

Relaxations of combinatorial problems via association schemes

Etienne de Klerk[‡], Fernando M. de Oliveira Filho[‡], and Dmitrii V. Pasechnik[†]

[‡]Tilburg University, The Netherlands; [†]Nanyang Technological University, Singapore

1.1 Introduction

Semidefinite programming relaxations of combinatorial problems date back to the work of Lovász [17] from 1979, who proposed a semidefinite programming relaxation for the maximum stable set problem which now is known as the Lovász theta number. More recently, Goemans and Williamson [9] showed how to use semidefinite programming to provide an approximation algorithm for the maximum-cut problem; this algorithm achieves the best known approximation ratio for the problem, which is moreover conjectured to be the best possible ratio under the unique games conjecture, a complexity-theoretical assumption (cf. Khot, Kindler, Mossel, and O'Donnell [12]).

The usual approach to obtaining semidefinite programming relaxations of combinatorial problems has been via binary variable reformulations. For example, the convex hull of the set $\{xx^T : x \in \{-1, 1\}^n\}$ is approximated by the convex *elliptope*

$$\{X \in \mathbb{R}^{n \times n} : X_{ii} = 1 \text{ for } i = 1, \dots, n \text{ and } X \succeq 0\},$$

where $X \succeq 0$ means that X is positive semidefinite. Stronger relaxations may be obtained by the use of lift-and-project techniques, which provide hierarchies of tighter and tighter semidefinite programming problems; see e.g. the article by Laurent and Rendl [16] for a description of such methods.

Recently, de Klerk, Pasechnik, and Sotirov [13] presented a different approach to derive semidefinite programming relaxations for the travelling salesman problem; their approach is based on the theory of association schemes.

An *association scheme* is a set $\{A_0, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ of 0–1 matrices that satisfy the following properties:

1. $A_0 = I$ and $A_0 + \dots + A_d = J$, where J is the all-ones matrix;
2. $A_i^T \in \{A_0, \dots, A_d\}$ for $i = 1, \dots, d$;
3. $A_i A_j = A_j A_i$ for $i, j = 0, \dots, d$;

4. $A_i A_j \in \text{span}\{A_0, \dots, A_d\}$ for $i, j = 0, \dots, d$.

We say that the association scheme $\{A_0, \dots, A_d\}$ has d classes. Association schemes where all the matrices A_i are symmetric are called *symmetric association schemes*. In what follows, we will only work with symmetric association schemes and will usually omit the word "symmetric". All the theory can be easily extended to non-symmetric association schemes, though.

In this chapter we show how to obtain semidefinite programming relaxations of the convex hull of an association scheme. To make ideas more precise, let $\{A_0, \dots, A_d\}$ be an association scheme. We show how to derive semidefinite programming relaxations of the set

$$\text{conv}(\{A_0, \dots, A_d\}) = \text{conv}\{(P^\top A_0 P, \dots, P^\top A_d P) : P \in \mathbb{R}^{n \times n} \text{ is a permutation matrix}\}. \quad (1.1)$$

Notice that $\{P^\top A_0 P, \dots, P^\top A_d P\}$, where $P \in \mathbb{R}^{n \times n}$ is a permutation matrix, is an association scheme as well.

Many combinatorial problems can be expressed as the problem of finding a maximum-weight matrix in an association scheme, and so by showing how to obtain relaxations of the convex hull of an association scheme, we provide a unified approach to obtaining semidefinite programming relaxations for many different combinatorial problems.

As an example, consider the maximum bisection problem: We are given a symmetric matrix $W \in \mathbb{R}^{2m \times 2m}$ with nonnegative entries, which we see as edge weights on the complete graph K_{2m} on $2m$ vertices, and we wish to find a maximum-weight copy of the complete bipartite graph $K_{m,m}$ in K_{2m} .

It is a simple matter to check that the $2m \times 2m$ matrices

$$A_0 = I, A_1 = \begin{pmatrix} 0 & J_m \\ J_m & 0 \end{pmatrix}, \text{ and } A_2 = \begin{pmatrix} J_m & 0 \\ 0 & J_m \end{pmatrix}, \quad (1.2)$$

where J_m is the all-ones $m \times m$ matrix, form an association scheme. Notice that A_1 is the adjacency matrix of $K_{m,m}$, and that A_2 is the matrix whose nonzero entries mark pairs of vertices of $K_{m,m}$ that are at distance 2 from each other.

One may now rewrite the maximum bisection problem as follows:

$$\begin{aligned} &\text{maximize } \frac{1}{2} \langle W, X_1 \rangle \\ &(X_0, X_1, X_2) \in \text{conv}(\{A_0, A_1, A_2\}), \end{aligned}$$

where for matrices $X, Y \in \mathbb{R}^{n \times n}$ we write $\langle X, Y \rangle$ for the trace inner product between X and Y .

In an analogous way, one may encode other combinatorial problems as optimization problems over the convex hull of an association scheme, and so by providing a way to derive semidefinite programming relaxations for the convex hull of an association scheme, we provide a way to derive semidefinite programming relaxations for many combinatorial problems. In particular, we

show how to obtain known relaxations for the travelling salesman and the maximum bisection problems, and new relaxations for the cycle covering and maximum p -section problems.

Finally, our approach may be further generalized by considering coherent configurations instead of association schemes. Coherent configurations are defined as association schemes, but there is no requirement for the matrices to commute. Problems such as the maximum (k, l) -cut problem, which cannot be encoded as optimization problems over the convex hull of association schemes, can be seen as optimization problems over the convex hull of coherent configurations.

1.2 Preliminaries and notation

All the graphs considered in this chapter are *simple*: They have no loops and no parallel edges.

We denote by A^\top the transpose of a matrix A . For complex matrices, we denote by A^* the conjugate transpose of A . By $e \in \mathbb{R}^n$ we denote the all-ones vector.

When we say that a real matrix is positive semidefinite, we mean that it is symmetric and has nonnegative eigenvalues. If a complex matrix is positive semidefinite, then it is Hermitian and has nonnegative eigenvalues. We often use the notation $A \succeq 0$ to denote that A is positive semidefinite.

For matrices A and $B \in \mathbb{R}^{n \times n}$, we denote by $\langle A, B \rangle$ the *trace inner product* between A and B , that is

$$\langle A, B \rangle = \text{trace}(B^\top A) = \sum_{i,j=1}^n A_{ij} B_{ij}.$$

When $A, B \in \mathbb{C}^{n \times n}$, the trace inner product is given by

$$\langle A, B \rangle = \text{trace}(B^* A) = \sum_{i,j=1}^n A_{ij} \overline{B_{ij}}.$$

We also need some basic properties of tensor products of vector spaces. For background on tensor products, see the book by Greub [10].

If U and V are vector spaces over a same field, we denote by $U \otimes V$ the tensor product of U and V . If $u \in U$ and $v \in V$, then $u \otimes v$ is an element of $U \otimes V$. The space $U \otimes V$ is spanned by elements of the form $u \otimes v$ for $u \in U$ and $v \in V$.

When $U = \mathbb{R}^{m_1 \times n_1}$ and $V = \mathbb{R}^{m_2 \times n_2}$, we identify $U \otimes V$ with the space $\mathbb{R}^{m_1 m_2 \times n_1 n_2}$, and for $A \in \mathbb{R}^{m_1 \times n_1}$ and $B \in \mathbb{R}^{m_2 \times n_2}$ we identify $A \otimes B$ with the *Kronecker product* between matrices A and B . This is the $m_1 \times n_1$ block matrix with blocks of size $m_2 \times n_2$, block (i, j) being equal to $A_{ij} B$.

So $A \otimes B$ is an $m_1 m_2 \times n_1 n_2$ matrix. The same definitions hold when \mathbb{R} is replaced by \mathbb{C} .

One property of the tensor product of matrices that we will use relates to the trace inner product. Namely, for $A_1, A_2 \in \mathbb{R}^{m \times m}$ and $B_1, B_2 \in \mathbb{R}^{n \times n}$ we have

$$\langle A_1 \otimes B_1, A_2 \otimes B_2 \rangle = \langle A_1, A_2 \rangle \langle B_1, B_2 \rangle.$$

The same holds when \mathbb{R} is replaced by \mathbb{C} , of course.

Finally, we denote by δ_{ij} the *Kronecker delta function*, which is equal to 1 if $i = j$, and equal to 0 otherwise.

1.3 Association schemes

We defined association schemes in the introduction. Now we give a brief overview of the relevant properties of association schemes that will be of use to us; for background on association schemes we refer the reader to Chapter 12 of the book by Godsil [7] or Section 3.3 in the book by Cameron [3].

1.3.1 Sources of association schemes

In the introduction we used an association scheme that was related to the complete bipartite graph $K_{m,m}$. This idea can be generalized: Association schemes can be obtained from some special classes of graphs. In this section, we discuss how to obtain association schemes from distance-regular graphs; these schemes will be important in Section 1.5 in the encoding of combinatorial problems as problems over the convex hull of association schemes.

Let $G = (V, E)$ be a graph of diameter d . We say that G is *distance-regular* if for any vertices $x, y \in V$ and any numbers $k, l = 0, \dots, d$, the number of vertices at distance k from x and l from y depends only on k, l , and the distance between x and y , not depending therefore on the actual choice of x and y . A distance-regular graph of diameter 2 is called *strongly regular*.

Now, say G is a distance-regular graph of diameter d and label its vertices $1, \dots, n$. Then the $n \times n$ matrices A_0, \dots, A_d such that

$$A_i(x, y) = \begin{cases} 1 & \text{if } \text{dist}_G(x, y) = i, \\ 0 & \text{otherwise} \end{cases}$$

form an association scheme with d classes (cf. Cameron [3], Theorem 3.6), to which we refer as the association scheme of the graph G . The matrices A_i above are the *distance matrices* of the graph G , and A_1 is just the adjacency matrix of G . The association scheme we used in the introduction to encode the maximum bisection problem was the association scheme of $K_{m,m}$.

So from distance-regular graphs one may obtain association schemes. Examples of distance-regular graphs are the complete bipartite graphs $K_{m,m}$ and the cycles.

1.3.2 Eigenvalues of association schemes

Let $\{A_0, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be an association scheme. The span of the matrices A_0, \dots, A_d is a matrix $*$ -algebra, that is, it is a subspace of $\mathbb{C}^{n \times n}$ closed under matrix multiplication and taking complex conjugate transposes. This algebra is the so-called *Bose-Mesner algebra* of the association scheme, and A_0, \dots, A_d is an orthogonal basis of this algebra. As we observed in the introduction, if $P \in \mathbb{R}^{n \times n}$ is a permutation matrix, then $\{P^\top A_0 P, \dots, P^\top A_d P\}$ is also an association scheme, and its Bose-Mesner algebra is isomorphic as an algebra to that of $\{A_0, \dots, A_d\}$.

The Bose-Mesner algebra of an association scheme is commutative, and therefore it can be diagonalized, that is, there is a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that the matrices

$$U^* A_0 U, \dots, U^* A_d U$$

are diagonal. This result implies that the algebra has a basis of idempotents, as we state below.

Theorem 1. *The Bose-Mesner algebra of an association scheme $\{A_0, \dots, A_d\}$ has an orthonormal basis E_0, \dots, E_d of Hermitian idempotents (that is, E_i is Hermitian and $E_i^2 = E_i$ for $i = 0, \dots, d$). Moreover, $E_0 = \frac{1}{n}J$ and $E_0 + \dots + E_d = I$.*

Note that the matrices E_i are positive semidefinite, since they are Hermitian and have zero-one eigenvalues.

We may write the basis A_0, \dots, A_d of the Bose-Mesner algebra in terms of the basis E_0, \dots, E_d and vice versa. So there are constants p_{ij} for $i, j = 0, \dots, d$ such that

$$A_j = \sum_{i=0}^d p_{ij} E_i$$

for $j = 0, \dots, d$. Notice that p_{ij} is actually the i -th eigenvalue of the matrix A_j . The constants p_{ij} are the *eigenvalues* of the association scheme $\{A_0, \dots, A_d\}$.

We may also write the basis E_0, \dots, E_d in terms of A_0, \dots, A_d . So there are constants q_{ij} for $i, j = 0, \dots, d$, the *dual eigenvalues* of the association scheme, such that

$$E_j = \frac{1}{n} \sum_{i=0}^d q_{ij} A_i \tag{1.3}$$

for $j = 0, \dots, d$. Our normalization above ensures that $q_{i0} = 1$ for all $i = 0, \dots, d$.

It is customary to define primal and dual matrices of eigenvalues of the association scheme as follows:

$$P = (p_{ij})_{i,j=0}^d \quad \text{and} \quad Q = (q_{ij})_{i,j=0}^d.$$

It is easy to check that $PQ = nI$.

Since we work with symmetric association schemes, we have that the numbers p_{ij} and q_{ij} are all real. Moreover, we have the identity

$$\frac{p_{ji}}{m_i} = \frac{q_{ij}}{n \operatorname{trace} E_j}. \quad (1.4)$$

Finally, notice that if $P \in \mathbb{R}^{n \times n}$ is a permutation matrix, then the eigenvalues and dual eigenvalues of both $\{A_0, \dots, A_d\}$ and $\{P^\top A_0 P, \dots, P^\top A_d P\}$ are the same.

There is a closed formula for the eigenvalues and dual eigenvalues of the association scheme of a strongly regular graph. If A is the adjacency matrix of a strongly regular graph, then e is an eigenvector of A , since a strongly regular graph is regular. Eigenvalues associated with eigenvectors orthogonal to e are said to be the *restricted* eigenvalues of the graph. Every strongly regular graph has exactly two distinct restricted eigenvalues. We then have the following result.

Theorem 2. *Let G be a strongly regular graph with n vertices and of degree k having restricted eigenvalues r and s with $r > s$. Suppose that the eigenvalue r occurs with multiplicity f and that s occurs with multiplicity g . Then*

$$P = \begin{pmatrix} 1 & k & |V| - k - 1 \\ 1 & r & -r - 1 \\ 1 & s & -s - 1 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 1 & f & g \\ 1 & fr/k & gs/k \\ 1 - \frac{f(r+1)}{|V|-k-1} & -\frac{g(s+1)}{|V|-k-1} \end{pmatrix}.$$

1.4 Semidefinite programming relaxations of association schemes

Let $\{A_0, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be an association scheme. We now present a semidefinite programming relaxation of its convex hull $\operatorname{conv}(\{A_0, \dots, A_d\})$, which we defined in (1.1).

Recall that we have expression (1.3) for the dual eigenvalues of an association scheme, and also that each E_i is positive semidefinite. Since the eigenvalues and dual eigenvalues of an association scheme are the same as the eigenvalues and dual eigenvalues of any permuted version of the association scheme, we see that any $(X_0, \dots, X_d) \in \operatorname{conv}(\{A_0, \dots, A_d\})$ satisfies

$$\begin{aligned} \sum_{i=0}^d X_i &= J, \\ \sum_{i=0}^d q_{ij} X_i &\succeq 0 \quad \text{for } j = 1, \dots, d, \\ X_0 &= I, X_i \text{ nonnegative and symmetric for } i = 1, \dots, d. \end{aligned} \quad (1.5)$$

The theorem below shows that many important constraints are already implied by (1.5).

Theorem 3. Suppose X_0, \dots, X_d satisfy (1.5). Write

$$Y_j = \frac{1}{n} \sum_{i=0}^d q_{ij} X_i.$$

Then the following holds:

1. $\sum_{j=0}^d Y_j = I$;
2. $Y_j e = 0$ for $j = 0, \dots, d$;
3. $\sum_{i=0}^d p_{ij} Y_i = X_j$ for $j = 0, \dots, d$;
4. $X_j e = A_j e = p_{0j} e$ for $j = 0, \dots, d$.

Proof. Direct from (1.5) and the properties of the matrices of eigenvalues and dual eigenvalues of the association scheme. \square

Also a whole class of linear matrix inequalities is implied by (1.5), as we show next.

Theorem 4. Let $\{A_0, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be an association scheme and suppose that the linear matrix inequality

$$\sum_{i=0}^d \alpha_i X_i \succeq 0, \tag{1.6}$$

where $\alpha_0, \dots, \alpha_d$ are given scalars, is satisfied by $X_i = A_i$. Then (1.6) may be written as a conic combination of the linear matrix inequalities

$$\sum_{i=0}^d q_{ij} X_i \succeq 0 \quad \text{for } j = 0, \dots, d.$$

Proof. We have

$$Y = \sum_{i=0}^d \alpha_i A_i \succeq 0.$$

But then since Y is in the Bose-Mesner algebra of the association scheme, there exist nonnegative numbers β_0, \dots, β_d such that

$$Y = \sum_{j=0}^d \beta_j E_j,$$

where E_0, \dots, E_d is the idempotent basis of the algebra. (Notice that, in fact, the numbers β_j are the eigenvalues of Y .) Substituting

$$E_j = \frac{1}{n} \sum_{i=0}^d q_{ij} A_i$$

yields

$$\sum_{i=0}^d \alpha_i A_i = \sum_{j=0}^d \frac{\beta_j}{n} \left(\sum_{i=0}^d q_{ij} A_i \right),$$

and since the A_i are pairwise orthogonal we see that

$$\alpha_i = \sum_{j=0}^d \frac{\beta_j}{n} q_{ij},$$

as we wanted. \square

1.5 Semidefinite programming relaxations of combinatorial problems

We now show how to derive semidefinite programming relaxations of many different combinatorial problems from the relaxation for the convex hull of an association scheme which we presented in the previous section.

The basic idea is to restate the combinatorial problem as the problem of finding a maximum/minimum-weight distance-regular subgraph of a weighted complete graph, and then use the fact that the distance matrices of a distance-regular graph give rise to an association scheme. For example, the travelling salesman problem is the problem of finding a minimum-weight Hamiltonian cycle in a graph, and the maximum bisection problem is the problem of finding a maximum-weight copy of $K_{m,m}$ in the complete graph K_{2m} .

More precisely, suppose we are given a symmetric matrix $W \in \mathbb{R}^{n \times n}$ that we see as giving weights to the edges of the complete graph K_n , and we are also given a distance-regular graph H on n vertices. Our problem is to find a maximum-weight copy of H in K_n .

Say H has diameter d , and let A_0, \dots, A_d be its distance matrices, so that A_1 is the adjacency matrix of H . Since H is distance-regular, $\{A_0, \dots, A_d\}$ is an association scheme. Our problem can be equivalently stated as the following *quadratic assignment problem* (QAP):

$$\begin{aligned} &\text{maximize } \frac{1}{2} \langle W, P^\top A_1 P \rangle \\ &P \in \mathbb{R}^{n \times n} \text{ is a permutation matrix.} \end{aligned}$$

The problem above can be equivalently stated in terms of the convex hull of the association scheme $\{A_0, \dots, A_d\}$:

$$\begin{aligned} &\text{maximize } \frac{1}{2} \langle W, X_1 \rangle \\ &(X_0, \dots, X_d) \in \text{conv}(\{A_0, \dots, A_d\}). \end{aligned}$$

Finally, one may consider relaxation (1.5) for the above problem, obtaining the following semidefinite programming relaxation for it:

$$\begin{aligned}
& \text{maximize } \frac{1}{2} \langle W, X_1 \rangle \\
& \sum_{i=0}^d X_i = J, \\
& \sum_{i=0}^d q_{ij} X_i \succeq 0 \quad \text{for } j = 1, \dots, d, \\
& X_0 = I, X_i \text{ nonnegative and symmetric for } i = 1, \dots, d.
\end{aligned}$$

This is the approach we employ in the following sections.

1.5.1 The travelling salesman problem and the k -cycle covering problem

We are given a symmetric nonnegative matrix $W \in \mathbb{R}^{n \times n}$, which we view as giving weights to the edges of the complete graph K_n . The travelling salesman problem asks us to find a minimum-weight copy of C_n , the cycle on n vertices, in the weighted graph K_n .

Now, C_n is a distance-regular graph, and its association scheme has $\lfloor n/2 \rfloor$ classes and is known as the *Lee scheme*, which is also the association scheme of symmetric circulant matrices. The dual eigenvalues of the association scheme of C_n are given by

$$q_{ij} = 2 \cos(2ij\pi/n) \quad \text{for } i, j = 1, \dots, \lfloor n/2 \rfloor.$$

Moreover, $q_{i0} = 1$ for $i = 0, \dots, \lfloor n/2 \rfloor$ and $q_{0j} = 2$ for $j = 1, \dots, \lfloor n/2 \rfloor$.

So, if we consider relaxation (1.5) for the convex hull of the Lee scheme, we obtain the following semidefinite programming relaxation for the travelling salesman problem:

$$\begin{aligned}
& \text{minimize } \frac{1}{2} \langle W, X_1 \rangle \\
& I + \sum_{i=1}^d X_i = J, \\
& I + \sum_{i=1}^d \cos(2ij\pi/n) X_i \succeq 0 \quad \text{for } j = 1, \dots, \lfloor n/2 \rfloor, \\
& X_i \text{ nonnegative and symmetric for } i = 1, \dots, \lfloor n/2 \rfloor.
\end{aligned} \tag{1.7}$$

This is precisely the semidefinite programming relaxation for the travelling salesman problem introduced by de Klerk, Pasechnik, and Sotirov [13]. One may use Theorem 4 to show that this relaxation is tighter than the one given by Cvetković, Cangalović, and Kovačević-Vujčić [2]. This fact was already known, but the proof can be greatly simplified by using the more general Theorem 4.

A simple modification in the objective function of (1.7) allows us to provide a relaxation for the minimum k -cycle covering problem. This is the problem of partitioning a weighted complete graph into k vertex-disjoint cycles of equal length so that the total weight of the cycles is minimized.

This problem has been studied by Goemans and Williamson [8] (see also Manthey [18]), who showed that a 4-approximation algorithm exists when the weights satisfy the triangle inequality.

Now suppose n is a multiple of k . One may check that the matrix A_k , the k -th distance matrix of the cycle C_n , is actually the incidence matrix of k

vertex-disjoint cycles of length n/k each. So, to obtain a relaxation of the k -cycle covering problem, one only has to replace the objective function of (1.7) by $\frac{1}{2}\langle W, X_k \rangle$.

1.5.2 The maximum bisection problem

We already considered the maximum bisection problem in the introduction. The problem consists of, given a nonnegative symmetric matrix $W \in \mathbb{R}^{2m \times 2m}$, which we see as giving weights to the edges of the complete graph K_{2m} , finding a maximum-weight copy of the complete bipartite graph $K_{m,m}$ in K_{2m} .

The graph $K_{m,m}$ is a strongly regular graph; the association scheme induced by it was described in (1.2). Since $K_{m,m}$ is strongly regular, its dual eigenvalues have a closed formula, as given in Theorem 2. Indeed, the restricted eigenvalues of $K_{m,m}$ are the eigenvalues $r = 0$ and $s = -m$ of A_1 , and so the matrices of eigenvalues and dual eigenvalues are given by

$$P = \begin{pmatrix} 1 & m & m-1 \\ 1 & 0 & -1 \\ 1 & -m & m-1 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 1 & 2(m-1) & 1 \\ 1 & 0 & -1 \\ 1 & -2 & 1 \end{pmatrix}.$$

Now, by adapting (1.5), we obtain the following relaxation of the maximum bisection problem:

$$\begin{aligned} &\text{maximize } \frac{1}{2}\langle W, X_1 \rangle \\ &\quad I + X_1 + X_2 = J, \\ &\quad (m-1)I - X_2 \succeq 0, \\ &\quad I - X_1 + X_2 \succeq 0, \\ &\quad X_1, X_2 \text{ nonnegative and symmetric.} \end{aligned} \tag{1.8}$$

Frieze and Jerrum [6] (see also Ye [24]) considered another semidefinite programming relaxation of the maximum bisection problem. Their relaxation is the following:

$$\begin{aligned} &\text{maximize } \frac{1}{4}\langle W, J - X \rangle \\ &\quad X_{ii} = 1 \text{ for } i = 1, \dots, 2m, \\ &\quad Xe = 0, \\ &\quad X \succeq 0. \end{aligned} \tag{1.9}$$

To see that this is a relaxation of the problem, notice that $X = vv^\top$ is a feasible solution, where $v \in \{-1, 1\}^{2m}$ gives a bisection of the vertex set.

Actually, the relaxation we give is equivalent to the one given by Frieze and Jerrum, as we show in the next theorem.

Theorem 5. *The optimal values of problems (1.8) and (1.9) coincide.*

Proof. Given a feasible solution X_1, X_2 of (1.8), set

$$X = I - X_1 + X_2 \succeq 0.$$

Using (4) of Theorem 3 we have that

$$Xe = e - X_1e + X_2e = e - me + (m-1)e = 0,$$

where we have used that $A_1e = me$ and $A_2e = (m-1)e$ for the association scheme of $K_{m,m}$. It is also easy to verify that the diagonal entries of X are all equal to 1 and that

$$\frac{1}{4}\langle W, J - X \rangle = \frac{1}{2}\langle W, X_1 \rangle.$$

Conversely, suppose X is a feasible solution of (1.9). We claim that

$$X_1 = \frac{1}{2}(J - X) \quad \text{and} \quad X_2 = \frac{1}{2}(J + X) - I$$

is a feasible solution of (1.8) with the same objective value.

Indeed, it is easy to see that X_1, X_2 has the same objective value as X . To see that it is a feasible solution of (1.8), notice that since X is positive semidefinite and has diagonal entries equal to 1, all entries of X have absolute value at most 1, and therefore both X_1 and X_2 are nonnegative. Moreover, it is easy to see that X_1 and X_2 are symmetric, and that $I + X_1 + X_2 = J$ and $I - X_1 + X_2 \succeq 0$. So we argue that $(m-1)I - X_2 \succeq 0$.

Indeed, $(m-1)I - X_2 = mI - \frac{1}{2}(J + X)$. Now, the matrix $M = \frac{1}{2}(J + X)$ is positive semidefinite, and e is one of its eigenvectors, with associated eigenvalue m . Since M has trace $2m$ and is positive semidefinite, this implies that any other eigenvalue of M is at most m . So it follows that $mI - M \succeq 0$, as we wanted. \square

1.5.3 The maximum p -section problem

Given a nonnegative symmetric matrix $W \in \mathbb{R}^{pm \times pm}$, which we see as giving weights to the edges of the complete graph K_{pm} , our goal is to find a maximum-weight copy of the complete p -partite graph in K_{pm} .

The complete p -partite graph is strongly regular, and it generates an association scheme whose matrix Q of dual eigenvalues is

$$Q = \begin{pmatrix} 1 & p(m-1) & p-1 \\ 1 & 0 & -1 \\ 1 & -p & p-1 \end{pmatrix}.$$

Then relaxation (1.5) simplifies to

$$\begin{aligned}
& \text{maximize } \frac{1}{2} \langle W, X_1 \rangle \\
& I + X_1 + X_2 = J, \\
& (m-1)I - X_2 \succeq 0, \\
& (p-1)I - X_1 + (p-1)X_2 \succeq 0, \\
& X_1, X_2 \text{ nonnegative and symmetric.}
\end{aligned}$$

Note that this coincides with relaxation (1.8) for the maximum bisection problem when $p = 2$. A different semidefinite programming relaxation for the maximum p -section problem was proposed by Andersson [1].

1.6 Semidefinite programming relaxations of coherent configurations

A *coherent configuration* is a set $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ of 0–1 matrices with the following properties:

1. There is a set $N \subseteq \{1, \dots, d\}$ such that $\sum_{i \in N} A_i = I$ and $A_1 + \dots + A_d = J$;
2. $A_i^\top \in \{A_1, \dots, A_d\}$ for $i = 1, \dots, d$;
3. $A_i A_j \in \text{span}\{A_1, \dots, A_d\}$ for $i, j = 1, \dots, d$.

Thus, association schemes are commutative coherent configurations.

Let $\{A_1, \dots, A_d\}$ be a coherent configuration. For $i \in \{1, \dots, d\}$ we let $i^* \in \{1, \dots, d\}$ be such that $A_{i^*} = A_i^\top$. We also write

$$m_i = \langle J, A_i \rangle = \langle A_i, A_i \rangle$$

for $i = 1, \dots, d$.

Matrices A_1, \dots, A_d generate a matrix $*$ -algebra containing the identity, which however need not be commutative, and therefore it might not be possible to diagonalize this algebra. It is always possible to block-diagonalize the algebra of a coherent configuration, however, as was shown by Wedderburn [23]. To precisely state Wedderburn's result we first introduce some notation.

Let $\mathcal{A}, \mathcal{B} \subseteq \mathbb{C}^{n \times n}$. We write

$$\mathcal{A} \oplus \mathcal{B} = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} : A \in \mathcal{A} \text{ and } B \in \mathcal{B} \right\}.$$

For $\mathcal{A} \subseteq \mathbb{C}^{n \times n}$ and for any positive integer r we write

$$r \odot \mathcal{A} = \{I_r \otimes A : A \in \mathcal{A}\},$$

where I_r is the $r \times r$ identity matrix, so that $I_r \otimes A$ is an $rn \times rn$ matrix with r repeated diagonal blocks equal to A and zeros everywhere else. Finally, let $U \in \mathbb{C}^{n \times n}$. Then for $\mathcal{A} \subseteq \mathbb{C}^{n \times n}$ we write

$$U^* \mathcal{A} U = \{U^* A U : A \in \mathcal{A}\}.$$

Wedderburn has shown the following theorem.

Theorem 6. *Let $\mathcal{A} \subseteq \mathbb{C}^{n \times n}$ be a matrix $*$ -algebra containing the identity. Then there is a unitary matrix $U \in \mathbb{C}^{n \times n}$ and positive integers r_1, \dots, r_k and s_1, \dots, s_k such that*

$$U^* \mathcal{A} U = \bigoplus_{i=1}^k r_i \odot \mathbb{C}^{s_i \times s_i}.$$

Notice that, if $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ is a coherent configuration and $P \in \mathbb{R}^{n \times n}$ is a permutation matrix, then also $\{P^T A_1 P, \dots, P^T A_d P\}$ is a coherent configuration. Our goal is to provide a semidefinite programming relaxation for the convex hull of a coherent configuration, which is the set

$$\text{conv}(\{A_1, \dots, A_d\}) = \text{conv}\{(P^T A_1 P, \dots, P^T A_d P) : P \in \mathbb{R}^{n \times n} \text{ is a permutation matrix}\}.$$

To provide a relaxation for this set, we use the following theorem, which describes a linear matrix inequality satisfied by a coherent configuration.

Theorem 7. *Let $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be a coherent configuration. Then*

$$\sum_{i=1}^d m_i^{-1} A_i \otimes A_i \succeq 0. \quad (1.10)$$

We prove the theorem in a moment; first we use it to describe our semidefinite programming relaxation of the convex hull of a coherent configuration. So let $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be a coherent configuration. From Theorem 7, it is clear that for $(X_1, \dots, X_d) \in \text{conv}(\{A_1, \dots, A_d\})$ we have

$$\sum_{i=1}^d m_i^{-1} A_i \otimes X_i \succeq 0. \quad (1.11)$$

Let $N \subseteq \{1, \dots, d\}$ be such that $\sum_{i \in N} A_i = I$. Then any $(X_1, \dots, X_d) \in \text{conv}(\{A_1, \dots, A_d\})$ satisfies

$$\begin{aligned} \sum_{i \in N} X_i &= I, \\ \sum_{i=1}^d X_i &= J, \\ \sum_{i=1}^d m_i^{-1} A_i \otimes X_i &\succeq 0, \\ \langle J, X_i \rangle &= m_i, \quad X_i^* = X_i^T, \text{ and } X_i \geq 0 \text{ for } i = 1, \dots, d. \end{aligned} \quad (1.12)$$

We observe that, in the linear matrix inequality (1.11), one may replace the matrices A_i from the coherent configuration by their block-diagonalizations, obtaining an equivalent constraint. Also, repeated blocks may be eliminated: Only one copy of each set of repeated blocks is necessary. So, from Wedderburn's theorem we see that in constraint (1.11) it is possible to replace the matrices A_i , which are $n \times n$ matrices, by matrices whose dimensions depend

only on the dimension of the algebra generated by A_1, \dots, A_d , which is equal to d .

In an analogous way to Theorem 4, the linear matrix inequality (1.11) implies a whole class of linear matrix inequalities that are valid for the convex hull of a coherent configuration, as we will see after the proof of Theorem 7.

To prove Theorem 7 we need the following theorem that holds in the more general context of C^* -algebras. We give a proof here for the case of matrix $*$ -algebras to make our presentation self-contained.

Theorem 8. *Let $\mathcal{A} \subseteq \mathbb{C}^{n \times n}$ be a matrix $*$ -algebra containing the identity and let $X \in \mathbb{C}^{n \times n}$ be a positive semidefinite matrix. Then the orthogonal projection of X onto \mathcal{A} is also positive semidefinite.*

Proof. From Wedderburn's theorem (Theorem 6) we may assume that

$$\mathcal{A} = \bigoplus_{i=1}^k r_i \odot \mathbb{C}^{s_i \times s_i}. \quad (1.13)$$

For $i = 1, \dots, k$ and $j = 1, \dots, r_i$, let X_{ij} be the $s_i \times s_i$ diagonal block of X corresponding to the j -th copy of $\mathbb{C}^{s_i \times s_i}$ in the decomposition of \mathcal{A} given in (1.13).

Since X is positive semidefinite, every matrix X_{ij} as defined above is also positive semidefinite. Moreover, the projection of X onto \mathcal{A} is explicitly given as

$$\bigoplus_{i=1}^k r_i \odot r_i^{-1} \sum_{j=1}^{r_i} X_{ij},$$

which is then also seen to be positive semidefinite. \square

We now can give a proof of Theorem 7.

Proof (Proof of Theorem 7). Consider the matrix

$$X = \sum_{k,l=1}^n E_{kl} \otimes E_{kl},$$

where E_{kl} is the $n \times n$ matrix with 1 in position (k, l) and 0 everywhere else.

We claim that X is positive semidefinite. To see this, let e_1, \dots, e_n be the canonical orthonormal basis of \mathbb{R}^n . Then $e_i \otimes e_j$ for $i, j = 1, \dots, n$ is a basis of $\mathbb{R}^n \otimes \mathbb{R}^n$.

Since $E_{kl}e_i = \delta_{li}e_k$, if $i \neq j$, then $X(e_i \otimes e_j) = 0$. So all vectors $e_i \otimes e_j$ with $i \neq j$ are in the kernel of X .

On the other hand, for $i = 1, \dots, n$ we have that

$$X(e_i \otimes e_i) = \sum_{k=1}^n e_k \otimes e_k.$$

Then it is clear that also the vectors $e_1 \otimes e_1 - e_i \otimes e_i$, for $i = 2, \dots, n$, lie in the kernel of X . Now, from the above identity we see that $e_1 \otimes e_1 + \dots + e_n \otimes e_n$ is an eigenvector of X with associated eigenvalue n , and hence it follows that X is positive semidefinite.

Now notice that, since $\{A_1, \dots, A_d\}$ is a coherent configuration, the matrices

$$(m_i m_j)^{-1/2} A_i \otimes A_j \quad \text{for } i, j = 1, \dots, d \quad (1.14)$$

span a matrix $*$ -algebra which moreover contains the identity. Actually, the matrices above form an orthonormal basis of this algebra.

By Theorem 8, the projection of X onto the algebra spanned by the matrices in (1.14) is a positive semidefinite matrix. We show now that this projection is exactly the matrix in (1.10). Indeed, the projection is explicitly given as

$$\begin{aligned} & \sum_{i,j=1}^d (m_i m_j)^{-1} A_i \otimes A_j \left\langle \sum_{k,l=1}^d E_{kl} \otimes E_{kl}, A_i \otimes A_j \right\rangle \\ &= \sum_{i,j=1}^d (m_i m_j)^{-1} A_i \otimes A_j \sum_{k,l=1}^d \langle E_{kl}, A_i \rangle \langle E_{kl}, A_j \rangle. \end{aligned}$$

Notice that, for $i, j \in \{1, \dots, d\}$ with $i \neq j$ we have $\langle E_{kl}, A_i \rangle \langle E_{kl}, A_j \rangle = 0$, because the matrices A_1, \dots, A_d have disjoint supports. So the above expression simplifies to

$$\sum_{i=1}^d m_i^{-2} A_i \otimes A_i \sum_{k,l=1}^n \langle E_{kl}, A_i \rangle^2 = \sum_{i=1}^d m_i^{-1} A_i \otimes A_i,$$

which is exactly the matrix in (1.10), as we wanted. \square

Theorem 8 also suggests a whole class of linear matrix inequalities that are valid for the convex hull of a coherent configuration. Indeed, let $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be a coherent configuration. Given a positive semidefinite matrix $Y \in \mathbb{R}^{n \times n}$, Theorem 8 implies that the projection of Y onto the algebra spanned by A_1, \dots, A_d is positive semidefinite, that is

$$\sum_{i=1}^d m_i^{-1} \langle Y, A_i \rangle A_i \succeq 0.$$

So we see that any $(X_1, \dots, X_d) \in \text{conv}(\{A_1, \dots, A_d\})$ satisfies

$$\sum_{i=1}^d m_i^{-1} \langle Y, A_i \rangle X_i \succeq 0.$$

All these infinitely many linear matrix inequalities are already implied by (1.11), however, as we show next.

Theorem 9. Let $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be a coherent configuration. Suppose matrices $X_1, \dots, X_d \in \mathbb{R}^{n \times n}$ which are such that $X_{i^*} = X_i^\top$ for $i = 1, \dots, d$ satisfy (1.11). Then we also have

$$\sum_{i=1}^d m_i^{-1} \langle Y, A_i \rangle X_i \succeq 0, \quad (1.15)$$

where $Y \in \mathbb{R}^{n \times n}$ is any positive semidefinite matrix.

Proof. Let M be the matrix in (1.15). It is clear that M is symmetric. To see that it is actually positive semidefinite, let $Z \in \mathbb{R}^{n \times n}$ be a positive semidefinite matrix. We show that $\langle Z, M \rangle \geq 0$.

Indeed, we have

$$\langle Z, M \rangle = \sum_{i=1}^d m_i^{-1} \langle Y, A_i \rangle \langle Z, X_i \rangle.$$

Now, notice that the matrix $Y \otimes Z$ is positive semidefinite, since both Y and Z are positive semidefinite. Then, since X_1, \dots, X_d satisfy (1.11) we have that

$$\begin{aligned} 0 &\leq \left\langle Y \otimes Z, \sum_{i=1}^d m_i^{-1} A_i \otimes X_i \right\rangle \\ &= \sum_{i=1}^d m_i^{-1} \langle Y, A_i \rangle \langle Z, X_i \rangle \\ &= \langle Z, M \rangle, \end{aligned}$$

as we wanted. \square

We also have the following theorem, which is analogous to Theorem 4.

Theorem 10. Let $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$ be a coherent configuration and suppose that the linear matrix inequality

$$\sum_{i=1}^d \alpha_i X_i \succeq 0, \quad (1.16)$$

where $\alpha_1, \dots, \alpha_d$ are scalars such that $\alpha_{i^*} = \alpha_i$ for $i = 1, \dots, d$, is satisfied by $X_i = A_i$.

If matrices $X_1, \dots, X_d \in \mathbb{R}^{n \times n}$ such that $X_{i^*} = X_i^\top$ for $i = 1, \dots, d$ satisfy (1.11), then X_1, \dots, X_d also satisfy (1.16).

Proof. We claim that there is a positive semidefinite matrix $Y \in \mathbb{R}^{n \times n}$ such that

$$\langle Y, A_i \rangle = \alpha_i m_i \quad \text{for } i = 1, \dots, d.$$

Indeed, from Farkas' lemma we know that either there exists such a matrix Y , or there are numbers β_1, \dots, β_d such that

$$\sum_{i=1}^d \beta_i A_i \succeq 0 \quad \text{and} \quad \sum_{i=1}^d \alpha_i \beta_i m_i < 0.$$

Now, since (1.16) is satisfied by $X_i = A_i$, we know that

$$\begin{aligned} 0 &\leq \left\langle \sum_{i=1}^d \alpha_i A_i, \sum_{j=1}^d \beta_j A_j \right\rangle \\ &= \sum_{i,j=1}^d \alpha_i \beta_j \langle A_i, A_j \rangle \\ &= \sum_{i=1}^d \alpha_i \beta_i m_i \\ &< 0, \end{aligned}$$

a contradiction. Here, we used the fact that $\langle A_i, A_j \rangle = \delta_{ij} m_i$ for $i, j = 1, \dots, d$. So there must be a matrix Y as claimed.

Now, if Y is given as in our claim, then we have that

$$\sum_{i=1}^d m_i^{-1} \langle Y, A_i \rangle X_i = \sum_{i=1}^d \alpha_i X_i,$$

and since X_1, \dots, X_d satisfy (1.11), from Theorem 9 we see that the matrix above is positive semidefinite, as we wanted. \square

Finally, we observe that for an association scheme $\{A_0, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$, relaxation (1.12) is equivalent to (1.5).

To see this, it suffices to show that if $\{A_0, \dots, A_d\}$ is an association scheme, then the constraint

$$\sum_{i=0}^d m_i^{-1} A_i \otimes X_i \succeq 0$$

is equivalent to the constraints

$$\sum_{i=0}^d q_{ij} X_i \succeq 0 \quad \text{for } j = 0, \dots, d. \quad (1.17)$$

Indeed, since $\{A_0, \dots, A_d\}$ is an association scheme, the matrices A_0, \dots, A_d can be simultaneously diagonalized via a unitary transformation. Then we may write

$$A_i = \sum_{j=0}^d p_{ji} E_j,$$

where the matrices E_0, \dots, E_d are 0–1 diagonal matrices that sum to the identity. So we see that the linear matrix inequality

$$\sum_{i=0}^d m_i^{-1} \sum_{j=0}^d p_{ji} E_j \otimes X_i = \sum_{i=0}^d m_i^{-1} A_i \otimes X_i \succeq 0$$

is equivalent to the linear matrix inequalities

$$\sum_{i=0}^d m_i^{-1} p_{ji} X_i \succeq 0 \quad \text{for } j = 0, \dots, d,$$

and by using (1.4) we get that the above constraints are equivalent to (1.17), as we wanted.

1.6.1 Relation to a relaxation of the quadratic assignment problem

Given symmetric matrices A and $W \in \mathbb{R}^{n \times n}$, a simplified version of the quadratic assignment problem is as follows:

$$\begin{aligned} & \text{maximize } \langle W, P^T A P \rangle \\ & P \in \mathbb{R}^{n \times n} \text{ is a permutation matrix.} \end{aligned}$$

Povh and Rendl [20] proposed the following semidefinite programming relaxation for this problem:

$$\begin{aligned} & \text{maximize } \langle A \otimes W, Y \rangle \\ & \langle I \otimes E_{jj}, Y \rangle = \langle E_{jj} \otimes I, Y \rangle = 1 \quad \text{for } j = 1, \dots, n, \\ & \langle I \otimes (J - I) + (J - I) \otimes I, Y \rangle = 0, \\ & \langle J \otimes J, Y \rangle = n^2, \\ & Y \geq 0 \text{ and } Y \succeq 0, \end{aligned} \tag{1.18}$$

where E_{jj} is the matrix with 1 in position (j, j) and 0 everywhere else. This relaxation is in fact equivalent to an earlier one due to Zhao et al. [25].

When the matrix A belongs to the span of a coherent configuration $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$, then one may use (1.12) to provide a semidefinite programming relaxation of the quadratic assignment problem as follows

$$\begin{aligned} & \text{maximize } \langle W, \sum_{i=1}^d m_i^{-1} \langle A, A_i \rangle X_i \rangle \\ & \sum_{i \in N} X_i = I, \\ & \sum_{i=1}^d X_i = J, \\ & \sum_{i=1}^d m_i^{-1} A_i \otimes X_i \succeq 0, \\ & \langle J, X_i \rangle = m_i, X_i^* = X_i^T, \text{ and } X_i \geq 0 \text{ for } i = 1, \dots, d. \end{aligned} \tag{1.19}$$

Actually, when A belongs to the span of a coherent configuration, relaxations (1.18) and (1.19) are equivalent. This result is essentially due to de Klerk and Sotirov [14], but we give a proof for the sake of completeness.

Theorem 11 (cf. de Klerk and Sotirov [14]). *If A belongs to the span of a coherent configuration $\{A_1, \dots, A_d\} \subseteq \mathbb{R}^{n \times n}$, then the optimal values of (1.18) and (1.19) coincide.*

Proof. We first show that the optimal value of (1.18) is at most that of (1.19).

To this end, let $Y \in \mathbb{R}^{n \times n} \otimes \mathbb{R}^{n \times n}$ be a feasible solution of (1.18). Write $\mathcal{A} = \text{span}\{A_1, \dots, A_d\}$ and consider the matrices

$$m_i^{-1/2} A_i \otimes E_{kl} \quad \text{for } i = 1, \dots, d \text{ and } k, l = 1, \dots, n,$$

where $E_{ij} \in \mathbb{R}^{n \times n}$ is the matrix with 1 at position (i, j) and 0 everywhere else. The matrices above form an orthonormal basis of the matrix $*$ -algebra $\mathcal{A} \otimes \mathbb{C}^{n \times n}$. From Theorem 8, we know that the orthogonal projection of Y onto this algebra is a positive semidefinite matrix. In view of the basis given above for the algebra, this projection may be explicitly written as

$$\bar{Y} = \sum_{i=1}^d m_i^{-1} A_i \otimes X_i,$$

where $X_i \in \mathbb{R}^{n \times n}$, and this is therefore a positive semidefinite matrix.

We now show that X_1, \dots, X_d is a feasible solution of (1.19). Indeed, it is immediate that $X_i^* = X_i^T$ for $i = 1, \dots, d$ and, since the A_i are 0–1 matrices with disjoint supports and Y is nonnegative, we also see that the X_i are nonnegative. Moreover, by construction we have

$$\sum_{i=1}^d m_i^{-1} A_i \otimes X_i \succeq 0.$$

To see that $\sum_{i \in N} X_i = I$, notice that for all $j = 1, \dots, n$, we have $I \otimes E_{jj} \in \mathcal{A} \otimes \mathbb{C}^{n \times n}$. Then, since Y is feasible for (1.18) and since $\{A_1, \dots, A_d\}$ is a coherent configuration, we have for $j = 1, \dots, n$ that

$$1 = \langle I \otimes E_{jj}, \bar{Y} \rangle = \sum_{i=1}^d m_i^{-1} \langle I, A_i \rangle \langle E_{jj}, X_i \rangle = \sum_{i \in N} \langle E_{jj}, X_i \rangle. \quad (1.20)$$

Now, since Y is feasible for (1.18) we also have that

$$\begin{aligned} 0 &= \langle I \otimes (J - I) + (J - I) \otimes I, Y \rangle \\ &= \langle I \otimes (J - I) + (J - I) \otimes I, \bar{Y} \rangle \\ &= \sum_{i=1}^d m_i^{-1} \langle I, A_i \rangle \langle J - I, X_i \rangle + \langle J - I, A_i \rangle \langle I, X_i \rangle. \end{aligned} \quad (1.21)$$

Since each X_i is nonnegative, this implies that whenever $i \in N$, all off-diagonal entries of X_i are equal to zero. But then together with (1.20) we have that $\sum_{i \in N} X_i = I$.

We now claim that for $i \in N$ one has $\langle J, A_i \rangle = m_i$. Indeed, since $\langle E_{jj} \otimes I, Y \rangle = 1$ for $j = 1, \dots, n$, for $i \in N$ we have that $\langle A_i \otimes I, Y \rangle = m_i$. But then we have that

$$m_i = \langle A_i \otimes I, \bar{Y} \rangle = \sum_{j=1}^d m_j^{-1} \langle A_i, A_j \rangle \langle I, X_j \rangle = \langle I, X_i \rangle.$$

But then, since $\sum_{i \in N} X_i = I$ and since each X_i is nonnegative, we must have that $\langle J, X_i \rangle = \langle I, X_i \rangle = m_i$ for $i \in N$, as we wanted.

Now we show that in fact $\langle J, X_i \rangle = m_i$ for $i = 1, \dots, d$. To see this, notice that the matrix

$$\sum_{i=1}^d m_i^{-1} \langle J, X_i \rangle A_i = (I \otimes e)^\top \left(\sum_{i=1}^d m_i^{-1} A_i \otimes X_i \right) (I \otimes e)$$

is positive semidefinite. But since $\langle J, X_i \rangle = m_i$ for $i \in N$, the diagonal entries of this matrix are all equal to 1. So since this is a positive semidefinite matrix, it must be that all of its entries have absolute value at most 1. But then, since $\{A_1, \dots, A_d\}$ is a coherent configuration, and since each X_i is nonnegative, it must be that $\langle J, X_i \rangle \leq m_i$ for $i = 1, \dots, n$.

Now, we know that

$$n^2 = \langle J \otimes J, Y \rangle = \langle J \otimes J, \bar{Y} \rangle = \sum_{i=1}^d \langle J, X_i \rangle,$$

and it follows that $\langle J, X_i \rangle = m_i$ for $i = 1, \dots, n$.

To prove that X_1, \dots, X_d is feasible for (1.19), we still have to show that $X_1 + \dots + X_d = J$. To this end, notice that, since the X_i are nonnegative, (1.21) implies that every X_i with $i \notin N$ has only zeros on the diagonal. So the matrix $X_1 + \dots + X_d$ has diagonal entries equal to 1 and then, since it is positive semidefinite (as \bar{Y} is positive semidefinite), and since

$$\sum_{i=1}^d \langle J, X_i \rangle = n^2,$$

we see immediately that $X_1 + \dots + X_d = J$, as we wanted.

So X_1, \dots, X_d is feasible for (1.19). Finally, the objective value of X_1, \dots, X_d coincides with that of Y , and we are done.

To see now that the optimal value of (1.19) is at most that of (1.18), one has to show that, given a solution X_1, \dots, X_d of (1.19), the matrix

$$Y = \sum_{i=1}^d m_i^{-1} A_i \otimes X_i$$

is a feasible solution of (1.18) with the same objective value, and showing this is rather straight-forward. \square

1.6.2 Coherent configurations and combinatorial optimization problems

In Section 1.5 we saw how the problem of finding a maximum/minimum-weight subgraph H of a complete graph can be modelled with the help of association schemes when the graph H is distance-regular. We now show how to model this problem as an optimization problem over the convex hull of a coherent configuration for general graphs H .

Let $H = (V, E)$ be a graph, with $V = \{1, \dots, n\}$. An *automorphism* of H is a permutation $\pi: V \rightarrow V$ that preserves the adjacency relation. The set of all automorphisms of H , denoted by $\text{Aut}(H)$, is a group under the operation of function composition.

For $(x, y), (x', y') \in V \times V$, we write

$$(x, y) \sim (x', y')$$

if there is $\pi \in \text{Aut}(H)$ such that $\pi(x) = x'$ and $\pi(y) = y'$. Relation \sim is an equivalence relation that partition $V \times V$ into equivalence classes called *orbits*. Say $V \times V / \sim = \{C_1, \dots, C_d\}$ and consider the matrices $A_1, \dots, A_d \in \mathbb{R}^{n \times n}$ such that

$$A_i(x, y) = \begin{cases} 1 & \text{if } (x, y) \in C_i, \\ 0 & \text{otherwise.} \end{cases}$$

We claim that $\{A_1, \dots, A_d\}$ is a coherent configuration.

Indeed, it is easy to see that $\{A_1, \dots, A_d\}$ satisfies items (1)–(2) in the definition of a coherent configuration. We show that item (3) is also satisfied, that is, that the span of $\{A_1, \dots, A_d\}$ is a matrix $*$ -algebra.

To see this, associate to each permutation $\pi \in \text{Aut}(H)$ a permutation matrix $P_\pi \in \mathbb{R}^{n \times n}$ in the obvious way. Then the set of matrices

$$\{A \in \mathbb{R}^{n \times n} : AP_\pi = P_\pi A \text{ for } \pi \in \text{Aut}(H)\}$$

is exactly the span of $\{A_1, \dots, A_d\}$, and it can be then easily seen that this span is also a matrix $*$ -algebra. So $\{A_1, \dots, A_d\}$ is a coherent configuration.

Now suppose we are given a symmetric matrix $W \in \mathbb{R}^{n \times n}$ which we see as giving weights to the edges of the complete graph K_n , and suppose we are asked to find a maximum-weight copy of H in K_n . Let A be the adjacency matrix of H . Then we may write this problem equivalently as

$$\begin{aligned} & \text{maximize } \frac{1}{2} \langle W, P^\top A P \rangle \\ & P \in \mathbb{R}^{n \times n} \text{ is a permutation matrix,} \end{aligned}$$

It can be easily seen that, for some subset $M \subseteq \{1, \dots, d\}$, one has

$$A = \sum_{i \in M} A_i.$$

So the problem above can be rewritten as

$$\begin{aligned} & \text{maximize } \frac{1}{2} \langle W, \sum_{i \in M} X_i \rangle \\ & (X_1, \dots, X_d) \in \text{conv}(\{A_1, \dots, A_d\}). \end{aligned}$$

Finally, when we consider relaxation (1.12) for this problem, we obtain

$$\begin{aligned} & \text{maximize } \frac{1}{2} \langle W, \sum_{i \in M} X_i \rangle \\ & \sum_{i \in N} X_i = I, \\ & \sum_{i=1}^d X_i = J, \\ & \sum_{i=1}^d m_i^{-1} A_i \otimes X_i \succeq 0, \\ & \langle J, X_i \rangle = m_i, X_{i^*} = X_i^\top, \text{ and } X_i \geq 0 \text{ for } i = 1, \dots, d. \end{aligned}$$

In our discussion it was not in any moment necessary to work with the full automorphism group of H ; any subgroup of this group will do. In any case, the larger the group one uses, the smaller the matrices A_i can be made to be after the block-diagonalization.

If the automorphism group of H is transitive, then it is possible to obtain stronger SDP relaxations by using a stabilizer subgroup in stead of the the full automorphism group; see de Klerk and Sotirov [15]. A similar idea was used by Schrijver [22] earlier to obtain improved SDP bounds for binary code sizes.

1.6.3 The maximum (k, l) -cut problem

The maximum (k, l) -cut problem is a generalization of the maximum bisection problem which consists of the following: Given a symmetric nonnegative matrix $W \in \mathbb{R}^{n \times n}$, which we see as giving weights to the edges of the complete graph K_n , where $n = k + l$, find a maximum-weight copy of $K_{k,l}$ in K_n .

We may use our semidefinite programming relaxation of the convex hull of a coherent configuration to give a relaxation of the maximum (k, l) -cut problem by considering the following coherent configuration associated with the complete bipartite graph $K_{k,l}$:

$$\begin{aligned} A_1 &= \begin{pmatrix} I_k & 0_{k \times l} \\ 0_{l \times k} & 0_{l \times l} \end{pmatrix}, \quad A_2 = \begin{pmatrix} J_k - I_k & 0_{k \times l} \\ 0_{l \times k} & 0_{l \times l} \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0_{k \times k} & J_{k \times l} \\ 0_{l \times k} & 0_{l \times l} \end{pmatrix}, \\ A_4 &= \begin{pmatrix} 0_{k \times k} & 0_{k \times l} \\ J_{l \times k} & 0_{l \times l} \end{pmatrix}, \quad A_5 = \begin{pmatrix} 0_{k \times k} & 0_{k \times l} \\ 0_{l \times k} & I_l \end{pmatrix}, \quad A_6 = \begin{pmatrix} 0_{k \times k} & 0_{k \times l} \\ 0_{l \times k} & J_l - I_l \end{pmatrix}. \end{aligned} \quad (1.22)$$

Now, relaxation (1.12) simplifies to:

$$\begin{aligned} & \text{maximize } \frac{1}{2} \langle W, X_3 + X_4 \rangle \\ & X_1 + X_5 = I, \\ & \sum_{i=1}^6 X_i = J, \\ & \sum_{i=1}^6 m_i^{-1} A_i \otimes X_i \succeq 0, \\ & \langle J, X_i \rangle = m_i, X_{i^*} = X_i^\top, \text{ and } X_i \geq 0 \text{ for } i = 1, \dots, 6. \end{aligned} \quad (1.23)$$

The algebra spanned by A_1, \dots, A_6 is isomorphic to the algebra $\mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}^{2 \times 2}$, so one may replace the matrices A_1, \dots, A_6 above by their respective images under this isomorphism, which we denote by ϕ . Their images are given by

$$\begin{aligned} \phi(A_1) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \phi(A_2) = \begin{pmatrix} -1 & 0 \\ 0 & k-1 \\ 0 & 0 \end{pmatrix}, \quad \phi(A_3) = \sqrt{kl} \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ \phi(A_4) &= \sqrt{kl} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \phi(A_5) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \phi(A_6) = \begin{pmatrix} 0 & -1 \\ 0 & l-1 \\ 0 & 0 \end{pmatrix}. \end{aligned} \quad (1.24)$$

Feige and Langberg [4] and Han, Ye, and Zhang [11] proposed another semidefinite programming relaxation for the maximum (k, l) -cut problem, namely

$$\begin{aligned} &\text{maximize } \frac{1}{4} \langle W, J - X \rangle \\ &\quad X_{ii} = 1 \quad \text{for } i = 1, \dots, n, \\ &\quad \langle J, X \rangle = (k - l)^2, \\ &\quad X \succeq 0. \end{aligned} \quad (1.25)$$

Our relaxation (1.23) is actually as strong as (1.25). Indeed, if (X_1, \dots, X_6) is a feasible solution of (1.23), one may construct a feasible solution of (1.25) by setting

$$X = X_1 + X_2 - X_3 - X_4 + X_5 + X_6.$$

Indeed, it can be easily seen that X has diagonal entries equal to 1 and satisfies $\langle J, X \rangle = (k - l)^2$. To see that X is positive semidefinite, notice that the matrices A_1, \dots, A_6 satisfy

$$A_1 + A_2 - A_3 - A_4 + A_5 + A_6 \succeq 0,$$

as one may easily check by using the isomorphism ϕ . Then, from Theorem 10 one has that X is positive semidefinite.

Now, also $\frac{1}{4} \langle W, J - X \rangle = \frac{1}{2} \langle W, X_3 + X_4 \rangle$, so X has the same objective value as X_1, \dots, X_6 , as we wanted. Finally, our relaxation is actually stronger, as can be checked numerically in some examples. In particular, we have the following theorem.

Theorem 12. *The optimal value of (1.23) is at most that of (1.25), and can be strictly smaller for specific instances.*

1.6.4 The maximum stable set problem

Let $G = (V, E)$ be a graph with adjacency matrix A and stability number $\alpha(G)$ and let an integer $k > 1$ be given. One has $\alpha(G) \geq k$, if and only if

$$\begin{aligned} 0 &= \min \langle A_1 + A_2, P^\top A P \rangle \\ &\quad P \in \mathbb{R}^{n \times n} \text{ is a permutation matrix,} \end{aligned}$$

where A_1 and A_2 are from the coherent configuration in (1.22).

Applying our procedure we obtain the SDP problem:

$$\begin{aligned}
& \text{minimize } \langle A, X_1 + X_2 \rangle \\
& X_1 + X_5 = I, \\
& \sum_{i=1}^6 X_i = J, \\
& \sum_{i=1}^6 m_i^{-1} \phi(A_i) \otimes X_i \succeq 0, \\
& \langle J, X_i \rangle = m_i, X_{i*} = X_i^T, \text{ and } X_i \geq 0 \text{ for } i = 1, \dots, 6,
\end{aligned} \tag{1.26}$$

where the $\phi(A_i)$ are as in (1.24). We denote by $\sigma_k(G)$ the optimal value of (1.26) when A is the adjacency matrix of the graph G .

If the optimal value of this problem is strictly positive, then $\alpha(G) < k$. So an upper bound for $\alpha(G)$ is given by

$$\max\{k : \sigma_k(G) = 0\}.$$

The bound provided above is at least as strong as the ϑ' bound of McEliece, Rodemich, and Rumsey [19] and Schrijver [21]. Recall that, given a graph $G = (V, E)$, we let $\vartheta'(G)$ be the optimal value of the semidefinite programming problem

$$\begin{aligned}
& \text{maximize } \langle J, X \rangle \\
& \text{trace } X = 1, \\
& X_{uv} = 0 \text{ if } uv \in E, \\
& X : V \times V \rightarrow \mathbb{R} \text{ is nonnegative and positive semidefinite.}
\end{aligned} \tag{1.27}$$

It is easy to see that

$$\max\{k : \sigma_k(G) = 0\} \leq \lfloor \vartheta'(G) \rfloor.$$

To see this, suppose $\sigma_k(G) = 0$. We show that $\vartheta'(G) \geq k$. So let X_1, \dots, X_6 be an optimal solution of (1.26) for the parameter k and set $X = k^{-1}(X_1 + X_2)$. Then X is a feasible solution of (1.27).

Indeed, X is nonnegative, and it is also positive semidefinite since $A_1 + A_2$ is a positive semidefinite matrix. Since the X_i sum up to J and $X_1 + X_5 = I$, we have that $\text{trace } X = k^{-1} \text{trace } X_1 = k^{-1} \langle J, X_1 \rangle = 1$. Finally since $\langle A, X \rangle = k^{-1} \langle A, X_1 + X_2 \rangle = 0$ and X is nonnegative, we see that $X_{uv} = 0$ whenever $uv \in E$, and therefore X is feasible for (1.27).

Now, notice that $\langle J, X \rangle = k^{-1} \langle J, X_1 + X_2 \rangle = k$, and we see that $\vartheta'(G) \geq k$, as we wanted.

1.6.5 The vertex separator problem

Let $G = (V, E)$ again be a graph with adjacency matrix A and $|V| = n$, and let $n_1, n_2 > 0$ be given integers satisfying $n_1 + n_2 < n$.

The vertex separator problem is to find disjoint subsets $V_1, V_2 \subset V$ such that $|V_i| = n_i$ ($i = 1, 2$) and no edge in E connects a vertex in V_1 with a vertex in V_2 , if such sets exist. (The set of vertices $V \setminus \{V_1 \cap V_2\}$ is called a vertex separator in this case; see Feige et al. [5] for a review of approximation results for various separator problems.)

Such a vertex separator exists if and only if

$$0 = \min \langle W, P^T A P \rangle$$

$P \in \mathbb{R}^{n \times n}$ is a permutation matrix,

where

$$W = \begin{pmatrix} 0_{(n-n_1-n_2) \times (n-n_1-n_2)} & 0_{(n-n_1-n_2) \times n_1} & 0_{(n-n_1-n_2) \times n_2} \\ 0_{n_1 \times (n-n_1-n_2)} & 0_{n_1 \times n_1} & J_{n_1 \times n_2} \\ 0_{n_2 \times (n-n_1-n_2)} & J_{n_2 \times n_1} & 0_{n_2 \times n_2} \end{pmatrix},$$

and where the subscripts again indicate the matrix sizes. It is easy to verify that the matrix W belongs to a coherent configuration with $d = 12$ relations, say $\{A_1, \dots, A_{12}\}$.

Thus one may again obtain an SDP problem where a strictly positive optimal value gives a certificate that the required vertex separator does not exist. The details of this example are left as an exercise to the reader.

1.7 Conclusion

In this chapter we have given a unified framework for deriving SDP relaxations of combinatorial problems. This approach is still in its infancy, and the relationships to various existing SDP relaxations is not yet fully understood. Moreover, since SDP relaxations play an important role in approximation algorithms, it would be very desirable to extend this framework with a general rounding technique that is amenable to analysis. It is the hope of the authors that these topics will spark the interest of other researchers in the SDP community.

Acknowledgement

Etienne de Klerk would like to thank Chris Godsil for many valuable discussions on the contents of this chapter.

References

1. G. Andersson, An approximation algorithm for Max p -Section, *Lecture Notes in Computer Science* **1563** (1999) 236–247.

2. D. Cvetković, M. Cangalović, and V. Kovačević-Vujčić, Semidefinite programming methods for the symmetric travelling salesman problem, in: *Proceedings of the 7th International Conference on Integer Programming and Combinatorial Optimization*, Springer-Verlag, London, 1999, pp. 126–136.
3. P.J. Cameron, *Permutation Groups*, London Mathematical Society Student Texts 45, Cambridge University Press, Cambridge, 1999.
4. U. Feige and M. Langberg, Approximation algorithms for maximization problems arising in graph partitioning, *Journal of Algorithms* **41** (2001) 174–211.
5. U. Feige, J.R. Lee and M.T. Hajiaghayi. Improved approximation algorithms for minimum-weight vertex separators. *SIAM Journal on Computing* **38**(2) (2008) 629–657.
6. A. Frieze and M. Jerrum, Improved approximation algorithms for MAX k -CUT and MAX BISECTION, *Algorithmica* **18** (1997) 67–81.
7. C.D. Godsil, *Algebraic Combinatorics*, Chapman & Hall, New York, 1993.
8. M.X. Goemans and D.P. Williamson, A general approximation technique for constrained forest problems, *SIAM Journal on Computing* **24** (1995) 296–317.
9. M.X. Goemans and D.P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of the ACM* **42** (1995) 1115–1145.
10. W. Greub, *Multilinear Algebra*, Springer-Verlag, New York, 1978.
11. Q. Han, Y. Ye, J. Zhang, An improved rounding method and semidefinite relaxation for graph partitioning, *Mathematical Programming* **92** (2002) 509–535.
12. S. Khot, G. Kindler, E. Mossel, R. O'Donnell, Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?, *SIAM Journal on Computing* **37** (2007) 319–357.
13. E. de Klerk, D.V. Pasechnik, and R. Sotirov, On semidefinite programming relaxations of the travelling salesman problem, *SIAM Journal on Optimization* **19** (2008) 1559–1573.
14. E. de Klerk and R. Sotirov, Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. *Mathematical Programming* **122**(2) (2010) 225–246.
15. E. de Klerk and R. Sotirov, Improved semidefinite programming bounds for quadratic assignment problems with suitable symmetry. Preprint, 2010. Available at: http://www.optimization-online.org/DB_HTML/2009/09/2399.html
16. M. Laurent and F. Rendl, Semidefinite programming and integer programming, in: *Handbook on Discrete Optimization* (K. Aardal, G. Nemhauser, and R. Weismantel, eds.), Elsevier, Amsterdam, 2005, pp. 393–514.
17. L. Lovász, On the Shannon capacity of a graph, *IEEE Transactions on Information Theory* **25** (1979) 1–7.
18. B. Manthey, Minimum-weight cycle covers and their approximability, *Discrete Applied Mathematics* **157** (2009) 1470–1480.
19. R.J. McEliece, E.R. Rodemich, H.C. Rumsey, Jr., The Lovász bound and some generalizations, *Journal of Combinatorics, Information & System Sciences* **1978** (134–152) .
20. J. Povh and F. Rendl, Copositive and semidefinite relaxations of the quadratic assignment problem, *Discrete Optimization* **6** (2009) 231–241.
21. A. Schrijver, A comparison of the Delsarte and Lovász bounds, *IEEE Transactions on Information Theory* **25** (1979) 425–429.

22. A. Schrijver, New code upper bounds from the Terwilliger algebra, *IEEE Transactions on Information Theory* **51** (2005) 2859–2866.
23. J.H.M. Wedderburn, On hypercomplex numbers, *Proceedings of the London Mathematical Society* **6** (1907) 77–118.
24. Y. Ye, A .699 approximation algorithm for max-bisection, *Mathematical Programming* **90** (2001) 101–111.
25. Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz. Semidefinite programming relaxations for the quadratic assignment problem. *Journal of Combinatorial Optimization* **2** (1998) 71–109.