Exploiting special structure in semidefinite programming: a survey of theory and applications

Etienne de Klerk*

January 15, 2009

Abstract

Semidefinite Programming (SDP) may be seen as a generalization of Linear Programming (LP). In particular, one may extend interior point algorithms for LP to SDP, but it has proven much more difficult to exploit structure in the SDP data during computation.

We survey three types of special structures in SDP data:

1. a common 'chordal' sparsity pattern of all the data matrices. This structure arises in applications in graph theory, and may also be used to deal with more general sparsity patterns in a heuristic way.
2. low rank of all the data matrices. This structure is common in SDP relaxations of combinatorial optimization problems, and SDP approximations of polynomial optimization problems.
3. the situation where the data matrices are invariant under the action of a permutation group, or, more generally, where the data matrices belong to a low dimensional matrix algebra. Such problems arise in truss topology optimization, particle physics, coding theory, computational geometry, and graph theory.

We will give an overview of existing techniques to exploit these structures in the data. Most of the paper will be devoted to the third situation, since it has received the least attention in the literature so far.

Keywords: semidefinite programming, chordal sparsity, algebraic symmetry

AMS classification: 90C22

*Department of Econometrics and OR, Tilburg University, The Netherlands. e.deklerk@uvt.nl

1 Introduction

What is the difference between method and device? A method is a device which you used twice.

George Pólya (1887 — 1985)

The phrase *semidefinite programming* (SDP) was coined in the 1990's, following the extension of interior point methods from linear programming to convex programming by Nesterov and Nemirovski [39], and independent work by Alizadeh [1]. The topic itself is somewhat older (see e.g. the 1963 paper by Bellman and Fan [5]), and arises naturally in the study of positive polynomials, which in turn has applications in system and control theory; for general surveys on SDP, see Vandenberghe and Boyd [50], and Todd [46].

It has been known since the 1970's that semidefinite programs may be solved to any fixed accuracy in polynomial time (in the real number model), by using the ellipsoid algorithm of Yudin and Nemirovski [38]. The ellipsoid algorithm was later used by Khachiyan [28] to prove the polynomial complexity of LP in the bit model.

The complexity results surrounding the ellipsoid method quickly drew attention in combinatorial optimization, and resulted in notable papers on SDP in the late 1970's. One of these papers, by Schrijver [41], introduced a methodology that is now known as symmetry reduction in SDP. This methodology may be seen as a way to reduce the size of SDP instances with 'group symmetric data'. After the development of practical interior point methods in the 1990's, this methodology gained new interest. The reason being that it proved much harder to exploit general sparsity in SDP problem data than in the LP case when using interior point methods. Thus, researchers became interested in structures in SDP data that allow problem size reduction.

In this paper, we will review three such structures, as well as some of their applications.

Notation and preliminaries

The space of $p \times q$ real (resp. complex) matrices will be denoted by $\mathbb{R}^{p\times q}$ (resp. $\mathbb{C}^{p\times q}$), and the space of $k \times k$ symmetric matrices by $\mathbb{S}^{k\times k}$.

We use I_n to denote the identity matrix of order n. Similarly, J_n denotes the $n \times n$ all-ones matrix. We will omit the subscript if the order is clear from the context. The standard unit vectors of $\mathbb{R}^n$ are denoted by $e_1, \ldots, e_n$, and e denotes the all-ones vector of size depending on the context.

A complex matrix $A \in \mathbb{C}^{n\times n}$ may be decomposed as

$$A = \mathrm{Re}(A) + \sqrt{-1}\mathrm{Im}(A),$$

where $\mathrm{Re}(A) \in \mathbb{R}^{n\times n}$ and $\mathrm{Im}(A) \in \mathbb{R}^{n\times n}$ are the real and imaginary parts of A, respectively. The *complex conjugate transpose* is defined as:

$$A^* = \mathrm{Re}(A)^T - \sqrt{-1}\mathrm{Im}(A)^T,$$

where the superscript 'T' denotes the transpose.

A matrix $A \in \mathbb{C}^{n\times n}$ is called *Hermitian* if $A^* = A$, i.e. if $\mathrm{Re}(A)$ is symmetric and $\mathrm{Im}(A)$ is skew-symmetric. A matrix $Q \in \mathbb{C}^{n\times n}$ is called *unitary* if $Q^*Q = I$. A real unitary matrix is called *orthogonal*.

The $\mathrm{vec}(\cdot)$ operator stacks the columns of a matrix, while the $\mathrm{diag}(\cdot)$ operator maps an $n \times n$ matrix to the n-vector given by its diagonal.

The *Kronecker product* $A \otimes B$ of matrices $A \in \mathbb{R}^{p\times q}$ and $B \in \mathbb{R}^{r\times s}$ is defined as the $pr \times qs$ matrix composed of pq blocks of size $r \times s$, with block ij given by $A_{ij}B$ $(i = 1, \ldots, p)$, $(j = 1, \ldots, q)$.

Structured instances in standard form

We consider the standard form SDP problem

$$\min_{X\succeq 0} \left\{ \mathrm{trace}(A_0 X) \ : \ \mathrm{trace}\,(A_k X) = b_k \ \ \forall\, k = 1, \ldots, m \right\}, \tag{1}$$

where the data matrices A_i are $n \times n$ and are linearly independent, and $X \succeq 0$ means X is positive semidefinite (or psd, for short). In most applications the A_i are real symmetric matrices, but the general case where the A_i are Hermitian matrices is often a useful theoretical setting. In the latter case X is a Hermitian positive semidefinite matrix. In what follows, we will assume the matrices to be real, unless otherwise specified.

Using the trace inner product, the dual problem is

$$\max_{y\in\mathbb{R}^m,\, S\succeq 0} \left\{ b^T y \ : \ \sum_{i=1}^{m} y_i A_i + S = A_0 \right\}. \tag{2}$$

At some point, we will also consider the additional constraint $X \geq 0$ (nonnegativity) in the real case.

The aim of this survey is to show how certain structures in the data matrices A_i $(i = 1, \ldots, m)$ may be exploited in order to reduce the computational requirements.

There are currently three types of structures (apart from general sparsity) that may be exploited in SDP.

Chordal structure

Here the matrices A_i $(i = 0, \ldots, m)$ have a common sparsity pattern, and this pattern is the same as the sparsity pattern of the adjacency matrix of some chordal graph. (Recall that a graph is called chordal if it does not contain a cycle of length 4 or more as an induced subgraph.)

Low rank

Here the matrices A_i $(i = 1, \ldots, m)$ have low rank. (The matrix A_0 may be arbitrary.)

Algebraic symmetry

Here the matrices A_i ($i = 0, \dots, m$) belong to a matrix $*$-algebra of low dimension. (Recall that a matrix $*$-algebra over a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is a subspace of $\mathbb{F}^{n \times n}$ that is also closed under matrix multiplication and taking (conjugate) transposes.) This structure mostly arises when the data matrices are invariant under the action of a permutation group.

2 Data matrices of low rank

Many SDP problems arising in combinatorial optimization involve rank one matrices in the constraints, say

$$A_i = a_i a_i^T, \quad a_i \in \mathbb{R}^n, \quad i = 1, \dots, m. \tag{3}$$

In this case the Schur complement matrix appearing in all interior point methods can be formed more efficiently than in the general case.

Indeed, the Schur complement matrix (say M) has entries of the form

$$m_{ij} = \mathrm{trace}(A_i Z_1 A_j Z_2) \quad (i, j = 1, \dots, m),$$

where Z_1 and Z_2 are positive definite matrices that depend on the choice of the search direction; for a survey on search directions of interior point methods for SDP, see the survey by Todd [45].

For example, in the dual logarithmic barrier (or in the dual scaling) method, one has $Z_1 = Z_2 = S^{-1}$, where S is the value of the dual matrix variable at the current iterate [6].

Substituting the expressions from (3), this reduces to:

$$\begin{aligned} m_{ij} &:= \mathrm{trace}\left(A_i Z_1 A_j Z_2\right) \\ &= \mathrm{trace}\left(a_i a_i^T Z_1 a_j a_j^T Z_2\right) \\ &= \left(a_i^T Z_1 a_j\right)\left(a_i^T Z_2 a_j\right). \end{aligned}$$

For the dual methods, where $Z_1 = Z_2 = S^{-1}$, this simplifies to

$$m_{ij} = \left(a_i^T S^{-1} a_j\right)^2 \quad i, j = 1 \dots, m.$$

For SDP relaxations of Boolean quadratic problems (like the Goemans-Williamson maximum cut relaxation in Example 1) this expression simplifies even more, as a_i is then simply the ith standard unit vector. In this case we have (for dual methods):

$$m_{ij} = \left(S^{-1}\right)_{ij}^2, \quad i, j = 1, \dots, m.$$

The Schur complement matrix M can therefore be formed efficiently if S is sparse. For many applications this is indeed the case, since

$$S = -\sum_{i=1}^{m} y_i A_i + A_0,$$

and C often has the same sparsity structure as the adjacency matrix of a sparse graph (see Example 1). If S is indeed sparse, one can compute S^{-1} by doing a sparse Choleski factorization of S with pre-ordering to reduce fill in. Although S^{-1} is not necessarily sparse if S is, this is often the case in practice, and then the Schur complement matrix is also sparse and can likewise be factored using sparse Choleski techniques.

Although we have only discussed the case where the constraint matrices are rank 1, it is obvious that this approach can be extended to the case where the constraint matrices have "low" rank.

Example 1. *Given a graph $G = (V, E)$ with Laplacian matrix L,*[1] *the following SDP relaxation by Goemans and Williamson [17] provides an upper bound on the cardinality of a maximum cut in the graph:*

$$\max_X \left\{ \frac{1}{4} \mathit{trace}(LX) \mid \operatorname{diag}(X) = e,\ X \succeq 0 \right\}. \tag{4}$$

This upper bound is known to be at most 1.14 *times the maximum cardinality of a cut in the graph [17].*

Note that the data matrices of this SDP problem, when cast in the standard form (1) are $A_0 := -\frac{1}{4}L$ and $A_i := e_i e_i^T$ $(i = 1, \ldots, |V| \equiv m)$. Thus, the matrices $A_1, \ldots, A_m$ are rank-one matrices. Moreover, the dual matrix variable takes the form

$$S = -\sum_{i=1}^{m} y_i (e_i e_i)^T - \frac{1}{4} L,$$

and therefore has the same sparsity pattern as the matrix L.

A second class of examples with rank-one data matrices arises in the study of nonnegative polynomials and polynomial optimization.

Example 2. *Consider a n-variate, degree d polynomial*

$$p(x) = \sum_{\alpha \in \mathbb{N}_0^n,\, \sum_i \alpha_i \le d} a_\alpha x^\alpha,$$

where $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$ and the $a_\alpha \in \mathbb{R}$ are given coefficients. Now consider the problem of determining if p may be written as a sum of squares of polynomials.

It is well known, and easy to show, that this is the case if and only if p may be written as:

$$p(x) = b(x)^T X b(x),$$

where $b : \mathbb{R}^n \mapsto \mathbb{R}^{\binom{n+d}{d}}$ is any basis of the n-variate polynomials of degree at most d, and X is some positive semidefinite matrix of order $\binom{n+d}{d}$.

The last equation is the same as requiring that

$$p(x_i) = b(x_i)^T X b(x_i) \equiv \mathit{trace}\left(b(x_i) b(x_i)^T X\right) \quad i = 1, \ldots, \binom{n+d}{d}, \tag{5}$$

[1] Recall that the Laplacian matrix of a graph G takes the form $L = D - A$, where D is the diagonal matrix of vertex degrees of G, and A is the adjacency matrix of G.

where the $x_i \in \mathbb{R}^n$ are points in general position, i.e. any n-variate polynomial of degree at most d is determined by the values it takes at these points.

This leads to an SDP problem where the data matrices

$$A_i := b(x_i)b(x_i)^T \quad i = 1, \ldots, \binom{n+d}{d} =: m$$

are all rank one matrices, namely, find an $X \succeq 0$ such that (5) holds.

Software and further reading

The SDP formulation in Example 2 is due to [34]. The same formulation can be applied to various SDP relaxations of polynomial optimization problems, e.g. to the SDP relaxations of Lasserre [30]. A recent survey of SDP relaxations for polynomial optimization problems is [32].

'Low rank' structure is already exploited by the solver DSDP [6] (that uses the dual-scaling direction), and the latest version of SDPT3 [47].

3 "Chordal" data structure

The data structure we are interested in here is the one where the nonzero elements in the data matrices correspond to the edges of a *chordal* graph.

Definition 1. *A graph is called chordal, if it does not contain a cycle of length four or more as an induced subgraph.*

To characterize positive semidefiniteness of matrices with a chordal sparsity pattern, we need to recall some definitions from graph theory. A *clique* in a graph $G = (V, E)$ is a subset $V' \subset V$ of vertices such that any pair of vertices in V' are adjacent. A clique in G is called *maximal* if it not a subset of a larger clique in G, and it is called *maximum* if it is a clique of maximum cardinality in G. The cardinality of a maximum clique in G is called the *clique number* of G, denoted by $\omega(G)$.

Similarly, a *co-clique* (or *stable set* or *independent set*) is a subset of vertices so that no two vertices in the subset are adjacent. Thus, one may define maximal and maximum co-cliques and the co-clique number of a graph as before. The co-clique number of G is usually denoted by $\alpha(G)$.

Theorem 1 ([20]). *Assume that the entries of a partially specified matrix X corresponds to a chordal graph G, and denote the maximal cliques of G by $K_1, ..., K_d$.*

Then the following two statements are equivalent:

- *X can be completed to a psd matrix;*
- *$X_{K_i,K_i} \succeq 0$ for all $i = 1, \ldots, d$, where X_{K_i,K_i} is the principal submatrix of X with rows and columns indexed by K_i.*

Formally, we now make the following assumption about the SDP problem data.

Assumption 1. *There exists a chordal graph $G = (V, E)$ with $|V| = n$ and clique number $\omega(G) \ll n$ such that $\{i, j\} \in E$ if $(A_k)_{ij} \neq 0$ for some k.*

We can therefore do a pre-processing step to find the maximal cliques of G. Subsequently, we can replace the primal SDP problem

$$\min_{X \succeq 0} \{\text{trace}(A_0 X) \ : \ \text{trace}\,(A_k X) = b_k \ \ \forall\, k = 1, \ldots, m\}\,,$$

by

$$\min_{X_{K_i,K_i} \succeq 0\ (i=1,\ldots,d)} \{\text{trace}(A_0 X) \ : \ \text{trace}\,(A_k X) = b_k \ \ \forall\, k = 1, \ldots, m\}\,.$$

Notice that, in solving the latter SDP, only the entries of X that correspond to edges in G need to be stored. If necessary, the full matrix X may be recovered at any point by solving a matrix completion problem.

Note that it is necessary to find all maximal cliques of G in order to do this pre-processing. In general, it is a NP-hard problem to find all maximal (or even one maximum) clique in a graph, but for chordal graphs this can be done efficiently (i.e. in polynomial time) via a combinatorial algorithm [40].

3.1 Chordal extensions

Notice that we did not specify how to find the chordal graph G in Definition 1. If the graph determined by the sparsity pattern in the data matrices is not chordal, then one may still construct a *chordal extension* (or *triangulation*) of it, i.e. add edges to it in order to obtain a chordal graph.

The difficulty here is that it is an NP-hard problem to find a minimum chordal extension, i.e. one where a minimum number of edges are added [53].

There are several algorithms available to perform a chordal extension, and we state one that is often used in practice. In this algorithm, a set of additional edges E' is generated, that gives the chordal extension.

Algorithm 1 (Elimination game)**.**

1. **Input:** *a graph $G = (V, E)$ and an ordering of V;*
2. *Set $G_1 = G$; $E' = \emptyset$;*
3. *for $i = 1$ to n do*

   Add edges to G_i as necessary to make all neighbors of vertex i pairwise adjacent, and add these new edges to the set E';

   Remove vertex i and call the resulting graph G_{i+1}.
4. **Output:** *$G' = (V, E \cup E')$, a chordal extension of G.*

It is not difficult to verify that the algorithm outputs a chordal extension of G.

Note that the chordal extension that is obtained depends on the initial numbering (ordering) of the vertices. A popular heuristic is the minimal degree ordering, i.e. to number the vertices V according to increasing degree; see e.g. [13].

Example 3. *Recall from Example 1, that the Goemans-Williamson SDP relaxation of the maximum cut problem for a graph $G = (V, E)$ takes the form:*

$$\max_X \left\{ \frac{1}{4} trace(LX) \mid \operatorname{diag}(X) = e, \ X \succeq 0 \right\}, \tag{6}$$

where L is the Laplacian matrix of the graph.

Now let K_i $(i = 1, \ldots, d)$ denote the maximal cliques of some chordal extension of G.

The constraint $X \succeq 0$ in (6) may now be replaced by $X_{K_i,K_i} \succeq 0$ $(i = 1, \ldots, d)$. If the maximal clique sizes are sufficiently small, then this results in a reduction of the SDP problem size.

Software and further reading

For a survey on matrix completion problems, see Laurent [31]. A recent application of the chordal completion strategy to the covariance selection problem is described in [9].

Chordal structure in SDP data is exploited by the solver SDPA-C [37].

4 Algebraic symmetry in SDP

We say that the SDP data matrices exhibit *algebraic symmetry* if they belong to a low dimensional matrix $*$-algebra, defined as follows.

Definition 2. *A set $\mathcal{A} \subseteq \mathbb{F}^{n \times n}$ is called a matrix *-algebra over a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ if, for all $X, Y \in \mathcal{A}$:*

- $\alpha X + \beta Y \in \mathcal{A} \quad \forall \alpha, \beta \in \mathbb{F}$;
- $X^* \in \mathcal{A}$;
- $XY \in \mathcal{A}$.

We may formally state the algebraic symmetry assumption for SDP as follows.

Assumption 2 (Algebraic symmetry)**.** *There exists a 'low dimensional' matrix $*$-algebra, say $\mathcal{A}_{SDP}$, that contains the data matrices $A_0, \ldots, A_m$.*

By 'low dimensional' we mean that $\dim(\mathcal{A}_{SDP}) \ll n$.

Under the algebraic symmetry assumption, we may restrict the optimization to the algebra $\mathcal{A}_{SDP}$, in the sense of the following theorem.

Theorem 2. *If the primal SDP problem (1) and its dual problem (2) meet the Slater condition, then there exists an optimal primal-dual pair* $(X, S) \in \mathcal{A}_{SDP} \times \mathcal{A}_{SDP}$.

Proof. For a given $\mu > 0$ the system:

$$\begin{aligned} \operatorname{trace}(A_k X) &= b_k \ \ (k = 1, \ldots, m) \\ \sum_{i=1}^{m} y_i A_i + S &= A_0 \\ XS &= \mu I, \ X \succ 0, \ S \succ 0, \end{aligned}$$

has a unique solution, denoted by $(X(\mu), y(\mu), S(\mu))$ (under the Slater condition assumption). The set

$$\{(X(\mu), y(\mu), S(\mu)) \ : \ \mu > 0\}$$

is known as the primal-dual central path, and is an analytic curve with limit point (as $\mu \downarrow 0$) in the primal-dual optimal set.

Note that

$$S(\mu) = A_0 - \sum_{i=1}^{m} y_i(\mu) A_i \in \mathcal{A}_{SDP}$$

and that $X(\mu) = \mu S(\mu)^{-1}$. Since $S(\mu) \in \mathcal{A}_{SDP}$ one also has $S(\mu)^{-1} \in \mathcal{A}_{SDP}$, since any matrix $*$-algebra is closed under taking inverses.

Consequently $X(\mu) \in \mathcal{A}_{SDP}$, and taking the limit as $\mu \downarrow 0$ completes the proof. □

The fact that the primal central path is contained in $\mathcal{A}_{SDP}$ was shown by Kanno et al. [27] for the case where $\mathcal{A}_{SDP}$ is the commuting algebra (centralizer ring) of a finite group (see Section 4.4).

It is also possible to drop the Slater assumption in Theorem 2. To be precise, if (1) has an optimal solution, then it has an optimal solution in $\mathcal{A}_{SDP}$. The proof is more elaborate than that of Theorem 2, but essentially follows from Theorem 3 in the next section.

4.1 Representations of matrix *-algebras

Matrix $*$-algebras have a canonical block-diagonal structure after a suitable 'coordinate transformation' (unitary transform). The details are a consequence of a theorem from 1907 by the algebraist Wedderburn.

Theorem 3 (Wedderburn (1907) [51, 52])**.** *Assume* $\mathcal{A} \subset \mathbb{C}^{n\times n}$ *is a matrix *-algebra over* $\mathbb{C}$ *that contains* I*. Then there is a unitary matrix* Q *and some integer* s *such that*

$$Q^* \mathcal{A} Q = \begin{pmatrix} \mathcal{A}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{A}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathcal{A}_s \end{pmatrix},$$

where each $\mathcal{A}_i$ is isomorphic to $\mathbb{C}^{n_i \times n_i}$ for some integers n_i, and takes the form

$$\mathcal{A}_i = \left\{ I_{k_i} \otimes A \mid A \in \mathbb{C}^{n_i \times n_i} \right\} \quad (i = 1, \ldots, s),$$

for some integers k_i $(i = 1, \ldots, s)$. Thus, one has $dim(\mathcal{A}) = \sum_{i=1}^{s} n_i^2$ and $n = \sum_{i=1}^{s} k_i n_i$.

This block-diagonal decomposition is canonical in the sense that it is unique up to the ordering of the blocks.

Interior point algorithms can exploit block diagonal structure of matrix variables.

For many matrix $*$-algebras that arise in applications, the canonical decomposition in Theorem 3 is known. One simple, but important example is the set of *circulant matrices.*

Example 4. *An $n \times n$ circulant matrix is defined by n numbers $c_0, \ldots, c_{n-1}$, and each row is a cyclic shift of the previous row:*

$$\begin{bmatrix} c_0 & c_1 & c_2 & \cdots & & c_{n-1} \\ c_{n-1} & c_0 & c_1 & & \cdots & \\ & c_{n-1} & c_0 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & c_1 \\ c_1 & & \cdots & & c_{n-1} & c_0 \end{bmatrix}.$$

The $n \times n$ circulant matrices form a commutative *matrix $*$-algebra of dimension n (over $\mathbb{R}$ or $\mathbb{C}$), that may be* diagonalized *by the unitary matrix:*

$$Q_{ij} := \frac{1}{\sqrt{n}} e^{-2\pi\sqrt{-1}ij/n} \quad (i, j = 0, \ldots, n-1).$$

The matrix Q is known as the discrete Fourier transform matrix*. Likewise, the real, symmetric circulant matrices form a matrix $*$-algebra over $\mathbb{R}$ of dimension $\lfloor n/2 \rfloor$; for more information on circulant matrices, see the review [19].*

If the canonical decomposition of $\mathcal{A}$ is not known, it may be obtained using (numerical) linear algebra, since the proof of Theorem 3 is constructive; see e.g. the proof in §2 of [14].

Randomized algorithms for the numerical decomposition of matrix $*$-algebras are described in [11] and [27]. One should note, however, that in many applications the order n of the matrices in $\mathcal{A}_{SDP}$ is too large to perform numerical linear algebra. In these cases a theoretical understanding of the structure of $\mathcal{A}_{SDP}$ is essential.

The regular *-representation

In general we do not know the unitary matrix Q in Theorem 3 that gives the canonical decomposition of $\mathcal{A}$, and the matrix size n may be too large to compute

Q using linear algebra. In these cases we may use other faithful representations of $\mathcal{A}$. One such faithful representation is the *regular $*$-representation* of $\mathcal{A}$.

Assume now that $B_1, \ldots, B_d \in \mathbb{R}^{n\times n}$ is a real, orthogonal basis of $\mathcal{A}$, seen as a matrix $*$-algebra over $\mathbb{R}$. We normalize this basis with respect to the *Frobenius norm*:

$$D_i := \frac{1}{\sqrt{\mathrm{trace}(B_i^T B_i)}} B_i \quad (i = 1, \ldots, d),$$

and define multiplication parameters $\gamma_{i,j}^k$ via:

$$D_i D_j = \sum_k \gamma_{i,j}^k D_k,$$

and subsequently define the $d \times d$ matrices L_k $(k = 1, \ldots, d)$ via

$$(L_k)_{ij} = \gamma_{k,j}^i \qquad (i, j = 1, \ldots, d).$$

The matrices L_k form a basis of a faithful (i.e. isomorphic) representation of $\mathcal{A}$, say $\mathcal{A}^{reg}$, that is also a matrix $*$-algebra, called the *regular $*$-representation of $\mathcal{A}$.*

Theorem 4 (cf. [23])**.** *The bijective linear mapping $\phi : \mathcal{A} \mapsto \mathcal{A}^{reg}$ such that $\phi(D_i) = L_i$ $(i = 1, \ldots, d)$ defines a $*$-isomorphism from $\mathcal{A}$ to $\mathcal{A}^{reg}$. Thus, ϕ is an algebra isomorphism with the additional property*

$$\phi(A^*) = \phi(A)^* \quad \forall A \in \mathcal{A}.$$

Since ϕ is a homomorphism, A and $\phi(A)$ have the same eigenvalues (up to multiplicities) for all $A \in \mathcal{A}$. As a consequence, one has

$$\sum_{i=1}^d x_i D_i \succeq 0 \Longleftrightarrow \sum_{i=1}^d x_i L_i \succeq 0.$$

It follows that we may work with $d \times d$ data matrices as opposed to $n \times n$.

Of course, this process requires the numerical calculation of the multiplication parameters $\gamma_{i,j}^k$, but this may often be done without storing the basis matrices $B_1, \ldots, B_d$ (see Section 5.4 for such an example). Thus, one may avoid calculations involving $n \times n$ matrices.

4.2 Symmetry reduction of SDP instances

By Theorem 2, we may rewrite the primal SDP problem (1) as:

$$\min_{X\succeq 0} \left\{ \mathrm{trace}(A_0 X) \;:\; \mathrm{trace}(A_k X) = b_k \quad (k = 1, \ldots, m),\; X \in \mathcal{A}_{SDP} \right\}.$$

Assume now that we have a basis $B_1, \ldots, B_d$ of $\mathcal{A}_{SDP}$. We set $X = \sum_{i=1}^d x_i B_i$ to get

$$\begin{aligned} &\min_{X\succeq 0} \left\{ \mathrm{trace}(A_0 X) \;:\; \mathrm{trace}(A_k X) = b_k \quad (k = 1, \ldots, m), X \in \mathcal{A}_{SDP} \right\} \\ = &\min_{\sum_{i=1}^d x_i B_i \succeq 0} \left\{ \sum_{i=1}^d x_i \mathrm{trace}(A_0 B_i) \;:\; \sum_{i=1}^d x_i \mathrm{trace}(A_k B_i) = b_k, (k = 1, \ldots, m) \right\}. \end{aligned}$$

We may now replace the linear matrix inequality (LMI) by $\sum_{i=1}^d x_i Q^* B_i Q \succeq 0$ to get block-diagonal structure, where Q is as described in Theorem 3, for $\mathcal{A} = \mathcal{A}_{SDP}$. Subsequently, we may delete any identical copies of blocks in the block structure to obtain a final reformulation.

Note that, even if the data matrices A_i are real symmetric, the final block diagonal matrices may in principle be complex Hermitian matrices, since Q may be unitary (as opposed to real orthogonal). This poses no problem in theory, since interior point methods apply to SDP with Hermitian data matrices as well. For example, the software SeDuMi [44] can handle Hermitian data matrices. If required, one may reformulate the LMI in terms of real matrices by applying the relation

$$A \succeq 0 \Longleftrightarrow \begin{bmatrix} \mathrm{Re}(A) & \mathrm{Im}(A)^T \\ \mathrm{Im}(A) & \mathrm{Re}(A) \end{bmatrix} \succeq 0 \quad (A = A^* \in \mathbb{C}^{n \times n})$$

to each block in the LMI. Note that this doubles the size of the block.

If the unitary matrix Q of the canonical decomposition in Theorem 3 is not available, one may use the regular $*$-representation of $\mathcal{A}_{SDP}$. In particular, the LMI $\sum_{i=1}^d x_i B_i \succeq 0$ is then replaced by an LMI of the form $\sum_{i=1}^d x_i L_i \succeq 0$, where the L_i's are $d \times d$ matrices (as opposed to $n \times n$) that are constructed as described in Section 4.1.

Further reading

Surveys on certain aspects of symmetry reduction in SDP are given by Gatermann and Parrilo [12], and Vallentin [48].

4.3 Coherent configurations and nonnegative variables

A basis $B_1, \ldots, B_d$ of a matrix *-algebra is called a *coherent configuration* if:

- The B_i's are 0-1 matrices;
- For each i, $B_i^T = B_{i^*}$ for some $i^* \in \{1, \ldots, d\}$;
- $\sum_{i=1}^d B_i = J$ (the all-ones matrix).

If the B_i's also commute, and $B_1 = I$, then we speak of an *association scheme*; see e.g. Godsil [16]. If $\mathcal{A}_{SDP}$ is spanned by an association scheme, then the SDP problem (1) reduces to an LP problem; see [18] for more details.

As another consequence, if $\mathcal{A}_{SDP}$ is spanned by a coherent configuration over the reals, and $X = \sum_{i=1}^d x_i B_i$, then $X \geq 0$ is equivalent to $x \geq 0$. Thus, additional nonnegativity of the matrix variable X requires only d additional linear inequality constraints to be added to the final SDP problem, as opposed to $\binom{n+1}{2}$. This is a very useful observation in several applications to be described later, where typically $d \ll n$.

4.4 Algebraic symmetry from permutation groups

Let $\mathcal{S}_n$ denote the symmetric group on n elements, i.e. the group of all permutations of $\{1, \ldots, n\}$.

We may represent any sub-group $\mathcal{G} \subseteq \mathcal{S}_n$ as a multiplicative group of $n \times n$ permutation matrices via

$$(P_\pi)_{i,j} := \begin{cases} 1 & \text{if } \pi(i) = j \\ 0 & \text{else.} \end{cases} \qquad \pi \in \mathcal{G},\ i, j = 1, \ldots, n.$$

The *commutant* of the representation is defined as

$$\{A \in \mathbb{C}^{n \times n} \ : \ AP_\pi = P_\pi A \quad \forall \pi \in \mathcal{G}\},$$

and forms a matrix $*$-algebra over $\mathbb{C}$, as is easy to show. This commutant is also called the *centralizer ring* of the group, or the *commuting algebra* of the group.

The commutant of the representation has a basis that is a coherent configuration. (A coherent configuration that arises in this way is sometimes called *Schurian* [8].) One may construct this 0-1 basis of the commutant from the *orbitals* of $\mathcal{G}$.

Definition 3. *The* two-orbit or orbital of an index pair (i, j) is defined as

$$\{(\pi(i), \pi(j)) \ : \ \pi \in \mathcal{G}\}.$$

The orbitals partition $\{1, \ldots, n\} \times \{1, \ldots, n\}$ and this partition yields the $0-1$ matrices of the coherent configuration in question.

Now consider the situation that the data matrices $A_0, \ldots, A_m$ are invariant under the action of some permutation group $\mathcal{G}_{SDP}$ in the sense that

$$(A_k)_{ij} = (A_k)_{\pi(i),\pi(j)} \quad \forall \pi \in \mathcal{G}_{SDP},\ k \in \{0, \ldots, m\},\ i, j \in \{1, \ldots, n\}.$$

This is equivalent to

$$P_\pi A_k P_\pi^T = A_k \quad \forall \pi \in \mathcal{G}_{SDP},\ k \in \{0, \ldots, m\}$$

or

$$P_\pi A_k = A_k P_\pi \quad \forall \pi \in \mathcal{G}_{SDP},\ k \in \{0, \ldots, m\}.$$

In other words, the data matrices belong to the commutant of the permutation matrix representation of $\mathcal{G}_{SDP}$, and we may define $\mathcal{A}_{SDP}$ as this commutant.

An equivalent way to describe this situation via the automorphism groups of the data matrices.

Definition 4. *We define the automorphism group of a given $A \in \mathbb{C}^{n \times n}$ as*

$$Aut(A) := \{\pi \in \mathcal{S}_n \ : \ A_{ij} = A_{\pi(i),\pi(j)}\ \forall\, i, j\}.$$

Thus, $\mathcal{G}_{SDP}$ may be defined as the intersection $\mathcal{G}_{SDP} = \bigcap_{i=0}^m \mathrm{Aut}(A_i)$.

Example 5. *We consider a fixed ordering $v_1, \ldots, v_{2^n}$ of $\{0,1\}^n$, i.e. of the binary vectors of length n. Now consider the matrix A with 2^n rows indexed by the elements of $\{0,1\}^n$, and A_{ij} given by the Hamming distance between $v_i \in \{0,1\}^n$ and $v_j \in \{0,1\}^n$. (Recall that the Hamming distance between two n-vectors is the number of positions where they differ.)*

The automorphism group of A arises as follows. Any permutation π of the index set $\{1, \ldots, n\}$ induces an isomorphism of A that maps row (resp. column) i of A to row (resp. column) $\pi(i)$ for all i. There are $n!$ such permutations. Moreover, there are an additional 2^n permutations that act on $\{0,1\}^n$ by either 'flipping' a given component from zero to one (and vice versa), or not.

Thus, Aut(A) has order $n!2^n$. The centralizer ring of Aut(A) is a commutative matrix $*$-algebra over $\mathbb{R}$ and is known as the Bose-Mesner algebra of the Hamming scheme.

A basis for the centralizer ring may be derived from the orbitals of Aut(A) and are given by

$$B_{ij}^{(k)} = \begin{cases} 1 & \text{if Hamming}(v_i, v_j) = k; \\ 0 & \text{else} \end{cases} \qquad (k = 0, \ldots, n),$$

where Hamming(v_i, v_j) is the Hamming distance between v_i and v_j. The basis matrices $B^{(k)}$ are simultaneously diagonalized by the real, orthogonal matrix Q defined by

$$Q_{ij} = 2^{-\frac{n}{2}}(-1)^{i^T j} \quad i, j = 1, \ldots, 2^n.$$

The distinct elements of the matrix $Q^T B^{(k)} Q$ equal $K_j(k)$ $(j = 0, \ldots, n)$, where

$$K_j(x) := \sum_{k=0}^{j} (-1)^k \binom{x}{k} \binom{n-x}{j-k}, \quad j = 0, \ldots, n,$$

are called Krawtchouk polynomials. Thus, a linear matrix inequality of the form

$$\sum_{k=0}^{n} x_k B^{(k)} \succeq 0$$

is equivalent to the system of linear inequalities

$$\sum_{k=0}^{n} x_k K_j(k) \geq 0 \qquad (j = 0, \ldots, n).$$

Further reading

The canonical decomposition of the centralizer ring of a group is determined by the so-called irreducible representations of the group. If all irreducible representations are known, then the canonical decomposition is also available. The details are beyond the scope of this survey, but are reviewed in §4 of [4], with reference to the standard work on finite group representation theory of Serre [43], §13.2.

5 Applications of symmetry reduction

In this section we will consider examples of symmetry reduction for SDP instances from various sources, namely:

- The Lovász ϑ-number for graphs;
- Error correcting binary codes;
- Kissing numbers and related problems;
- Crossing numbers of completely bipartite graphs;
- Quadratic assignment problems;
- Truss topology design.

5.1 The Lovász ϑ-number for graphs

The ϑ number of a given graph $G = (V, E)$ may be defined as the optimal value of the following semidefinite program:

$$\vartheta(G) := \max \operatorname{trace}(JX)$$

subject to

$$\begin{array}{rcl} X_{ij} & = & 0, \ \{i,j\} \in E \ (i \neq j) \\ \operatorname{trace}(X) & = & 1 \\ X & \succeq & 0. \end{array}$$

The many interesting properties of the ϑ number are surveyed in [29]. One of these properties is known as the *sandwich theorem.*

Theorem 5 (Lovász 'sandwich theorem' [35])**.** *Let $G = (V, E)$ be a graph with stability number $\alpha(G)$, and let $\bar{G}$ denote its complementary graph. One has*

$$\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G}),$$

where $\chi(\bar{G})$ denotes the chromatic number of $\bar{G}$ (the minimum number of colors required to color the vertices of $\bar{G}$ such that adjacent vertices have different colors).

For example, for the Pentagon graph (5-cycle), usually denoted by C_5, one has

$$2 \equiv \alpha(C_5) \leq \vartheta(C_5) \equiv \sqrt{5} \leq \chi(\bar{C}_5) \equiv 3.$$

Thus, one may view $\vartheta(G)$ as an approximation of $\alpha(G)$ or of $\chi(\bar{G})$.

The associated ϑ' number of G is obtained by adding nonnegativity constraints $X \geq 0$ to the formulation of the ϑ SDP problem.

An equivalent formulation for $\vartheta'(G)$ is:

$$\vartheta'(G) = \max_{X \succeq 0, X \geq 0} \{\operatorname{trace}(JX) \mid \operatorname{trace}(A + I)X = 1\}, \tag{7}$$

where A is the adjacency matrix of the graph G. Thus, here $\mathcal{G}_{SDP}$ is simply the automorphism group $\text{Aut}(G)$ of the graph G, since the data matrices of the SDP problem (7) are simply J and $A + I$, that are invariant under $\text{Aut}(G)$.

Thus, for graphs G where the commutant of $\text{Aut}(G)$ is low dimensional, the SDP problem (7) may be reduced in size via symmetry reduction. This has been done in the literature for graphs on the hypercube (see the next section), as well as the so-called Erdös-Rényi graphs [26].

5.2 Upper bounds on binary code sizes

The Hamming graph $G(k, \delta)$ has vertices indexed by $\{0,1\}^k$ and vertices adjacent if they are at Hamming distance less than δ (see Figure 1 for an example).

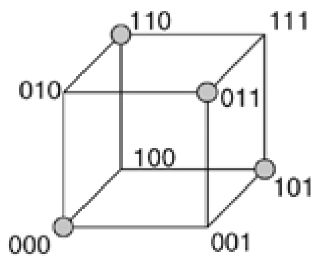

Figure 1: Hamming graph with $k = 3$ and $\delta = 2$. A maximum stable set of size 4 is shown, implying that $A(3, 2) = 4$.

In coding theory, the usual notation is $\alpha(G(k, \delta)) =: A(k, \delta)$. Thus, $A(k, \delta)$ is the maximum size of a binary code on k letters such that any two words are at a Hamming distance of at least δ.

For the Hamming graph, one has $|\text{Aut}(G(k, \delta))| = 2^k k!$, and the commutant of $\text{Aut}(G(k, \delta))$ is the commutative *Bose-Mesner algebra of the Hamming scheme* that has dimension $k+1$ (see Example 5 for the details). Thus, the SDP matrices may be reduced from the original size $n = 2^k$ to diagonal matrices of size $k + 1$.

The resulting LP problem coincides with the LP relaxation of Delsarte [10], as was first shown in a seminal paper by Schrijver [41] in 1979. In fact, the symmetry reduction methodology introduced in [41] was arguably even more interesting than the main result of the paper, and pre-dated the other applications described in this survey by more than twenty years!

In the paper [42], a stronger SDP bound for $A(k, \delta)$ is obtained via the following steps:

- a stronger SDP relaxation is constructed via lift-and-project such that some symmetry is retained in the resulting SDP.

- in the stronger relaxation, $\mathcal{A}_{SDP}$ becomes the Terwilliger algebra of the Hamming scheme, a non-commutative algebra that contains the Bose-Mesner algebra of the Hamming scheme as a sub-algebra. The Terwilliger algebra has dimension $\binom{k+3}{3}$ and its canonical block-diagonalization is described in [42].

Thus, improved upper bounds were computed for $A(19, 6)$, $A(23, 6)$, $A(25, 6)$, etc.; see [42] for more details.

Using other lift-and-project schemes, slightly better SDP bounds were obtained by Laurent [33] for some values of $A(k, \delta)$, and the approach in [42] was extended to non-binary codes by Gijswijt, Schrijver, and Tanaka [15].

5.3 SDP bounds on kissing numbers

The kissing number of $\mathbb{R}^k$ is defined as the maximum number of unit balls that can simultaneously touch a unit ball centered at the origin, without any overlap.

Thus, the kissing number of $\mathbb{R}^2$ is 6 and in $\mathbb{R}^3$ it is 12. (There was a famous disagreement between Newton and Gregory on whether the correct answer in $\mathbb{R}^3$ is 12 or 13).

Not much is known about kissing numbers for general values of k, and it is interesting to compute upper bounds for fixed k.

In a seminal paper by Bachoc and Vallentin [3], the authors introduce new SDP relaxations of this problem, and succeeded to compute improved upper bounds on the kissing number in the dimensions $n = 5, 6, 7, 9$ and 10.

Although, the details are beyond the scope of this survey, the basic methodology is as follows.

The kissing number problem may first be formulated as a maximum stable set problem in an infinite graph. The vertices of the graph in question are the points on the unit sphere in $\mathbb{R}^k$, and two vertices adjacent if the angle between them is below a certain value (see Figure 2).

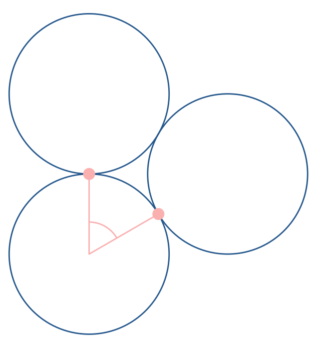

Figure 2: The bottom sphere represents a infinite vertex set, with two vertices adjacent if the angle between them are at most $\pi/3$.

The rough idea is to generalize certain SDP relaxations of the maximum stable set problem to infinite graphs, and subsequently exploit the symmetry of the sphere (orthogonal group) to obtain a (finite) SDP relaxation. This analysis requires tools from harmonic analysis and the representation theory for compact topological groups. The interested reader may find the necessary background in the lecture notes [49].

5.4 Bounding the crossing number of complete bipartite graphs

The *crossing number* $\mathrm{cr}(G)$ of a graph G is the minimum number of intersections of edges (at a point other than a vertex) in a drawing of G in the plane.

Paul Turán first raised the problem of calculating the crossing number of the complete bipartite graph $K_{r,s}$ in the 1940's. $K_{r,s}$ may be drawn in the plane with at most $Z(r,s)$ edges crossing, where

$$Z(r,s) = \left\lfloor \frac{r-1}{2} \right\rfloor \left\lfloor \frac{r}{2} \right\rfloor \left\lfloor \frac{s-1}{2} \right\rfloor \left\lfloor \frac{s}{2} \right\rfloor. \tag{8}$$

For example, Figure 3 shows a drawing of $K_{4,5}$ with 8 crossings.

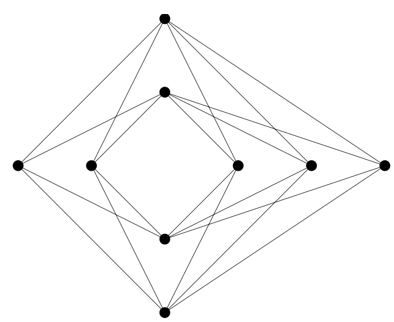

Figure 3: A drawing of $K_{4,5}$ with 8 crossings. A similar strategy can be used to construct drawings of $K_{r,s}$ with exactly $Z(r,s)$ crossings.

The *Zarankiewicz's conjecture* is that equality holds, i.e:

$$\mathrm{cr}(K_{r,s}) = Z(r,s).$$

This has been verified for $\min\{r,s\} = 5$, but remains a famous open problem for most other values of r and s.

De Klerk et al. [22] showed that one may obtain a lower bound on $\mathrm{cr}(K_{r,s})$ via the optimal value of a suitable SDP problem, namely:

$$\mathrm{cr}(K_{r,s}) \geq \frac{s}{2}\left(s \min_{X \geq 0,\ X \succeq 0} \{\mathrm{trace}(QX) \ \mid \ \mathrm{trace}(JX) = 1\} - \left\lfloor \frac{r}{2} \right\rfloor \left\lfloor \frac{r-1}{2} \right\rfloor \right),$$

where Q is a certain (given) matrix of order $(r-1)!$, and J is the all-ones matrix of the same size. The rows and columns of Q are indexed by all the cyclic orderings of r elements, which we denote by $u_1, \ldots, u_{(r-1)!}$. Entry Q_{ij}

in the matrix Q is the distance between the cyclic orderings u_i and u_j. (The distance between two cyclic orderings is the number of neighbor swaps needed to go from one to the other; thus, the distance between 123 and 132 is one, for example.)

It follows that $\mathcal{G}_{SDP}$ is a certain representation of $\mathcal{S}_2 \times \mathcal{S}_r$, so that $|\mathcal{G}_{SDP}| = 2(r!)$ For this example the canonical decomposition of the centralizer ring of $\mathcal{G}_{SDP}$ is not known in closed form.

De Klerk and al. [22] solved the SDP problem for $r = 7$ by using partial symmetry reduction, to obtain the bound:

$$\mathrm{cr}(K_{7,s}) > 2.1796s^2 - 4.5s.$$

Using an averaging argument, the bound for $\mathrm{cr}(K_{7,s})$ implies the following asymptotic bound on $\mathrm{cr}(K_{r,s})$:

$$\lim_{s\to\infty} \frac{\mathrm{cr}(K_{r,s})}{Z(r,s)} \geq 0.83\frac{r}{r-1}.$$

Thus, loosely speaking, asymptotically, $Z(r,s)$ and $\mathrm{cr}(K_{r,s})$ do not differ by more than 17%.

In subsequent, related work, De Klerk et al. [23] improved the constant 0.83 to 0.859 by solving the SDP for $r = 9$. This was possible by using the regular $*$-representation of the commutant of $\mathcal{G}_{SDP}$, as described in Section 4.1.

5.5 SDP relaxation of the quadratic assignment problem

The quadratic assignment problem (QAP) may be stated in the following form:

$$\min_{X\in\Pi_k} \mathrm{trace}(AXBX^T), \tag{9}$$

where A and B are given symmetric $k \times k$ matrices (called the *distance* and *flow* matrices, respectively), and Π_k is the set of $k \times k$ permutation matrices.

The QAP has many applications in facility location, circuit design, graph isomorphism and other problems, but is NP-hard in the strong sense, and hard to solve in practice for $k \geq 30$; for a review, see Anstreicher [2].

An SDP relxation of (QAP) from [54] and [25] takes the form:

$$\left.\begin{array}{rcl}
\min \mathrm{trace}(B\otimes A)Y & & \\
\text{subject to} & & \\
\mathrm{trace}((I\otimes(J-I))Y + ((J-I)\otimes I)Y) & = & 0 \\
\mathrm{trace}(Y) - 2e^T y & = & -k \\
\begin{pmatrix} 1 & y^T \\ y & Y \end{pmatrix} & \succeq & 0, \\
Y & \geq & 0.
\end{array}\right\} \tag{10}$$

It is easy to verify that this is indeed a relaxation of QAP, by noting that setting $Y = \mathrm{vec}(X)\mathrm{vec}(X)^T$ and $y = \mathrm{diag}(Y)$ gives a feasible solution if $X \in \Pi_k$.

instance	k	previous l.b.	SDP l.b. (10)	best known u.b.	CPU time(s)
esc64a	64	47	98	116	13
esc128a	128	2	54	64	140

Table 1: SDP lower bounds for two QAP instances from QAPlib. l.b. = lower bound, u.b. = upper bound

These SDP problems are hard to solve even for small values of k, since the matrix variable Y is nonnegative and of order k^2.

The automorphism groups of A and B determine the SDP symmetry group $\mathcal{G}_{SDP}$. Thus, symmetry reduction is possible if $\mathrm{Aut}(A)$ and/or $\mathrm{Aut}(B)$ is large. Several instances in the QAPlib library [7] of QAP instances have algebraic symmetry, e.g. the distance matrix is a Hamming distance matrix (see Example 5).

Some numerical results, after doing the SDP symmetry reduction are shown in Table 1. These results are taken from [25], and involve QAP instances with Hamming distance matrices. We emphasize that, for the problems in Table 1, the SDP relaxation (10) is too large even to solve by lower order methods. After symmetry reduction, however, they may be solved using interior point methods in a few seconds on a standard Pentium IV PC.

The traveling salesman problem

It is well-known that the QAP contains the symmetric traveling salesman problem (TSP) as a special case. To show this, we denote the complete graph on n vertices with edge lengths $D_{ij} = D_{ji} > 0$ $(i \neq j)$, by $K_n(D)$, where D is called the matrix of edge lengths. The TSP is to find a Hamiltonian circuit of minimum length in $K_n(D)$.

To see that TSP is a special case of QAP, let C denote the adjacency matrix of the standard circuit on n vertices:

$$C := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & & 0 & 1 \\ 1 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

Now the TSP problem is obtained from the QAP problem (9) by setting $A = \frac{1}{2}D$ and $B = C$. To see this, note that every Hamiltonian circuit in a complete graph has adjacency matrix XCX^T for some $X \in \Pi_n$. Thus, we may concisely state the TSP as

$$\min_{X \in \Pi_n} \mathrm{trace}\left(\frac{1}{2} DXCX^T\right).$$

The matrix C is symmetric circulant, and therefore belongs to a matrix $*$-algebra of dimension $d = \lfloor \frac{1}{2}n \rfloor$ (see Example 4). By exploiting this symmetry, De Klerk et al. [24] showed that the SDP relaxation (10) reduces to the following problem in the special case of TSP.

$$\min \frac{1}{2} \operatorname{trace}\left(DX^{(1)}\right)$$

subject to

$$\left.\begin{array}{rcll} X^{(k)} & \geq & 0, & k = 1, \ldots, d \\ \sum_{k=1}^{d} X^{(k)} & = & J - I, & \\ I + \sum_{k=1}^{d} \cos\left(\frac{2\pi i k}{n}\right) X^{(k)} & \succeq & 0, & i = 1, \ldots, d \\ X^{(k)} & \in & \mathbb{S}^{n\times n}, & k = 1, \ldots, d, \end{array}\right\} \qquad (11)$$

where $d = \lfloor \frac{1}{2}n \rfloor$. Note that this problem only involves matrix variables $X^{(1)}, \ldots, X^{(d)}$ of order n as opposed to the matrix variable of order n^2 in (10), i.e. the problem size is reduced by a factor n in this sense.

In addition to the size reduction, the variables of the reduced problem have an interesting interpretation in terms of association schemes.

In particular, one may construct a feasible solution by setting:

$$X_{ij}^{(k)} = \begin{cases} 1 & \text{if } \operatorname{dist}(i,j) = k \\ 0 & \text{else,} \end{cases} \qquad (i, j = 1, \ldots, n),\ k = 0, \ldots, d,$$

where $\operatorname{dist}(i,j)$ is the length of the shortest path from i to j in the minimum length Hamiltonian cycle. It is well-known (see e.g. Chapter 12 in [16]), that the matrices constructed in this way form an association scheme, since a cycle is a distance regular graph.

Thus, the symmetry reduction process may also provide unexpected theoretical insight into a specific SDP problem.

5.6 Truss topology design problems

A truss structure (like the Eiffel tower) is defined by a ground structure of nodes and bars. Letting t denote the number of bars, and $\ell \in \mathbb{R}^t$ the vector of bar lengths, a truss topology design problem is to find a vector $z \in \mathbb{R}^t$ of cross-sectional areas of the bars, such that some objective is optimized.

A specific topology optimization problem, introduced in [21], is to find a truss of minimum volume such that the fundamental frequency of vibration is higher than some prescribed critical value:

$$\begin{array}{lll} \text{(TOP)} & \min & \sum\limits_{i=1}^{t} \ell_i z_i \\ & \text{s.t.} & S = \sum_{i=1}^{t} (K_i - \bar{\Omega} M_i) z_i - \bar{\Omega} M_0 \\ & & z_i \geq 0 \quad i = 1, \ldots, t \\ & & S \succeq 0, \end{array} \qquad (12)$$

where $\bar{\Omega}$ is a lower bound on the (squared) fundamental frequency of vibration of the truss, and M_0 is the so-called non-structural mass matrix. The matrices $x_i K_i$ and $x_i M_i$ are called the *stiffness* and *mass* matrices of bar i, respectively. The order of these matrices is the number of free nodes in the structure times the degrees of freedom.

If the ground structure of nodes and bars has isometries, then the SDP problem has algebraic symmetry that may be exploited. An example of such a truss from [27] is shown in Figure 4.

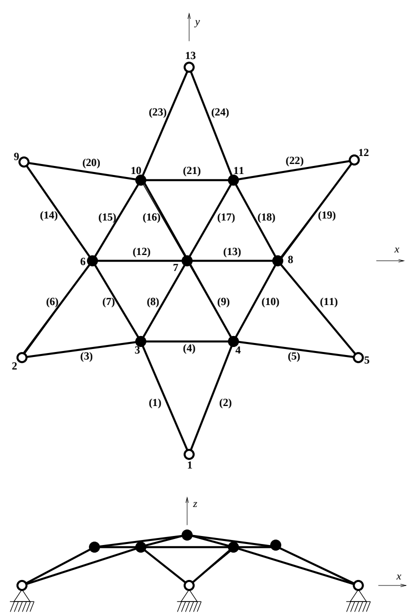

Figure 4: Top and side views of a spherical lattice dome with D_6 symmetry. The black nodes are free and the white nodes fixed.

The ground structure of nodes in Figure 4 has the same symmetry as a hexagon, i.e. the associated group is the dihedral group D_6.

To understand the size reduction that is possible for ground structures of this type, consider the generalization where the hexagon in the figure is replaced by a regular k-gon, so that the symmetry group becomes D_k. Bai et al. [4] have shown that the SDP problem (12) reduces to one with only 3 scalar variables, and block diagonal matrix variables with block sizes at most 4, regardless of the value of k. Without symmetry reduction, the SDP problem (12) has $3k$ scalar variables, and a matrix variable of order $3k$.

6 Summary and conclusion

There are three types of structure in SDP data that may be exploited by interior point methods, namely low rank, chordal sparsity structure and algebraic symmetry. The first two are well-established and have been incorporated in some software packages.

The third (symmetry reduction in SDP) may best be described as the application of representation theory to reduce the size of specially structured SDP instances. The most notable applications so far are in computer assisted proofs (bounds on crossing numbers, kissing numbers, code sizes, etc.), but also in pre-processing of some SDP problems arising in optimal design (truss design, QAP, etc.)

The symmetry reduction 'device' introduced in 1979 has become a method!

Acknowledgements

This survey grew from a plenary talk given at the 2008 SIAM conference on Optimization held in Boston, MA. The author would like to thank the organizers Kurt Anstreicher and Sven Leyffer for the opportunity to give that talk. The author would also like to thank Masakazu Kojima, for pointing out that the Slater assumption in Theorem 2 may be weakened, and Monique Laurent for pointing out an error in an earlier version of this paper.

References

[1] F. Alizadeh. *Combinatorial optimization with interior point methods and semi–definite matrices.* PhD thesis, University of Minnesota, Minneapolis, USA, 1991.

[2] K.M. Anstreicher. Recent advances in the solution of quadratic assignment problems. *Mathematical Programming, Ser. B*, **97**:27–42, 2003.

[3] C. Bachoc and F. Vallentin. New upper bounds for kissing numbers from semidefinite programming. *Journal of the AMS*, **21**:909-924, 2008.

[4] Y.-Q. Bai, E. de Klerk, D.V. Pasechnik, and R. Sotirov. Exploiting group symmetry in truss topology optimization. *Optimization and Engineering* (to appear).

[5] R.E. Bellman and K. Fan. On systems of linear inequalities in Hermitian matrix variables. In V.L. Klee, editor, *Convexity*, Vol. 7 Proc. Symposia in Pure Mathematics, pages 1–11. Amer. Math. Soc., Providence, RI, 1963.

[6] S. J. Benson, Y. Ye, and X. Zhang. Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM J. Optim.*, **10**(2):443–461, 2000.

[7] R.E. Burkard, S.E. Karisch, and F. Rendl. QAPLIB — a quadratic assignment problem library. *Journal on Global Optimization*, **10**:291–403, 1997; see also `http://www.seas.upenn.edu/qaplib/`.

[8] P.J. Cameron. Coherent configurations, association schemes, and permutation groups, pp. 55-71 in Groups, Combinatorics and Geometry (A. A. Ivanov, M. W. Liebeck and J. Saxl eds.), World Scientific, Singapore, 2003.

[9] J. Dahl, L. Vandenberghe, V. Roychowdhury. Covariance selection for non-chordal graphs via chordal embedding. *Optimization Methods and Software*, **23**(4):501–520, 2008.

[10] P. Delsarte, Bounds for unrestricted codes, by linear programming, Philips Research Report 27, 272–289, 1972.

[11] W. Eberly and M. Giesbrecht. Efficient decomposition of separable algebras. *Journal of Symbolic Computation*, 37(1):35–81, 2004.

[12] K. Gatermann and P.A. Parrilo, Symmetry groups, semidefinite programs, and sums of squares. *J. Pure and Applied Algebra*, **192**:95–128, 2004.

[13] A.J. George and J. W. H. Liu. The evolution of the minimum degree ordering algorithm, *SIAM Review*, **31**:1–19, 1989.

[14] D. Gijswijt. Matrix Algebras and Semidefinite Programming Techniques for Codes. PhD thesis, University of Amsterdam, The Netherlands, 2005. `http://staff.science.uva.nl/~gijswijt/promotie/thesis.pdf`

[15] D. Gijswijt, A. Schrijver, H. Tanaka, New upper bounds for nonbinary codes based on the Terwilliger algebra and semidefinite programming, *Journal of Combinatorial Theory, Series A*, **113**:1719–1731, 2006.

[16] C. Godsil. *Algebraic Combinatorics*. Chapman and Hall, 1993.

[17] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, **42**(6):1115–1145, 1995.

[18] M.X. Goemans and F. Rendl. Semidefinite Programs and Association Schemes. *Computing*, **63**:331–340, 1999.

[19] R.M. Gray. Toeplitz and Circulant Matrices: A review. *Foundations and Trends in Communications and Information Theory*, **2**(3):155–239, 2006.

[20] R. Grone, C.R. Johnson, E.M. Sá, and H. Wolkowicz. Positive definite completions of partial Hermitian matrices. *Linear Algebra and its Applications*, **58**:109–124, 1984.

[21] Y. Kanno, M. Ohsaki, K. Murota and N. Katoh. Group symmetry in interior-point methods for semidefinite program, *Optimization and Engineering*, **2**(3): 293–320, 2001.

[22] E. de Klerk, J. Maharry, D.V. Pasechnik, B. Richter and G. Salazar. Improved bounds for the crossing numbers of $K_{m,n}$ and K_n. *SIAM Journal on Discrete Mathematics* **20**:189–202, 2006.

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

[24] E. de Klerk, D.V. Pasechnik and R. Sotirov. On semidefinite programming relaxations of the traveling salesman problem. *SIAM Journal on Optimization*, to appear.

[25] E. de Klerk and R. Sotirov. Exploiting Group Symmetry in Semidefinite Programming Relaxations of the Quadratic Assignment Problem. *Mathematical Programming, Series A*, to appear.

[26] E. de Klerk, M.W. Newman, D.V. Pasechnik, and R. Sotirov. On the Lovász ϑ-number of almost regular graphs with application to Erdös-Rényi graphs. *European Journal of Combinatorics*, to appear.

[27] Y. Kanno, M. Ohsaki, K. Murota and N. Katoh, Group symmetry in interior-point methods for semidefinite programming, *Optimization and Engineering*, **2**(3): 293–320, 2001.

[28] L. Khachiyan. A polynomial time algorithm in linear programming. *Soviet Mathematics Doklady*, **20**:191–194, 1979.

[29] D.E. Knuth, The sandwich theorem, *Electron. J. Combin.*, **1**, Article 1, approx. 48 pp. (electronic), 1994.

[30] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J.Optim.*, **11**(3):796–817, 2001.

[31] M. Laurent. A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems. In P. Pardalos and H. Wolkowicz, editors, *Topics in Semidefinite and Interior-Point Methods*, Volume 18 of The Fields Institute for Research in Mathematical Science, Communications Series. Providence, Rhode Island, pages 51–76, 1998.

[32] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, Volume 149 of IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivant (eds.), Springer, 2008.

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

[34] J. Löfberg and P.A. Parrilo, From coefficients to samples: a new approach to SOS optimization, IEEE Conference on Decision and Control, 2004.

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

[36] K. Murota, Y. Kanno, M. Kojima and S. Kojima, A Numerical Algorithm for Block-Diagonal Decomposition of Matrix *-Algebras, Technical report B-445, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, September 2007 (Revised June 2008).

[37] K. Nakata, K. Fujisawa, M. Fukuda, M. Kojima, and K. Murota. Exploiting sparsity in semidefinite programming via matrix completion II: Implementation and numerical results," *Mathematical Programming* Series B, **95**:303–327, 2003.

[38] A.S. Nemirovsky and D.B. Yudin. *Problem Complexity and Method Efficiency in Optimization.* (Wiley-Interscience, 1983).

[39] Yu.E. Nesterov and A.S. Nemirovski. *Interior point polynomial algorithms in convex programming.* SIAM Studies in Applied Mathematics, Vol. 13. SIAM, Philadelphia, USA, 1994.

[40] D. Rose, G. Lueker, R.E. Tarjan. "Algorithmic aspects of vertex elimination on graphs". *SIAM Journal on Computing*, **5**(2):266-283, 1976.

[41] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Transactions on Information Theory*, **25**:425–429, 1979.

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

[43] J.-P. Serre. *Linear representations of finite groups.* Graduate Texts in Mathematics, Vol. 42, Springer-Verlag, New York, 1977.

[44] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, **11-12**:625–653, 1999.

[45] M.J. Todd. A Study of Search Directions in Primal-Dual Interior-Point Methods for Semidefinite Programming. *Optimization Methods and Software*, **11**:1–46, 1999.

[46] M.J. Todd. Semidefinite optimization. Acta Numerica, **10**:515–560, 2001.

[47] K. C. Toh, M. J. Todd, and R. H. Tütüncü. SDPT3—a MATLAB software package for semidefinite programming, version 1.3. *Optimization Methods and Software*, **11/12**(1-4):545–581, 1999.

[48] F. Vallentin. Symmetry in semidefinite programs. *Linear Algebra and its Applications*, to appear. Preprint version: arXiv:0706.4233

[49] F. Vallentin. Lecture notes: Semidefinite programs and harmonic analysis. arXiv:0809.2017 preprint, 2008.

[50] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, **38**:49–95, 1996.

[51] J.H.M. Wedderburn. On hypercomplex numbers. *Proc. London Math. Soc.* **6**(2):77–118, 1907.

[52] J.H.M. Wedderburn. *Lectures on Matrices.* AMS Colloquium Publications, Volume 17, AMS publishers, 1934.

[53] M. Yannakakis. Computing the minimum fill-in is NP-complete, *SIAM Journal on Algebraic and Discrete Methods*, **2**(1):77–79, 1981.

[54] Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz. Semidefinite Programming Relaxations for the Quadratic Assignment Problem. *Journal of Combinatorial Optimization*, **2**:71–109, 1998.