$ and $y = 0$ define a matrix $2nI - J \succ 0$ in the strict interior of $\mathcal{K} \supseteq \mathcal{P}$. The conic duality theorem (see [5]) also states that the optimum of (23) is indeed attained. Likewise the dual of (24) satisfies

$$\Theta^{\mathcal{K}}(G) = \sup\{\langle X - W, J\rangle : \text{tr } X = 1, A_{\bar{G}} X = 0, X - W \succeq 0, W \in \mathcal{K}\}, \tag{26}$$

and the optimum of (24) is attained as $X - W = W = \frac{1}{2n} I \succ 0$ is strictly feasible for (26) while $t = 2n$, $y = 0$ is feasible for (24).

Compare $\vartheta(\bar{G})$ computed from (25) to (17). The resulting well-known equation $\Theta(G) = \vartheta(\bar{G})$ was proved already in [22]. The next proposition is a generalization of the inequality $\vartheta(G)\vartheta(\bar{G}) \ge n$ [22]. The case $\mathcal{K} = \mathcal{P} + \mathcal{N}$ was proved in [32].

Proposition 8 *Let $\mathcal{K}$ be a cone such that $\mathcal{P} \subseteq \mathcal{K} \subseteq \mathcal{C}$. Then $\vartheta^{\mathcal{K}}(G)\Theta^{\mathcal{K}}(G) \ge n$.*

Proof Let the optimum $t^* := \Theta^{\mathcal{K}}(G)$ of (24) be achieved at (t^*, y^*). Since $Z^* := t^* I + A_{\bar{G}}^T y - J \succeq 0$, its diagonal elements $t^* - 1 \ge 0$. Therefore $X := \frac{1}{nt^*}(t^* I + A_{\bar{G}}^T y^*) = \frac{1}{n} I + A_{\bar{G}}^T(\frac{1}{nt^*} y^*)$ is feasible for (23). Since $X = \frac{1}{nt^*}(Z^* + J)$

$$\vartheta^{\mathcal{K}}(G) \ge \langle X, J\rangle = e^T X e = \frac{e^T Z^* e + e^T J e}{nt^*} \ge \frac{0 + n^2}{nt^*} = \frac{n}{\Theta^{\mathcal{K}}(G)}.$$

□

Remark 9 A polynomial bound $p(G)$ provably improves $\Theta(G)$ toward the chromatic number, if whenever

$$\Theta(G) \neq \chi(G) \Rightarrow \Theta(G) < p(G) \leq \chi(G).$$

The existence of such a bound remains an open question. Busygin and Pasechnik [4] however prove that no polynomial bound improves $\vartheta(G)$ toward $\alpha(G)$ unless P = NP. Likewise Gvozdenović and Laurent prove that there is no polynomial bound between $\chi_f(G)$ and $\chi(G)$ unless P = NP [15]. Next, let a polynomial bound $p(G)$ lay between $\Theta(G)$ and $\chi_f(G)$. (Notice that all conic bounds described in this paper indeed satisfy $\Theta(G) \leq \Theta^{\mathcal{K}}(G) \leq \chi_f(G)$.) Since $\Theta(\bar{H}(6,2)) = \chi_f(\bar{H}(6,2)) = 10 + 2/3$ where $H(m,n)$ stands for the Hamming graph described in Sect. 9, see [8] for details, $p(G)$ does not provably improve $\Theta(G)$ toward the chromatic number $\chi(G)$ either.

6 Vertex transitive graphs

The numerical experiments as reported in [10,14] suggest that strengthening the Lovász theta number on highly structured vertex transitive graphs can be substantial while the computational effort can be kept reasonable. So in the rest of the paper we will consider only the vertex transitive graphs.

Definition 10 Let $G(V,E)$ be a simple graph. A permutation $\pi\colon V \to V$ of the vertices such that $ij \in E \iff \pi(i)\pi(j) \in E$ is an automorphism of the graph. The group of all automorphisms is denoted by $Aut(G)$. Group $H \leq Aut(G)$ is *vertex transitive*, if for every pair $i, j \in V$ there exists an automorphism

$$\pi_{ij} \in H \text{ such that } \pi_{ij}(i) = j. \tag{27}$$

Graph G is *vertex transitive*, if $Aut(G)$ is a vertex transitive group.

Vertex transitive graphs possess a lot of symmetry and thus enable us to apply the technique which goes back at least to [30]. Its main idea is roughly sketched in Lemma 12.

Definition 11 Let $\bar{\mathcal{K}}$ be a subcone of the cone $\mathcal{K}$, i.e., $\bar{\mathcal{K}} \subseteq \mathcal{K}$. A (continuous) mapping $r : \mathcal{K} \to \bar{\mathcal{K}}$ is a *retraction* to its subcone $\bar{\mathcal{K}}$ and $\bar{\mathcal{K}}$ a *retract* of $\mathcal{K}$, if $r|_{\bar{\mathcal{K}}} = id$.

Lemma 12 *Let $r : \mathcal{K} \to r(\mathcal{K})$ be a retraction. Let $c : \mathcal{K} \to \mathbb{R}$ be such a map that $c(A) = c(r(A))$ for each $A \in \mathcal{K}$. Then*

$$\max\{c(A) : A \in \mathcal{K}\} = \max\{c(\bar{A}) : \bar{A} \in \bar{\mathcal{K}}\}. \tag{28}$$

Proof $\max\{c(A) : A \in \mathcal{K}\} = \max\{c(r(A)) : A \in \mathcal{K}\} = \max\{c(r(A)) : r(A) \in \bar{\mathcal{K}}\}$. □

In the most interesting case the dimension of the retract $\bar{\mathcal{K}}$ is considerably smaller than the dimension of the original cone $\mathcal{K}$ implying that $\bar{\mathcal{K}}$ can be described by substantially fewer variables. As the space complexity alone determines which conic programs

considered in this paper are practically solvable replacing the former program in (28) by the latter enables us to compute bounds on much larger (vertex transitive) graphs. (Note that the time complexities of all programs considered here are cubic in their space complexities as we solve them by an interior point method.)
Inspired by [11], we study in this section cases where the structure of the cone $\mathcal{K}$ is described by a vertex transitive group $H \leq Aut(G)$. To transmit the structure from permutations of $n = |V|$ vertices to the corresponding cone of $n \times n$ matrices we need to define for each permutation $\pi : V \to V$ its *permutation matrix* P_π by

$$P_\pi e_i = e_{\pi(i)}.$$

Definition 13 Matrix M is H-invariant,[2] if $P_\beta^T M P_\beta = M$ for any $\beta \in H$, i.e., $m_{ij} = m_{\beta(i)\beta(j)}$.

Definition 14 Cone $\mathcal{K}$ *preserves graph isomorphisms*, if it satisfies

$$Z \in \mathcal{K} \Rightarrow P_\pi^T Z P_\pi \in \mathcal{K} \tag{29}$$

for any automorphism $\pi \in \text{Aut}\ (G)$.

Obviously the cones $\mathcal{N}, \mathcal{P}, \mathcal{C}$ and $\mathcal{N} + \mathcal{P}$ preserve graph isomorphisms, see (1), (3) and (4). For further examples of such cones see, e.g., [6,10,26]. Next, consider the dual $\mathcal{K}^*$ (see definition (2)) of a cone $\mathcal{K}$ which preserves graph isomorphisms. Let $K \in \mathcal{K}$, $D \in \mathcal{K}^*$ and $\pi \in \text{Aut}\ (G)$. Since $P_\pi^T = P_{\pi^{-1}}, \pi^{-1} \in \text{Aut}\ (G)$ and $P_{\pi^{-1}}^T K P_{\pi^{-1}} \in \mathcal{K}$

$$\langle K, P_\pi^T D P_\pi \rangle = \langle P_\pi K P_\pi^T, D \rangle = \langle P_{\pi^{-1}}^T K P_{\pi^{-1}}, D \rangle \geq 0,$$

implying $P_\pi^T D P_\pi \in \mathcal{K}^*$. So the dual cone $\mathcal{K}^*$ also preserves graph isomorphisms. In fact we are not aware of any subcone of $\mathcal{S}^n$ applied to the graph coloring or stable set problems which does not preserve graph isomorphisms.[3]
Let cone $\mathcal{K}$ preserve graph isomorphisms. Let matrix $Z \in \mathcal{K}$. Then the cone $\mathcal{K}$ contains its H*-average* defined by

$$r(Z) := \frac{1}{|H|} \sum_{\pi \in H} P_\pi^T Z P_\pi. \tag{30}$$

Lemma 15 *Let cone $\mathcal{K} \subseteq \mathcal{S}^n$ preserve graph isomorphisms. Let $\bar{\mathcal{K}} := \{A \in \mathcal{K} : P_\pi^T A P_\pi = A\ \forall \pi \in H\}$ be its subcone consisting of the H-invariant matrices. Then*

(a) *Map r defined by (30) is a retraction of $\mathcal{K}$ to $\bar{\mathcal{K}}$.*
(b) *Let A be an H-invariant matrix. Let $c(Z) := \langle Z, A \rangle$. Then $c(Z) = c(r(Z))$.*

[2] or is invariant under the action of group H or possesses the symmetry of group H.

[3] In the papers [19,23] larger square matrices are considered. Their rows do not represent only vertices of the graph G but its edges as well. However, if we additionally define $Pe_{ij} = e_{\pi(i)\pi(j)}$, the cones considered in [19,23] satisfy (29), too.

Proof (a) Let H be a group, $\circ$ be the group operation and $\beta \in H$. Then $H \circ \beta = \{\pi \circ \beta | \pi \in H\} = H$.

So

$$P_\beta^T r(Z) P_\beta = \frac{1}{|H|} \sum_{\pi \in H} P_\beta^T P_\pi^T Z P_\pi P_\beta = \frac{1}{|H|} \sum_{\pi \in H} P_{\pi \circ \beta}^T Z P_{\pi \circ \beta} = r(Z), \quad (31)$$

i.e., the image $r(Z)$ is H-invariant. Since any H-invariant matrix is by definition (30) obviously a fixed point of map r, r is indeed the described retraction.
(b) By (30) and since A is H invariant $c(r(Z)) = \langle r(Z), A\rangle = \frac{1}{|H|} \sum_{\pi \in H} \langle Z, P_\pi A P_\pi^T \rangle$ $= \langle Z, A\rangle = c(Z)$. □

Corollary 16 *Let $H \leq Aut(G)$ be a group of automorphisms. Let cone $\mathcal{K}$ preserve graph isomorphisms and satisfy $\mathcal{P} \subseteq \mathcal{K} \subseteq \mathcal{C}$. Then*

$$\vartheta^{\mathcal{K}}(G) = \max\{\langle X, J\rangle : tr\ X = 1, A_G X = 0, X \in \mathcal{K}^*, x_{ij} = x_{\pi(i)\pi(j)} \forall \pi \in H\}, \quad (32)$$

and

$$\Theta^{\mathcal{K}}(G) = \sup\{\langle X - W, J\rangle : tr\ X = 1, A_{\bar{G}} X = 0, X - W \succeq 0, W \in \mathcal{K}, \\ x_{ij} = x_{\pi(i)\pi(j)} \text{ and } w_{ij} = w_{\pi(i)\pi(j)} \forall \pi \in H\}. \quad (33)$$

Proof By definition (23), $\vartheta^{\mathcal{K}}(G) = \max\{\langle X, J\rangle : \text{tr } X = 1, A_G X = 0, X \in \mathcal{K}^*\}$. Matrices J and identity I are obviously group invariant for any permutation group. By Lemma 15 the retraction $r : \mathcal{K} \to \{X \in \mathcal{K} | P_\pi^T X P_\pi = X\ \forall \pi \in H\}$ defined in (30) satisfies $\langle X, J\rangle = \langle r(X), J\rangle$ and $\text{tr } X = \langle X, I\rangle = \langle r(X), I\rangle = \text{tr } r(X)$. If $ij \in E$, then $\{\pi(ij) : \pi \in H\} \subseteq E$. So $A_G X = 0$ implies

$$\langle E_{ij}, r(X)\rangle = \frac{1}{|H|} \sum_{\pi \in H} \langle e_i e_j^T + e_j e_i^T, P_\pi^T X P_\pi \rangle = \frac{1}{|H|} \sum_{\pi \in H} \langle E_{\pi(i)\pi(j)}, X\rangle \\ = \sum 0 = 0,$$

i.e., $A_G\, r(X) = 0$. Also by Lemma 15 r maps $\mathcal{K}^*$ into H-invariant matrices. So r maps the feasible space of (23) onto the feasible space of (32). Now Lemma 12 establishes (32). Likewise (33) is established by applying the retraction r in (26). □

In the case of a highly symmetric vertex transitive graph this observation reduces the number of variables and therefore the space and time complexity, dramatically. We will exploit (33) in Sect. 9.

Lemma 17 *Let H be a vertex transitive group and $M \in \mathcal{S}^n$ be an H-invariant matrix. Then e is an eigenvector of M, and all diagonal elements of M are equal.*

Proof Take any $i, j \in V$. Since $H \leq Aut(G)$ is a vertex transitive group by Definition 10 there is a permutation $\pi \in H$ such that $\pi(j) = i$. Since M H-invariant, by Definition 13 $m_{jk} = m_{\pi(j)\pi(k)} = m_{i\pi(k)}$, i.e., the elements in jth row are just permuted elements from the ith row. Their sum equals $\sum_k m_{jk} = \sum_k m_{ik}$, so $Me = (\sum_i m_{1\pi(i)})e$, and e is an eigenvector. Likewise $z_{jj} = z_{\pi(j)\pi(j)} = z_{ii}$, i.e., all diagonal elements are equal. □

Theorem 18 *Let G be a vertex transitive graph. Let cone $\mathcal{K}$ preserve graph isomorphisms and satisfy $\mathcal{P} \subseteq \mathcal{K} \subseteq \mathcal{C}$. Then $\vartheta^{\mathcal{K}}(G)\Theta^{\mathcal{K}}(G) = n$.*

Proof Take $H := Aut(G)$ in Corollary 16 and let an optimum of (32) be attained at $\bar{X} \in \mathcal{K}^*$. As $\bar{X}$ is H-invariant and since tr $\bar{X} = 1$ by Lemma 17 any diagonal entry $\bar{x}_{ii} = \frac{1}{n}$, and also $\bar{X}e = \lambda e$. Therefore the sum of all entries in $\bar{X}$

$$\vartheta^{\mathcal{K}}(G) = \langle \bar{X}, J \rangle = e^T \bar{X} e = e^T \lambda e = n\lambda.$$

The only nonzero eigenvalue of J is n, $Je = ne$. So vector e is also an eigenvector of $Z := \frac{n}{\lambda}\bar{X} - J$,

$$Ze = \frac{n}{\lambda}\bar{X}e - Je = ne - ne = 0.$$

Next, let v be any eigenvector of $\bar{X}$ orthogonal to e. Then $Jv = 0$ and $Zv = (n/\lambda)\bar{X}v$, i.e., v is eigenvector of Z, too. Therefore $\bar{X} \in \mathcal{K}^* \subseteq \mathcal{P}$ implies $Z \succeq 0$. As $A_G \bar{X} = 0$, there are $t \in \mathbb{R}$ and $y \in \mathbb{R}^{\bar{E}}$ such that $Z = tI + A_{\bar{G}}^T y - J$. So (y, t) is feasible for (24) and $t = \frac{n}{\lambda}\bar{x}_{ii} = \frac{1}{\lambda} = \frac{n}{\vartheta^{\mathcal{K}}(G)}$, so $\Theta^{\mathcal{K}}(G) \leq \frac{n}{\vartheta^{\mathcal{K}}(G)}$. Now Proposition 8 completes the proof. □

By taking $\mathcal{K} = \mathcal{P}$ we see that Theorem 18 generalizes the well-known equality for vertex transitive graphs $\vartheta(G)\vartheta(\bar{G}) = n$, see [18]. We notice that the same technique has already been used in [10] to prove Theorem 18 for the cone $\mathcal{K}_0 = \mathcal{P} + \mathcal{N}$.

Remark 19 Other approaches to bound the stability number also have related approaches to bound the fractional chromatic number. Lasserre's hierarchy of bounds on the stability number $\vartheta(G) = las^{(1)}(G) \geq las^2(G) \geq \cdots \geq \alpha(G)$ introduced in [19] and the corresponding hierarchy of bounds on the chromatic number $\Theta(G) = \psi^{(1)}(G) \leq \psi^{(2)}(G) \leq \cdots \leq \chi_f(G)$ introduced in [15] also satisfy the inequality $las^{(i)}(G)\psi^{(i)}(G) \geq n$ which is again tight on the vertex transitive graphs [15]. The semidefinite programs $las^{(i)}(G)$ and $\psi^{(i)}(G)$ have equal computational complexity. Gvozdenović and Laurent [15] also define an operator which turns any hierarchy toward the stability number into a hierarchy toward the chromatic number where the latter is not bounded above by $\chi_f(G)$. However the effort to compute a bound in this new hierarchy is much larger than for the corresponding bound on the stability number.

7 Cone $\mathcal{K}_1$

Parrilo [26] has introduced a chain of increasingly more complex cones squeezed between $\mathcal{P}$ and $\mathcal{C}$. The simplest of them is $\mathcal{K}_0 = \mathcal{P} + \mathcal{N}$ with $\mathcal{K}_0^* = \mathcal{P} \cap \mathcal{N}$ [26].

So $\vartheta^{\mathcal{K}_0}(G)$ is just Schrijver's number introduced in [30], and $\Theta^{\mathcal{K}_0}(G)$ is the Szegedy number introduced in [32]. Since efficient algorithms for optimizing over the cone $\mathcal{K}_0$ are already known (see [10]), we will in the rest of the paper address optimizing over $\mathcal{K}_1$.

The cone $\mathcal{K}_1$ is characterized by *n supporting matrices* $M^{(1)}, \dots, M^{(n)} \in \mathcal{S}^n$ [1,26]

$$\begin{aligned} M \in \mathcal{K}_1 \iff \exists M^{(1)}, \dots, M^{(n)}\colon M \succeq M^{(i)} \quad &\forall i \in V,\\ m_{ii}^{(i)} = 0 \quad &\forall i \in V,\\ m_{ii}^{(j)} + 2m_{ij}^{(i)} = 0 \quad &\forall i, j \in V,\ i \neq j\\ m_{jk}^{(i)} + m_{ik}^{(j)} + m_{ij}^{(k)} \geq 0 \quad &\forall i, j, k \in V\colon i < j < k. \end{aligned} \tag{34}$$

To gain more insight into this cone we will give a constructive proof that preserves graph isomorphisms.

Proposition 20 *Let* $\pi \in Aut\ (G)$ *and* $M \in \mathcal{K}_1$ *be supported by matrices* $M^{(1)}, \dots, M^{(n)}$. *Then* $P_\pi^T M P_\pi \in \mathcal{K}_1$ *is supported by the matrices* $M^{[i]} := P_\pi^T M^{(\pi(i))} P_\pi$, $i = 1, \dots, n$.

Proof Obviously $P_\pi^T M P_\pi \succeq M^{[i]}$ for any $i \in V$. The other three conditions in (34) regarding $M^{[1]}, \dots, M^{[n]}$ also follow trivially from the corresponding three conditions for $M^{(1)}, \dots, M^{(n)}$ since $m_{bc}^{[a]} = e_b^T P_\pi^T M^{\pi(a)} P_\pi e_c = m_{\pi(b)\pi(c)}^{(\pi(a))}$. For example $m_{ii}^{[i]} = m_{\pi(i)\pi(i)}^{\pi(i)} = 0$. □

Remark 21 Let us note that de Klerk and Pasechnik [6] do not relax (8) but the copositive program $\alpha(G) = \min\{t\colon t(I + A_G^T e) - J \in \mathcal{C}\}$. So their relaxation $\vartheta^{(i)}$, Gvozdenović-Laurent's $\underline{\vartheta}^{(i)}$ [13] and our refinement of the same relaxation $\vartheta^{\mathcal{K}_i}$

$$\vartheta^{(i)}(G) = \inf\{t : tI + A_G^T(te) - J \in \mathcal{K}_i\}, \tag{35}$$

$$\underline{\vartheta}^{(i)}(G) = \inf\{t : tI + A_G^T(se) - J \in \mathcal{K}_i\}, \tag{36}$$

$$\vartheta^{\mathcal{K}_i}(G) = \inf\{t : \ tI + A_G^T y - J \ \in \mathcal{K}_i\} \tag{37}$$

slightly differ. While (35) has only one variable (and is therefore better suited for theoretical considerations) (37) is obviously the tightest as any feasible solution of (35) or (36) is feasible for (37), too. However, as already observed in [6] $\vartheta^{(0)}(G)$ also equals Schrijver number $\vartheta^{\mathcal{K}_0}(G)$. Moreover $\vartheta^{\mathcal{K}_1}(G) = \vartheta^{(1)}(G)$ on the vertex transitive graphs, see [8]. So all three versions of the de Klerk–Pasechnik bound $\vartheta^{(i)}$ coincide on all computationally tractable instances.

Though polynomial-time (semidefinite programming), algorithms for optimizing over cones $\mathcal{K}_1, \mathcal{K}_2, \dots$ are increasingly demanding while their remarkable theoretical properties have been intensively studied [1,6,13,26,27]. Applying copositive programming to optimization problems is only a bit older [2,29]. Consider computing $\Theta^{\mathcal{K}_1}(G)$ on an unstructured (random) graph. Though it is an SDP which can be solved by efficient primal-dual interior point methods notice that its description needs roughly $n^3/2$ variables. Therefore the size of the system matrix is $O(n^6)$, the Cholesky decomposition

of which requires $O(n^9)$ flops. So computing a $\mathcal{K}_1$ bound on a random graph with 100 vertices requires more than 1000G RAM computer.
However the main result of the next section, Lemma 22, states that for computing $\Theta^{\mathcal{K}_1}(G)$ or $\vartheta^{\mathcal{K}_1}(G)$ in the case of a vertex transitive graphs we do not need n supporting matrices $M^{(1)}, \ldots, M^{(n)}$ like in (34) for a general graph. We need only one supporting matrix defined by (38), see (39). This decreases the time complexity to $O(n^6)$. These results apply also to the recently introduced copositive programming formulation of the chromatic number [15] while a similar result for codes is derived in [20,30].

8 Cone $\mathcal{K}_1$ in the case of a vertex transitive graph

Let us restate characterization (34) of $\mathcal{K}_1$. We define

$$\begin{aligned}\hat{\mathcal{K}}^1 := \{&(X^{(0)}, X^{(1)}, \ldots, X^{(n)}) : X^{(0)} \succeq X^{(i)} \forall i,\ X^{(i)} \in \mathcal{S}^n \forall i,\\ &x_{ii}^{(i)} = 0 \forall i,\ \ x_{ii}^{(j)} + 2x_{ij}^{(i)} = 0 \forall i, j, x_{jk}^{(i)} + x_{ik}^{(j)} + x_{ij}^{(k)} \ge 0 \forall i, j, k{:}\, i < j < k\}.\end{aligned}$$

Obviously $X^{(0)} \in \mathcal{K}_1$, if and only if there are $(X^{(0)}, X^{(1)}, \ldots, X^{(n)}) \in \hat{\mathcal{K}}^1$. So we can rewrite any conic program over the cone $\mathcal{K}_1$ into an SDP, e.g.,

$$\min\{c(X^{(0)}) : X^{(0)} \in \mathcal{K}_1\} = \min\{c(X^{(0)}) : (X^{(0)}, \ldots, X^{(n)}) \in \hat{\mathcal{K}}^1\}.$$

Notice the "explosion" of the number of variables as an element of the latter semidefinite cone is determined by $O(n^3)$ real numbers. In order to keep optimizing over $\mathcal{K}_1$ tractable, we need to decrease the number of variables by finding a suitable retraction. In this section we describe such a retraction for any vertex transitive graph G with a known vertex transitive group $H \le \text{Aut}\ (G)$. Note that finding Aut (G) is an isomorphism complete problem. There is however a very efficient heuristic named NAUTY which can find large subgroups of Aut (G) [24].
Motivated by the definition of the retract r in Lemma 12 we consider the mapping $R : (\mathcal{S}^n)^n \to (\mathcal{S}^n)^n$ with the image $(\bar{X}^{(1)}, \ldots, \bar{X}^{(n)}) := R(X^{(1)}, \ldots, X^{(n)})$ componentwise defined by

$$\bar{X}^{(i)} := \frac{1}{|H|} \sum_{\pi \in H} P_\pi^T X^{(\pi(i))} P_\pi. \tag{38}$$

Lemma 22 *Let H be a vertex transitive group of graph G and the map $r \times R : \hat{\mathcal{K}}_1 \to \hat{\mathcal{K}}_1$ defined by (30) and (38). Then*

1. *Image $r \times R(\hat{\mathcal{K}}^1) \subseteq \hat{\mathcal{K}}^1$.*
2. *Image $r \times R(\hat{\mathcal{K}}^1)$ is a retract of the cone $\hat{\mathcal{K}}^1$.*
3. *Image*

$$\begin{aligned}r \times R(\hat{\mathcal{K}}^1) = \{(\bar{X}^{(0)}, \bar{X}^{(1)}, \ldots, \bar{X}^{(n)}) \in \hat{\mathcal{K}}^1{:}\ &P_\pi^T X^{(0)} P_\pi = X^{(0)} \forall \pi \in H,\\ &P_\pi^T \bar{X}^{(\pi(i))} P_\pi = \bar{X}^{(i)} \forall \pi \in H, \forall i > 0\}.\end{aligned}$$

Proof Let $(X^{(0)}, X^{(1)}, \dots, X^{(n)}) \in \mathcal{K}_1$, and $(\bar{X}^{(0)}, \bar{X}^{(1)}, \dots, \bar{X}^{(n)}) := r \times R(X^{(0)}, X^{(1)}, \dots, X^{(n)})$.

1. By Lemma 15 retract $r(\mathcal{K}_1) \subseteq \mathcal{K}_1$. Let $i > 0$. Since $X^{(0)} \succeq X^{(i)}$, $P_\pi^T X^{(0)} P_\pi \succeq P_\pi^T X^{(i)} P_\pi$ implying $\bar{X}^{(0)} \succeq \bar{X}^{(i)}$.
   Next,

$$\bar{x}_{ii}^{(i)} = e_i^T \bar{X}_{ii}^{(i)} e_i = \frac{1}{|H|} \sum_{\pi \in H} e_i P_\pi^T X^{(\pi(i))} P_\pi e_i = \frac{1}{|H|} \sum_{\pi \in H} e_{\pi(i)}^T X^{(\pi(i))} e_{\pi(i)}$$
$$= \frac{1}{|H|} \sum_{\pi \in H} x_{\pi(i)\pi(i)}^{\pi(i)} = \sum 0$$

   by (34). An analog proof,

$$\bar{x}_{ii}^{(j)} + \bar{x}_{ij}^{(i)} = \cdots = \frac{1}{|H|} \sum_{\pi \in H} e_{\pi(i)}^T X^{(\pi(j))} e_{\pi(i)} + 2 e_{\pi(i)}^T X^{(\pi(i))} e_{\pi(j)} = \cdots$$
$$= \sum 0 = 0$$

   establishes the second equality. And similar for inequality.
2. Since r is by Lemma 15 a retract, $r(r(X^{(0)})) = r(X^{(0)})$. To prove that $R(R(X^{(1)}, \dots, X^{(n)})) = R(X^{(1)}, \dots, X^{(n)})$ observe its kth component (by twice applying (38)) equals

$$\frac{1}{|H|} \sum_{\sigma \in H} P_\sigma^T \bar{X}^{(\sigma(k))} P_\sigma = \frac{1}{|H|^2} \sum_{\sigma, \pi \in H} P_\sigma^T P_\pi^T X^{(\pi(\sigma(k)))} P_\pi P_\sigma$$
$$= \frac{1}{|H|^2} \sum_{\sigma, \pi \in H} P_{\pi \circ \sigma}^T X^{(\pi \circ \sigma(k))} P_{\pi \circ \sigma} = \bar{X}^{(k)}$$

   as in the above sum over $\sigma, \pi \in H$ compositums $\pi \circ \sigma$ produce each element of H exactly $|H|$ times.
3. Let $(\bar{X}_0, \bar{X}_1, \dots, \bar{X}_n) \in r \times R(\mathcal{K}_1)$. The equality $P_\pi^T \bar{X}^{(0)} P_\pi = \bar{X}^{(0)}$ was proved in Lemma 15, see (31). Likewise

$$P_\sigma^T \bar{X}^{(\sigma(i))} P_\sigma = \frac{1}{|H|} \sum_{\pi \in H} P_\sigma^T P_\pi^T X^{(\pi(\sigma(i)))} P_\pi P_\sigma$$
$$= \frac{1}{|H|} \sum_{\pi \circ \sigma \in H} P_{\pi \circ \sigma}^T X^{(\pi \circ \sigma(i))} P_{\pi \circ \sigma} = \bar{X}^{(i)}.$$

□

As G is a vertex transitive graph and H its vertex transitive group by applying $\pi_{ij} \in H$ such that $\pi_{ij}(i) = j$ in the above formula we get

$$\bar{X}^{(i)} = P_{\pi_{ij}}^T \bar{X}^{(j)} P_{\pi_{ij}}, \tag{39}$$

Table 1 Computational results on the complements of Hamming graphs

Graph	n	Orbits	$\Theta^{\mathcal{K}_0}(\bar{G})$	$\mathcal{K}_1$-bound	$\chi(\bar{G})$	Times	Seconds
$H(7,6)$	128	120	53.33	63.9	64	2	0
$H(8,6)$	256	165	85.33	127.9	128	22	1
$H(9,4)$	512	220	51.19	53.9		228	4
$H(10,8)$	1024	286	383.99	511.9	512	3290	11
$H(11,10)$	2048	364	921.59	1023.5	1024	46135	19
$H(12,4)$	4096	455	211.86	255.5			37
$H(13,10)$	8192	560	2867.20	4095.5	4096		64
$H(14,4)$	16384	680	614.40	930.8			99
$H(15,4)$	32768	816	1000.72	1846.9			174
$H(16,14)$	65536	969	28086.86	32668.9	32768		250

i.e., all supporting matrices $\bar{X}^{(1)}, \ldots, \bar{X}^{(n)}$ have the same (just permuted) entries. However as the number of vertices in Hamming graphs $H(n, 2)$ grows exponentially with n, for large n this does not suffice to compute the $\mathcal{K}_1$ bounds considered in the next section. In [8] further technical details are given on how to exploit the richness of $Aut(G)$ (i.e., existence of $\pi_{ii} \in H$ such that $\pi_{ii} \neq \text{id}, \pi_{ii}(i) = i$) to further decrease the number of variables as many entries in

$$P_{\pi_{ii}}^T \bar{X}^{(i)} P_{\pi_{ii}} = \bar{X}^{(i)}$$

are equal, i.e., $\bar{x}_{jk}^{(i)} = \bar{x}_{\pi_{ii}(j)\pi_{ii}(k)}^{(i)}$ is in the reduced model (28) only one variable. These simplifications reduce the number of the remaining inequalities and equalities in classification (34), too (e.g., all variables in n equalities $\bar{x}_{11}^{(1)} = 0, \ldots, \bar{x}_n^{(n)} = 0$ are equal thus reducing these n equalities to a single equality $\bar{x}_{11}^{(1)} = 0$). In fact we noticed that each remaining variable appears in exactly one remaining equality or inequality, see [8].

9 Example: Hamming graphs

We tested the algorithm on the binary Hamming graphs (Table 2) and on the complements of such graphs (Table 1). We denote by $H(b, d)$ the binary Hamming graph in which vertices are all $n = 2^b$ strings made of b bits and edges connect strings at Hamming distance (the number of different bits) d. Such graphs have rich automorphism groups, a fact already exploited by Schrijver in [30].
In addition to a vertex transitive group[4] isomorphic to $\mathbb{Z}_2^b$ such a graph also has a rich group of automorphisms isomorphic to the symmetric group S_b.[5] We used the

[4] generated by $\{\pi_i : i = 1, \ldots, b\}$ where π_i switches the ith bit and leaves the other bits unchanged.

[5] consisting of all $n!$ permutations of b bits.

Table 2 Computational results on Hamming graphs

Graph	$\Theta(G)$	$\Theta^{\mathcal{K}_0}(G)$	$\mathcal{K}_1$-bound	$\psi(G)$	$\psi_{\geq 0}(G)$
$H(10, 6)$	6.00	8.73	10.5	10.44	10.89
$H(10, 8)$	2.67	3.20	3.4	$^{+}$3.92	3.92
$H(11, 4)$	16.00	21.57	24.7	$^{+}$25.74	25.74
$H(11, 6)$	12.00	12.00	14.1	12.00	15.28
$H(11, 8)$	3.20	4.94	5.4	$^{+}$5.78	5.78
$H(13, 8)$	5.33	9.41	12.5	12.14	13.65
$H(15, 6)$	27.76	30.74	43.0	$^{+}$46.43	50.30
$H(16, 8)$	16.00	16.00	24.1	16.00	28.44
$H(17, 6)$	35.00	48.22	62.5	$^{+}$86.31	88.32
$H(17, 8)$	18.00	18.00	34.5	32.00	46.51
$H(17, 10)$	6.67	12.63	20.5	15.88	25.84
$H(18, 10)$	10.00	16.00	28.8	18.31	38.88

semidirect product $\mathbb{Z}_2^b \times S_b$ of both subgroups to decrease the number of variables in (26) rewritten as SDP

$$\Theta^{\mathcal{K}_1}(H(b,d)) = \sup\{\langle X - W, J\rangle : \text{tr } X = 1, A_{\bar{G}}X = 0, X - W_0 \succeq 0, (W_0, U_1, \ldots, W_n) \in \hat{\mathcal{K}}^1\} \tag{40}$$

by applying the retraction from Sect. 8, see [8] for technical details. As all matrix variables in SDP (40) possess group symmetry even further reductions described in [7] can be made. (Actually we implemented its implicit version described in [9].) Both reductions have decreased the numerical stability of the algorithm (the faster was less stable, see also [10] or [8]), but the computation times have been reduced substantially. We implemented a primal-dual predictor-corrector interior point method encoded in Matlab with C interfaces, and ran it on a 3 GHz PC running Linux. We stopped the algorithm when numerical instability was encountered. So the numbers in column "$\mathcal{K}_1$-bound" in Tables 1 and 2 are the costs of the last (feasible) dual (40). Even though these numbers are only lower bounds on Θ^{K_1} large improvements over the Szegedy number $\Theta^{\mathcal{K}_0}$ are easily observed.[6]

In Table 1 the times in last column correspond to applying reductions described in [9]. On the other hand, the times in the last-but-one column refer to solving (40) without any further simplifications with the blank entries corresponding to the programs whose variables did not fit into 1G RAM.

As Lasserre's semidefinite program $las^{(2)}(G)$ is a strengthening of de Klerk–Pasechnik's bound on the stability number $\vartheta^{(1)}(G)$ [13] Gvozdenović and Laurent consider the corresponding lower bound on the fractional chromatic number $\psi^{(2)}(G)$ [14]. In Table 2 we compare these relaxations $\psi(G)$ and $\psi_{\geq 0}(G)$, reported in [14],

[6] The Szegedy number $\Theta^{\mathcal{K}_0}(\bar{G})$ happens to be equal to Lovász theta number $\Theta(\bar{G})$ on all graphs in Table 1.

with our bound $\Theta^{\mathcal{K}_1}(G)$. While $\psi_{\geq 0}(G)$ is always stronger than $\Theta^{\mathcal{K}_1}(G)$, the bound $\psi(G)$ is stronger only on the instances marked by +.
Hamming graphs are vertex transitive. So we could interpret these results in terms of bounds $\vartheta^{\mathcal{K}_1}$ on the stability numbers, since $\vartheta^{\mathcal{K}_1}(G) = n/\Theta^{\mathcal{K}_1}(G)$ by Theorem 18 and Proposition 20. The same algorithm could be applied to compute $\vartheta^{\mathcal{K}_1}$ of a vertex transitive graph. For example, $\vartheta^{\mathcal{K}_1}$ of powers of an odd cycle C_k with the automorphism groups $(S_p \times Z_2^p \leq Aut(C_k^p, 1))$ could be used in studying Shannon capacities of C_k. The same holds for $\vartheta^{\mathcal{K}_i}$, $i > 1$, which also involves symmetries.

Acknowledgments We thank Bojan Mohar for inspiring discussion and encouragement to investigate the relation proved in Lemma 3, and Nebojša Gvozdenović and Monique Laurent for communicating to us their proof of Lemma 4 which generalizes our previous result (see [8]). Finally we thank two anonymous referees for their careful reading of earlier versions of the manuscript, resulting in numerous suggestions for improvements, which resulted in the present version of the paper.

References

1. Bomze, I.M., de Klerk, E.: Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. J. Global Optim. **24**, 163–185 (2001)
2. Bomze, I.M., Dür, M., de Klerk, E., Roos, C., Quist, A.J., Terlaky, T.: On copositive programming and standard quadratic optimization problems. J. Global Optim. **18**, 301–320 (2000)
3. Burer, S.: On the copositive representation of binary and continuos nonconvex quadratic Programs. Working paper (2007)
4. Busygin, S., Pasechnik, D.V.: On NP-hardness of the clique partition—independence number gap recognition and related problems. Discrete Math. **306**(4), 460–463 (2006)
5. de Klerk, E.: Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. Kluwer, Dordrecht (2002)
6. de Klerk, E., Pasechnik, D.V.: Approximation of the stability number of a graph via copositive programming. SIAM J. Optim. **12**, 875–892 (2002)
7. de Klerk, E., Pasechnik, D.V., Schrijver, A.: Reduction of symmetric semidefinite programs using the regular ∗-representation. Math. Program. **109**(2–3), 613–624 (2007)
8. Dukanovic, I.: Semidefinite programming applied to graph coloring problem. Ph.D. thesis, University of Klagenfurt, Austria (2008, forthcoming)
9. Dukanovic, I., Rendl, F.: Reductions of group symmetric semidefinite programs. Working paper (2005)
10. Dukanovic, I., Rendl, F.: Semidefinite programming relaxations for graph coloring and maximal clique problems. Math. Program. B **109**(2), 345–366 (2007)
11. Gatermann, K., Parrilo, P.A.: Symmetry groups, semidefinite programs, and sums of squares. J. Pure Appl. Algebra **192**, 95–128 (2004)
12. Gvozdenović, N., Laurent, M.: Personal communication (2005)
13. Gvozdenović, N., Laurent, M.: Semidefinite bounds for the stability number of a graph via sums of squares of polynomials. Lect. Notes Comput. Sci. **3509**, 136–151 (2005)
14. Gvozdenović, N., Laurent, M.: Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization. SIAM J. Optim. (2008, to appear)
15. Gvozdenović, N., Laurent, M.: The operator Ψ for the chromatic number of a graph. SIAM J. Optim. (2008, to appear)
16. Halperin, E., Nathaniel, R., Zwick, U.: Coloring *k*-colorable graphs using relatively small palettes. J. Algorithms **45**, 72–90 (2002)
17. Karger, D., Motwani, R., Sudan, M.: Approximate graph coloring by semidefinite programming. J. Assoc. Comput. Machinery **45**, 246–265 (1998)
18. Knuth, D.E.: The sandwich theorem. Electron. J. Combinatorics **1**, 1–48 (1994)
19. Lasserre, J.B.: On explicit exact SDP relaxation for nonlinear 0-1 programs. In: Ardal, K., Gerards, A.M.H. (eds.) Lecture Notes in Computer Science, vol. 2081, pp. 293–303 (2001)
20. Laurent, M.: Strengthened semidefinite bounds for codes. Preprint, p. 17 (2005)

21. Laurent, M., Rendl, F.: Semidefinite programming and integer programming. In: Weismantel, R., Aardal, K., Nemhauser, G. (eds.) Handbook on Discrete Optimization, pp. 393–514. Elsevier B. V., Amsterdam (2005)
22. Lovász, L.: On the Shannon capacity of a graph. IEEE Trans. Inform. Theory **25**, 1–7 (1979)
23. Lovász, L., Schrijver, A.: Cones of matrices and set-functions and 0-1 optimization. SIAM J. Optim. **1**, 166–190 (1991)
24. McKay, B.D.: Practical graph isomorphism. Congressus Numerantium **30**, 45–87 (1981)
25. Murty, K.G., Kabadi, S.N.: Some NP-complete problems in quadratic and nonlinear programming. Math. Program. **39**, 117–129 (1987)
26. Parrilo, P.A.: Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. Ph.D. thesis, California Institute of Technology (2000)
27. Peña, J., Vera, J., Zuluaga, L.F.: Computing the stability number of a graph via linear and semidefinite programming. SIAM J. Optim. **18**, 87–105 (2007)
28. Povh, J., Rendl, F.: Copositive and semidefinite relaxations of the quadratic assignment problem. Working paper (2006)
29. Quist, A.J., de Klerk, E., Roos, C., Terlaky, T.: Copositive relaxation for general quadratic programming. Optim. Methods Softw. **9**, 185–209 (1998)
30. Schrijver, A.: A comparison of the Delsarte and Lovász bounds. IEEE Trans. Inform. Theory IT **25**, 425–429 (1979)
31. Schrijver, A.: Combinatorial Optimization—Polyhedra and Efficiency, vol. B. Springer, New York (2003)
32. Szegedy, M.: A note on the theta number of Lovász and the generalized Delsarte bound. In: 35th Annual Symposium on Foundations of Computer Science, pp. 36–39 (1994)
33. Todd, M.J.: Semidefinite optimization. Acta Numerica **10**, 515–560 (2001)