# Invariant semidefinite programs

Christine Bachoc[1], Dion C. Gijswijt[2], Alexander Schrijver[3], and Frank Vallentin[4]

[1] Laboratoire A2X, Université Bordeaux I, 351, cours de la Libération, 33405 Talence, France bachoc@math.u-bordeaux1.fr
[2] CWI and Department of Mathematics, Leiden University; Centrum voor Wiskunde en Informatica (CWI), Sciencepark 123, 1098 XG Amsterdam, The Netherlands dion.gijswijt@gmail.com
[3] CWI and Department of Mathematics, University of Amsterdam; Centrum voor Wiskunde en Informatica (CWI), Sciencepark 123, 1098 XG Amsterdam, The Netherlands lex@cwi.nl
[4] Delft Institute of Applied Mathematics, Technical University of Delft, P.O. Box 5031, 2600 GA Delft, The Netherlands f.vallentin@tudelft.nl

## 1 Introduction

In the last years many results in the area of semidefinite programming were obtained for invariant semidefinite programs — semidefinite programs which have symmetries. This was done for a variety of problems and applications. The purpose of this handbook chapter is to give the reader the necessary background for dealing with semidefinite programs which have symmetry. Here the focus is on the basic theory and on representative examples. We do not aim at completeness of the presentation.

In all applications the underlying principles are similar: one simplifies the original semidefinite program which is invariant under a group action by applying an algebra isomorphism mapping a "large" matrix algebra to a "small" matrix algebra. Then it is sufficient to solve the semidefinite program using the smaller matrices.

We start this chapter by developing the general framework in the introduction where we give a step-by-step procedure for simplifying semidefinite programs: Especially Step 2 (first version), Step 2 (second version), and Step $1\frac{1}{2}$ will be relevant in the later discussion. Both versions of Step 2 are based on the main structure theorem for matrix $*$-algebras. Step $1\frac{1}{2}$ is based on the regular $*$-representation.

In Section 2 we give a proof of the main structure theorem for matrix $*$-algebras and we present the regular $*$-representation. Strictly speaking the framework of matrix $*$-algebras is slightly too general for the applications we have in mind. However, working with matrix $*$-algebras does not cause much extra work and it also gives a numerical algorithm for finding an explicit algebra isomorphism. Section 2 is mainly concerned with finite dimensional invariant semidefinite programs. In Section 3 we show how one can extend this to special classes of infinite dimensional invariant semidefinite programs, namely those which arise from permutation actions of compact groups. We focus on this case because of space limitations and because it suffices for our examples. This section is connected to Step 2 (second version) of the introduction.

The later sections contain examples coming from different areas: In Section 4 we consider finding upper bounds for finite error-correcting codes and in Section 5 we give lower bounds for the crossing number of complete bipartite graphs. Both applications are based on the methods explained in Section 2 (Step 2 (first version) and Step $1\frac{1}{2}$ in the introduction). In Section 6 we use Step 2 (second version) for finding upper bounds for spherical codes and other geometric packing problems on the sphere. For this application the background in Section 3 is relevant. Section 4 and Section 6 both use the theta number of Lovász for finding upper bounds for the independence number of highly symmetric graphs. In Section 7 we show how one can exploit symmetry in polynomial optimization: We give particular sum of squares representations of polynomials which have symmetry.

This list of applications is not complete, and many more applications can be found in the literature. In the last Section 8 we give literature pointers to more applications.

### 1.1 Complex semidefinite programs

In order to present the theory as simple as possible we work with complex semidefinite programs. We give the necessary definitions. A complex matrix $X \in \mathbb{C}^{n\times n}$ is *Hermitian* if $X = X^*$, where $X^*$ is the *conjugate transpose* of $X$, i.e. $X_{ij} = \overline{X_{ji}}$. It is called *positive semidefinite*, we write $X \succeq 0$, if for all (column) vectors $(\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$ we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i X_{ij} \overline{\alpha_j} \geq 0.$$

The space of complex matrices is equipped with a complex inner product, the trace product $\langle X, Y\rangle = \mathrm{trace}(Y^* X)$, which is linear in the first entry. The inner product of two Hermitian matrices is always real.

**Definition 1.1** *A* (complex) semidefinite program *is an optimization problem of the form*

$$\max\{\langle X, C\rangle : X \succeq 0, \langle X, A_1\rangle = b_1, \ldots, \langle X, A_m\rangle = b_m\}, \tag{1}$$

*where* $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$, *and* $C \in \mathbb{C}^{n\times n}$ *are given Hermitian matrices,* $(b_1, \ldots, b_m) \in \mathbb{R}^m$ *is a given vector and* $X \in \mathbb{C}^{n\times n}$ *is a variable Hermitian matrix.*

A Hermitian matrix $X \in \mathbb{C}^{n\times n}$ is called a *feasible solution* of (1) if it is positive semidefinite and fulfills all $m$ linear constraints. It is called an *optimal solution* if it is feasible and if for every feasible solutions $Y$ we have $\langle X, C\rangle \geq \langle Y, C\rangle$.

There is an easy reduction from complex semidefinite programs to semidefinite programs involving real matrices only, as noticed by Goemans and Williamson [43]. A complex matrix $X \in \mathbb{C}^{n\times n}$ defines a real matrix

$$\begin{pmatrix} \Re(X) & -\Im(X) \\ \Im(X) & \Re(X) \end{pmatrix} \in \mathbb{R}^{2n\times 2n},$$

where $\Re(X) \in \mathbb{R}^{n\times n}$ and $\Im(X) \in \mathbb{R}^{n\times n}$ are the real and imaginary parts of $X$. Then the properties of being Hermitian and being complex positive semidefinite translate into being symmetric and being real positive semidefinite: We have for all real vectors $\alpha = (\alpha_1, \ldots, \alpha_{2n}) \in \mathbb{R}^{2n}$:

$$\alpha^{\mathsf{T}} \begin{pmatrix} \Re(X) & -\Im(X) \\ \Im(X) & \Re(X) \end{pmatrix} \alpha \geq 0.$$

On the other hand, complex semidefinite programs fit into the framework of conic programming (see e.g. Nemirovski [76]). Here one uses the cone of positive semidefinite Hermitian matrices instead of the cone of real positive semidefinite matrices. There are implementations available, SeDuMi (Sturm [89]) for instance, which can deal with complex semidefinite programs directly.

### 1.2 Semidefinite programs invariant under a group action

Now we present the basic framework for simplifying a complex semidefinite program which has symmetry, i.e. which is invariant under the action of a group.

Let us fix some notation first. Let $G$ be a finite group. Let $\pi : G \to U_n(\mathbb{C})$ be a *unitary representation* of $G$, that is, a group homomorphism from the group $G$ to the group of unitary matrices $U_n(\mathbb{C})$. The group $G$ is acting on the set of Hermitian matrices by

$$(g, A) \mapsto \pi(g)A\pi(g)^*.$$

In general, a (left) *action* of a group $G$ on a set $M$ is a map

$$G \times M \to M, \quad (g, x) \mapsto gx,$$

that satisfies the following properties: We have $1x = x$ for all $x \in M$ where 1 denotes the neutral element of $G$. Furthermore, $(g_1g_2)(x) = g_1(g_2x)$ for all $g_1, g_2 \in G$ and all $x \in M$.

A matrix $X$ is called *$G$-invariant* if $X = gX$ for all $g \in G$, and we denote the set of all $G$-invariant matrices by $(\mathbb{C}^{n\times n})^G$. We say that the semidefinite program (1) is *$G$-invariant* if for every feasible solution $X$ and for every $g \in G$ the matrix $gX$ is again a feasible solution and if it satisfies $\langle gX, C\rangle = \langle X, C\rangle$ for all $g \in G$.

One example, which will receive special attention because of its importance, is the case of a permutation action: The set of feasible solutions is invariant under simultaneous permutations of rows and columns. Let $G$ be a finite group which acts on the index set $[n] = \{1, \ldots, n\}$. So we can see $G$ as a subgroup of the permutation group on $[n]$. To a permutation $\sigma$ on $[n]$ corresponds a matrix permutation $P_\sigma \in U_n(\mathbb{C})$ defined by

$$[P_\sigma]_{ij} = \begin{cases} 1, & \text{if } i = \sigma(j), \\ 0, & \text{otherwise.} \end{cases}$$

and the underlying unitary representation is $\pi(\sigma) = P_\sigma$. In this case, the action on matrices $X \in \mathbb{C}^{n\times n}$ is

$$\sigma(X) = P_\sigma X P_\sigma^*, \text{ where } \sigma(X)_{ij} = X_{\sigma^{-1}(ij)} = X_{\sigma^{-1}(i),\sigma^{-1}(j)}.$$

In the following, we give some background on unitary representations. This part may be skipped at first reading. Let $G$ be a finite group. A *representation of* $G$ is a finite dimensional complex vector space $V$ together with a homomorphism $\pi : G \to \mathrm{Gl}(V)$ from $G$ to the group of invertible linear maps on $V$. The space $V$ is also called a *$G$-space* or a *$G$-module* and $\pi$ may be dropped in notations, replacing $\pi(g)v$ by $gv$. In other words, the group $G$ acts on $V$ and this action has the additional property that $g(\lambda v + \mu w) = \lambda gv + \mu gw$ for all $(\lambda, \mu) \in \mathbb{C}^2$ and $(v, w) \in V^2$. The *dimension* of a representation is the dimension of the underlying vector space $V$.
If $V$ is endowed with an inner product $\langle v, w\rangle$ which is invariant under $G$, i.e. satisfies $\langle gv, gw\rangle = \langle v, w\rangle$ for all $(v, w) \in V^2$ and $g \in G$, $V$ is called a *unitary representation* of $G$. An inner product with this property always exists on $V$, since it is obtained by taking the average

$$\langle v, w\rangle = \frac{1}{|G|} \sum_{g\in G} \langle gv, gw\rangle_0$$

of an arbitrary inner product $\langle v, w\rangle_0$ on $V$. So, with an appropriate choice of a basis of $V$, any representation of $V$ is isomorphic to one of the form $\pi : G \to U_n(\mathbb{C})$, the form in which unitary representations of $G$ where defined above.

**Step 1: Restriction to invariant subspace**

Because of the convexity of (1), one can find an optimal solution of (1) in the set of $G$-invariant matrices. In fact, if $X$ is an optimal solution of (1), so is its *group average* $\frac{1}{|G|} \sum_{g\in G} gX$. Hence, (1) is equivalent to

$$\max\{\langle X, C\rangle : X \succeq 0, \langle X, A_1\rangle = b_1, \ldots, \langle X, A_m\rangle = b_m, X \in (\mathbb{C}^{n\times n})^G\}. \quad (2)$$

The $G$-invariant matrices intersected with the Hermitian matrices form a vector space. Let $B_1, \ldots, B_N$ be a basis of this space. Step 1 of simplifying a $G$-invariant semidefinite program is rewriting (2) in terms of this basis.

***Step 1:*** *If the semidefinite program* (1) *is $G$-invariant, then it is equivalent to*

$$\max \Big\{ \langle X, C \rangle : x_1, \ldots, x_N \in \mathbb{C}, \\ X = x_1 B_1 + \cdots + x_N B_N \succeq 0, \\ \langle X, A_i \rangle = b_i,\ i = 1, \ldots, m \Big\}. \tag{3}$$

In the case of a permutation action there is a canonical basis of $(\mathbb{C}^{n\times n})^G$ which one can determine by looking at the orbits of the group action on pairs. Then, performing Step 1 using this basis amounts to coupling the variable matrix entries of $X$.

The *orbit* of the pair $(i, j) \in [n] \times [n]$ under the group $G$ is given by

$$O(i, j) = \{(\sigma(i), \sigma(j)) : \sigma \in G\}.$$

The set $[n]\times[n]$ decomposes into the orbits $R_1, \ldots, R_M$ under the action of $G$. For every $r \in \{1, \ldots, M\}$ we define the matrix $C_r \in \{0, 1\}^{n\times n}$ by $(C_r)_{ij} = 1$ if $(i, j) \in R_r$ and $(C_r)_{ij} = 0$ otherwise. Then $C_1, \ldots, C_M$ forms a basis of $(\mathbb{C}^{n\times n})^G$, the *canonical basis*. If $(i, j) \in R_r$ we also write $C_{[i,j]}$ instead of $C_r$. Then, $C_{[j,i]}^{\mathsf{T}} = C_{[i,j]}$.

Note here that $C_1, \ldots, C_M$ is a basis of $(\mathbb{C}^{n\times n})^G$. In order to get a basis of the space of $G$-invariant Hermitian matrices we have to consider the orbits of unordered pairs: We get the basis by setting $B_{\{i,j\}} = C_{[i,j]}$ if $(i, j)$ and $(j, i)$ are in the same orbit and $B_{\{i,j\}} = C_{[i,j]} + C_{[j,i]}$ if they are in different orbits. The matrix entries of $X$ in (3) are constant on the (unordered) orbits of pairs: $X_{ij} = X_{\sigma(ij)} = X_{\sigma(ji)} = X_{ji}$ for all $(i, j) \in [n] \times [n]$ and $\sigma \in G$.

**Step 2: Reducing the matrix sizes by block diagonalization**

The $G$-invariant subspace $(\mathbb{C}^{n\times n})^G$ is closed under matrix multiplication. This can be seen as follows: Let $X, Y \in (\mathbb{C}^{n\times n})^G$ and let $g \in G$, then

$$g(XY) = \pi(g)XY\pi(g)^* = (\pi(g)X\pi(g)^*)(\pi(g)Y\pi(g)^*) = (gX)(gY) = XY.$$

Moreover, it is also closed under taking the conjugate transpose, since, for $X \in (\mathbb{C}^{n\times n})^G$, $\pi(g)X^*\pi(g)^* = (\pi(g)X\pi(g)^*)^* = X^*$. Thus $(\mathbb{C}^{n\times n})^G$ has the structure of a *matrix $*$-algebra*. (However, not all matrix $*$-algebras are coming from group actions.)

In general, a *matrix $*$-algebra* is a set of complex matrices that is closed under addition, scalar multiplication, matrix multiplication, and taking the conjugate transpose. The main structure theorem of matrix $*$-algebras is the following:

**Theorem 1.2** *Let $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ be a matrix $*$-algebra. There are numbers $d$, and $m_1, \ldots, m_d$ so that there is $*$-isomorphism between $\mathcal{A}$ and a direct sum of complete matrix algebras*

$$\varphi : \mathcal{A} \to \bigoplus_{k=1}^{d} \mathbb{C}^{m_k \times m_k}.$$

We will give a proof of this theorem in Section 2.5 which also will give an algorithm for determining $\varphi$. In the case $\mathcal{A} = (\mathbb{C}^{n\times n})^G$, the numbers $d$, and $m_1, \ldots, m_d$ have a representation theoretic interpretation: The numbers are determined by the unitary representation $\pi : G \to U_n(\mathbb{C})$, where $d$ is the number of pairwise non-isomorphic irreducible representations contained in $\pi$ and where $m_k$ is the multiplicity of the $k$-th isomorphism class of irreducible representations contained in $\pi$.

In the following we present background in representation theory which is needed for the understanding of the numbers $d$ and $m_1, \ldots, m_d$ in Theorem 1.2. Again, this part may be skipped at first reading.

A *$G$-homomorphism* $T$ between two $G$-spaces $V$ and $W$ is a linear map that commutes with the actions of $G$ on $V$ and $W$: for all $g \in G$ and all $v \in V$ we have $T(gv) = gT(v)$. If $T$ is invertible, then it is a *$G$-isomorphism* and $V$ and $W$ are *$G$-isomorphic*. The set of all $G$-homomorphisms is a linear space, denoted $\mathrm{Hom}^G(V, W)$. If $V = W$, then $T$ is said to be a *$G$-endomorphism* and the set $\mathrm{End}^G(V)$ of $G$-endomorphisms of $V$ is moreover an algebra under composition. If $V = \mathbb{C}^n$, and if $G$ acts on $\mathbb{C}^n$ by unitary matrices, then we have $\mathrm{End}^G(V) = (\mathbb{C}^{n\times n})^G = \mathcal{A}$.

A representation $\pi : G \to \mathrm{Gl}(V)$ of $G$ (and the corresponding $G$-space $V$) is *irreducible* if it contains no proper subspace $W$ such that $gW \subset W$ for all $g \in G$, i.e. $V$ contains no *$G$-subspace*. If $V$ contains a proper $G$-subspace $W$, then also the orthogonal complement $W^\perp$ relative to a $G$-invariant inner product, is a $G$-subspace and $V$ is the direct sum $W \oplus W^\perp$. Inductively, one obtains Maschke's theorem:

*Every $G$-space is the direct sum of irreducible $G$-subspaces.*

This decomposition is generally not unique: For example, if $G$ acts trivially on a vector space $V$ (i.e. $gv = v$ for all $g \in G$, $v \in V$) of dimension at least 2, the irreducible subspaces are the 1-dimensional subspaces and $V$ can be decomposed in many ways.

From now on, we fix a set $\mathcal{R} = \{R_k : k = 1, \ldots, d\}$ of representatives of the isomorphism classes of irreducible $G$-subspaces which are direct summands of $V$. Starting from an arbitrary decomposition of $V$, we consider, for $k = 1, \ldots, d$, the sum of the irreducible subspaces which are isomorphic to $R_k$. One can prove that this $G$-subspace of $V$, is independent of the decomposition. It is called the *isotypic component* of $V$ associated to $R_k$ and is denoted $\mathcal{MI}_k$. The integer $m_k$ such that $\mathcal{MI}_k \simeq R_k^{m_k}$ is called the *multiplicity* of $R_k$ in $V$. In other words, we have

$$V = \bigoplus_{k=1}^{d} \mathcal{MI}_k \quad \text{and} \quad \mathcal{MI}_k = \bigoplus_{i=1}^{m_k} H_{k,i}$$

where, $H_{k,i}$ is isomorphic to $R_k$, and $m_k \geq 1$. The first decomposition is orthogonal with respect to an invariant inner product and is uniquely determined by $V$, while the decomposition of $\mathcal{MI}_k$ is not unique unless $m_k = 1$.

*Schur's lemma* is the next crucial ingredient for the description of $\mathrm{End}^G(V)$:

*If $V$ is $G$-irreducible, then* $\mathrm{End}^G(V) = \{\lambda \,\mathrm{Id} : \lambda \in \mathbb{C}\} \simeq \mathbb{C}$.

In the general case, when $V$ is not necessarily $G$-irreducible,

$$\mathrm{End}^G(V) \simeq \bigoplus_{k=1}^{d} \mathbb{C}^{m_k \times m_k}.$$

This result will be derived in the next section, in an algorithmic way, as a consequence of the more general theory of matrix $*$-algebras.

So we consider a $*$-isomorphism $\varphi$ given by Theorem 1.2 applied to $\mathcal{A} = (\mathbb{C}^{n\times n})^G$:

$$\varphi : (\mathbb{C}^{n\times n})^G \to \bigoplus_{k=1}^{d} \mathbb{C}^{m_k \times m_k}. \tag{4}$$

Notice that since $\varphi$ is a $*$-isomorphism between matrix algebras with unity, $\varphi$ preserves also eigenvalues and hence positive semidefiniteness. Indeed, let $X \in (\mathbb{C}^{n\times n})^G$ be $G$-invariant, then also $X - \lambda I$ is $G$-invariant, and $X - \lambda I$ has an inverse if and only if $\varphi(X) - \lambda I$ has a inverse. This means that a test whether a (large) $G$-invariant matrix is positive semidefinite can be reduced to a test whether $d$ (small) matrices are positive semidefinite. Hence, applying $\varphi$ to (3) gives the second and final step of simplifying (1):

***Step 2 (first version):*** *If the semidefinite program* (1) *is $G$-invariant, then it is equivalent to*

$$\max\Big\{\langle X, C\rangle : x_1, \ldots, x_N \in \mathbb{C}, \\ X = x_1 B_1 + \cdots + x_N B_N \succeq 0, \\ \langle X, A_i\rangle = b_i,\ i = 1, \ldots, m, \\ x_1\varphi(B_1) + \cdots + x_N\varphi(B_N) \succeq 0\Big\}. \tag{5}$$

Applying $\varphi$ to a $G$-invariant semidefinite program is also called *block diagonalization.* The advantage of (5) is that instead of dealing with matrices of size $n \times n$ one only has to deal with block diagonal matrices with $d$ block matrices of size $m_1, \ldots, m_d$, respectively. So one reduces the dimension from $n^2$ to $m_1^2 + \cdots + m_d^2$. In the case of a permutation action this sum is also the number of distinct orbits $M$. In many applications the latter is much smaller than the former. In particular many practical solvers take advantage of the block structure to speed up the numerical calculations.

Instead of working with the $*$-isomorphism $\varphi$ and a basis $B_1, \dots, B_N$ of the Hermitian $G$-invariant matrices, one can also work with the inverse $\varphi^{-1}$ and the standard basis of $\bigoplus_{k=1}^d \mathbb{C}^{m_k \times m_k}$. This is given by the matrices $E_{k,uv} \in \mathbb{C}^{m_k \times m_k}$ where all entries of $E_{k,uv}$ are zero except the $(u,v)$-entry which equals 1. This gives an *explicit parametrization* of the cone of $G$-invariant positive semidefinite matrices.

***Step 2 (second version):*** *If the semidefinite program* (1) *is $G$-invariant, then it is equivalent to*

$$\begin{aligned}\max \Big\{ \langle X, C\rangle : \\ & X = \sum_{k=1}^d \sum_{u,v=1}^{m_k} x_{k,uv} \varphi^{-1}(E_{k,uv}) \\ & x_{k,uv} = \overline{x_{k,vu}},\ u,v = 1,\dots,m_k, \\ & \big(x_{k,uv}\big)_{1 \le u,v \le m_k} \succeq 0,\ k = 1,\dots,d, \\ & \langle X, A_i\rangle = b_i,\ i = 1,\dots,m \Big\}.\end{aligned} \tag{6}$$

Hence, every $G$-invariant positive semidefinite matrix $X$ is of the form

$$X = \sum_{k=1}^d \sum_{u,v=1}^{m_k} x_{k,uv} \varphi^{-1}(E_{k,uv}),$$

where the $d$ matrices

$$X_k = \big(x_{k,uv}\big)_{1 \le u,v \le m_k},\quad k = 1,\dots,d,$$

are positive semidefinite. Define for $(i,j) \in [n] \times [n]$ the matrix $E_k(i,j) \in \mathbb{C}^{m_k \times m_k}$ componentwise by

$$[E_k(i,j)]_{uv} = \left[\varphi^{-1}(E_{k,uv})\right]_{ij}.$$

By definition we have $E_k(i,j)^* = E_k(j,i)$. Then, in the case of a permutation action, $E_k(i,j) = E_k(\sigma(i), \sigma(j))$ for all $(i,j) \in [n] \times [n]$ and $\sigma \in G$. With this notation one can write the $(i,j)$-entry of $X$ by

$$X_{ij} = \sum_{k=1}^d \langle X_k, E_k(i,j)\rangle. \tag{7}$$

In summary, finding a block diagonalization of a $G$-invariant semidefinite program amounts to first identifying a basis of the Hermitian $G$-invariant matrices and then in finding an explicit $*$-isomorphism (4) between the algebra of $G$-invariant matrices and the direct sum of complete matrix algebras. In the following sections we will mainly be concerned with different strategies to find such a $*$-isomorphism.

**Step $1\frac{1}{2}$: Reducing the matrix sizes by the regular $*$-representation**

In general finding a block diagonalization of a $G$-invariant semidefinite program is a non-trivial task, especially because one has to construct an explicit $*$-isomorphism. In cases where one does not have this one can fall back to a simpler $*$-isomorphism coming from the regular $*$-representation. In general this does not provide the maximum possible simplification. However, for instance in the case of a permutation action, it has the advantage that one can compute it on the level of knowing the orbit structure of the group action only.

For this we consider an orthogonal basis (with respect to the trace inner product $\langle \cdot, \cdot \rangle$) of the $G$-invariant algebra $(\mathbb{C}^{n\times n})^G$. For instance, we can use the canonical basis $C_1, \ldots, C_M$ in the case of a permutation action. By considering the multiplication table of the algebra we define the *multiplication parameters* $p_{rs}^t$, sometimes also called *structural parameters*, by

$$C_r C_s = \sum_{t=1}^{M} p_{rs}^t C_t.$$

In the case of a permutation action the structural parameters can be computed by knowing the structure of orbits (if one chose the canonical basis):

$$p_{rs}^t = |\{k \in [n] : (i,k) \in R_r, (k,j) \in R_s\}|,$$

where $(i,j) \in R_t$. Here, $p_{rs}^t$ does not depend on the choice of $i$ and $j$. The norms $||C_r|| = \sqrt{\langle C_r, C_r \rangle}$ equal the sizes of the corresponding orbits. We define the matrices $L(C_r)_{st} \in \mathbb{C}^{M\times M}$ by

$$(L(C_r))_{st} = \frac{\langle C_r C_t, C_s \rangle}{\|C_t\| \, \|C_s\|} = \frac{\|C_s\|}{\|C_t\|} p_{rt}^s.$$

**Theorem 1.3** *Let $\mathcal{L}$ the algebra generated by the matrices $L(C_1), \ldots, L(C_M)$. Then the linear map*

$$\phi : (\mathbb{C}^{n\times n})^G \to \mathcal{L}, \quad \phi(C_r) = L(C_r), \quad r = 1, \ldots, M,$$

*is a $*$-isomorphism.*

We will give a proof of this theorem in Section 2.6. There we will show that the $*$-isomorphism is the regular $*$-representation of the $G$-invariant algebra associated with the orthonormal basis $C_1/\|C_1\|, \ldots, C_M/\|C_M\|$. Again, since $\phi$ is a $*$-isomorphism between algebras with unity it preserves eigenvalues. This means that a test whether a $G$-invariant matrix of size $n \times n$ is positive semidefinite can be reduced to testing whether an $M \times M$ matrix is positive semidefinite.

***Step*** $\mathbf{1\frac{1}{2}}$ : *If the semidefinite program* (1) *is $G$-invariant, then it is equivalent to* (5) *where the $*$-isomorphism $\varphi$ is replaced by $\phi$.*

# 2 Matrix $*$-algebras

In Section 1.2 we saw that the process of block diagonalizing a semidefinite program can be naturally done in the framework of matrix $*$-algebras using the main structure theorem (Theorem 1.2). In this section we prove this main structure theorem. Although we are mainly interested in the case when the matrix $*$-algebra comes from a group, working in the more general framework here, does not cause much extra work. Furthermore the proof of the main structure theorem we give here provides an algorithmic way for finding a block diagonalization.

We start by giving the basic definitions, examples, and results of matrix $*$-algebras (Section 2.1–Section 2.4). In Section 2.5 we prove the main structure theorem which gives a very efficient representation of a matrix $*$-algebra $\mathcal{A}$: We show that $\mathcal{A}$ is $*$-isomorphic to a direct sum of full matrix $*$-algebras. The corresponding $*$-isomorphism is called a *block diagonalization* of $\mathcal{A}$. This corresponds to Step 2 in the introduction. After giving the proof we interpret it in the context of groups and we discuss a numerical algorithm for finding a block diagonalization which is based on the proof. In Section 2.6 we consider the regular $*$-representation, which embeds $\mathcal{A}$ into $\mathbb{C}^{M\times M}$, where $M = \dim \mathcal{A}$. This corresponds to Step $1\frac{1}{2}$ in the introduction.

## 2.1 Definitions and examples

**Definition 2.1** *A* matrix $*$-algebra *is a linear subspace $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ of complex $n \times n$ matrices, that is closed under matrix multiplication and under taking the conjugate transpose. The conjugate transpose of a matrix $A$ is denoted $A^*$.*

Matrix $*$-algebras are finite dimensional $C^*$-algebras and many results here can be extended to a more general setting. For a gentle introduction to $C^*$-algebras we refer to Takesaki [92].

Trivial examples of matrix $*$-algebras are the *full matrix algebra* $\mathbb{C}^{n\times n}$ and the *zero algebra* $\{0\}$. From given matrix $*$-algebras $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ and $\mathcal{B} \subseteq \mathbb{C}^{m\times m}$, we can construct the *direct sum* $\mathcal{A} \oplus \mathcal{B}$ and *tensor product* $\mathcal{A} \otimes \mathcal{B}$ defined by

$$\mathcal{A} \oplus \mathcal{B} = \left\{ \left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right) : A \in \mathcal{A}, B \in \mathcal{B} \right\},$$

$$\mathcal{A} \otimes \mathcal{B} = \left\{ \sum_{i=1}^{k} A_i \otimes B_i : k \in \mathbb{N}, A_i \in \mathcal{A}, B_i \in \mathcal{B} \right\},$$

where $A \otimes B \in \mathbb{C}^{nm \times nm}$ denotes the Kronecker- or tensor product. The *commutant* of $\mathcal{A}$ is the matrix $*$-algebra

$$\mathcal{A}' = \{B \in \mathbb{C}^{n \times n} : BA = AB \text{ for all } A \in \mathcal{A}\}.$$

Many interesting examples of matrix $*$-algebras come from unitary group representations, as we already demonstrated in the introduction: Given a unitary representation $\pi : G \to \mathrm{Gl}(n, \mathbb{C})$, the set of invariant matrices $(\mathbb{C}^{n \times n})^G = \{A \in \mathbb{C}^{n \times n} : \pi(g)A\pi(g)^{-1} = A\}$ is a matrix $*$-algebra. It is the commutant of the matrix $*$-algebra linearly spanned by the matrices $\pi(g)$ with $g \in G$. If the unitary group representation is given by permutation matrices then the canonical basis of the algebra $(\mathbb{C}^{n \times n})^G$ are the zero-one incidence matrices of orbits on pairs $C_1, \ldots, C_M$, see Step 1 in Section 1.2.

Other examples of matrix $*$-algebras, potentially not coming from groups, include the (complex) Bose-Mesner algebra of an *association scheme*, see e.g. Bannai, Ito [13], and Brouwer, Cohen, Neumaier [20] and more generally, the adjacency algebra of a *coherent configuration*, see e.g. Cameron [22].

## 2.2 Commutative matrix $*$-algebras

A matrix $*$-algebra $\mathcal{A}$ is called *commutative* (or *Abelian*) if any pair of its elements commute: $AB = BA$ for all $A, B \in \mathcal{A}$. Recall that a matrix $A$ is *normal* if $AA^* = A^*A$. The spectral theorem for normal matrices states that if $A$ is normal, there exists a unitary matrix $U$ such that $U^*AU$ is a diagonal matrix. More generally, a set of commuting normal matrices can be simultaneously diagonalized (see e.g. Horn, Johnson [48]). Since any algebra of diagonal matrices has a basis of zero-one diagonal matrices with disjoint support, we have the following theorem.

**Theorem 2.2** *Let $\mathcal{A} \subseteq \mathbb{C}^{n \times n}$ be a commutative matrix $*$-algebra. Then there exist a unitary matrix $U$ and a partition $[n] = S_0 \cup S_1 \cup \cdots \cup S_k$ with $S_1, \ldots, S_k$ nonempty, such that*

$$U^*\mathcal{A}U = \{\lambda_1 I_1 + \cdots + \lambda_k I_k : \lambda_1, \ldots, \lambda_k \in \mathbb{C}\},$$

*where $I_i$ is the zero-one diagonal matrix with ones in positions from $S_i$ and zeroes elsewhere.*

The matrices $E_i = UI_iU^*$ satisfy $E_0 + E_1 + \cdots + E_k = I$, $E_iE_j = \delta_{ij}E_i$ and $E_i = E_i^*$. The matrices $E_1, \ldots, E_k$ are the *minimal idempotents* of $\mathcal{A}$ and form an orthogonal basis of $\mathcal{A}$. Unless $S_0 = \emptyset$, $E_0$ does not belong to $\mathcal{A}$.

Geometrically, we have an orthogonal decomposition

$$\mathbb{C}^n = V_0 \oplus V_1 \oplus \cdots \oplus V_k,$$

where $E_i$ is the orthogonal projection onto $V_i$ or equivalently, $V_i$ is the space spanned by the columns of $U$ corresponding to $S_i$. The space $V_0$ is the maximal subspace contained in the kernel of all matrices in $\mathcal{A}$.

### 2.3 Positive semidefinite elements

Recall that a matrix $A$ is positive semidefinite ($A \succeq 0$) if and only if $A$ is *Hermitian* (that is $A^* = A$) and all eigenvalues of $A$ are nonnegative. Equivalently, $A = U^*DU$ for some unitary matrix $U$ and nonnegative diagonal matrix $D$.

Let $\mathcal{A}$ be a matrix $*$-algebra. Positive semidefiniteness can be characterized in terms of $\mathcal{A}$.

**Proposition 2.3** *An element $A \in \mathcal{A}$ is positive semidefinite if and only if $A = B^*B$ for some $B \in \mathcal{A}$.*

*Proof.* The 'if' part is trivial. To see the 'only if' part, let $A \in \mathcal{A}$ be positive semidefinite and write $A = U^*DU$ for some unitary matrix $U$ and (nonnegative real) diagonal matrix $D$. Let $p$ be a polynomial with $p(\lambda_i) = \sqrt{\lambda_i}$ for all eigenvalues $\lambda_i$ of $A$. Then taking $B = p(A) \in \mathcal{A}$ we have $B = U^*p(D)U$ and hence $B^*B = U^*p(D)p(D)U = U^*DU = A$. $\square$

Considered as a cone in the space of Hermitian matrices in $\mathcal{A}$, the cone of positive semidefinite matrices is self-dual:

**Theorem 2.4** *Let $A \in \mathcal{A}$ be Hermitian. Then $A \succeq 0$ if and only if $\langle A, B\rangle \geq 0$ for all $B \succeq 0$ in $\mathcal{A}$.*

*Proof.* Necessity is clear. For sufficiency, let $A \in \mathcal{A}$ be Hermitian and let $\mathcal{B} \subseteq \mathcal{A}$ be the $*$-subalgebra generated by $A$. By Theorem 2.2 we can write $A = \lambda_1 E_1 + \cdots + \lambda_k E_k$, where the $E_i$ are the minimal idempotents of $\mathcal{B}$. The $E_i$ are positive semidefinite since their eigenvalues are zero or one. Hence by assumption, $\lambda_i = \frac{\langle A, E_i\rangle}{\langle E_i, E_i\rangle} \geq 0$. Therefore $A$ is a nonnegative linear combination of positive semidefinite matrices and hence positive semidefinite. $\square$

As a corollary we have:

**Corollary 2.5** *The orthogonal projection $\pi_{\mathcal{A}} : \mathbb{C}^{n\times n} \to \mathcal{A}$ preserves positive semidefiniteness.*

This implies that if the input matrices of the semidefinite program (1) lie in some matrix $*$-algebra, then we can assume that the optimization variable $X$ lies in the same matrix $*$-algebra: If (1) is given by matrices $C, A_1, \ldots, A_m \in \mathcal{A}$ for some matrix $*$-algebra $\mathcal{A}$, the variable $X$ may be restricted to $\mathcal{A}$ without changing the objective value. Indeed, any feasible $X$ can be replaced by $\pi_{\mathcal{A}}(X)$, which is again feasible and has the same objective value. When $\mathcal{A}$ is the invariant algebra of a group, this amounts to replacing $X$ by the average under the action of the group.

### 2.4 $*$-Homomorphisms and block diagonalization

**Definition 2.6** *A map $\phi : \mathcal{A} \to \mathcal{B}$ between two matrix $*$-algebras $\mathcal{A}$ and $\mathcal{B}$ is called a* $*$-homomorphism *if*

*(i)* *$\phi$ is linear,*
*(ii)* *$\phi(AB) = \phi(A)\phi(B)$ for all $A, B \in \mathcal{A}$,*
*(iii)* *$\phi(A^*) = \phi(A)^*$ for all $A \in \mathcal{A}$.*

*If $\phi$ is a bijection, the inverse map is also a $*$-homomorphism and $\phi$ is called a* $*$-isomorphism*.*

It follows directly from Proposition 2.3 that $*$-homomorphisms preserve positive semidefiniteness: Let $\phi : \mathcal{A} \to \mathcal{B}$ be a $*$-homomorphism. Then $\phi(A)$ is positive semidefinite for every positive semidefinite $A \in \mathcal{A}$.

The implication of this for semidefinite programming is the following. Given a semidefinite program with matrix variable restricted to a matrix $*$-algebra $\mathcal{A}$ and a $*$-isomorphism $\mathcal{A} \to \mathcal{B}$, we can rewrite the semidefinite program in terms of matrices in $\mathcal{B}$. This can be very useful if the matrices in $\mathcal{B}$ have small size compared to those in $\mathcal{A}$. In the following, we will discuss two such efficient representations of a general matrix $*$-algebra.

### 2.5 Block diagonalization

In this section, we study the structure of matrix $*$-algebras in some more detail. The main result is, that any matrix $*$-algebra is $*$-isomorphic to a direct sum of full matrix algebras:

**Theorem 2.7** ***(= Theorem 1.2)*** *Let $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ be a matrix $*$-algebra. There are numbers $d$, and $m_1, \ldots, m_d$ so that there is $*$-isomorphism between $\mathcal{A}$ and a direct sum of complete matrix algebras*

$$\varphi : \mathcal{A} \to \bigoplus_{k=1}^{d} \mathbb{C}^{m_k \times m_k}.$$

This theorem is well-known in the theory of $C^*$-algebras, where it is generalized to $C^*$-algebras of compact operators on a Hilbert space (cf. Davidson [28, Chapter I.10]).

Some terminology: Let $\mathcal{A}$ be a matrix $*$-algebra. A matrix $*$-algebra $\mathcal{B}$ contained in $\mathcal{A}$ is called a $*$-*subalgebra* of $\mathcal{A}$. An important example is the *center* of $\mathcal{A}$ defined by

$$\mathcal{C}_{\mathcal{A}} = \{A \in \mathcal{A} : AB = BA \text{ for all } B \in \mathcal{A}\}.$$

There is a unique element $E \in \mathcal{A}$ such that $EA = AE = A$ for every $A \in \mathcal{A}$, which is called the *unit* of $\mathcal{A}$. If $\mathcal{A}$ is non-zero and $\mathcal{C}_{\mathcal{A}} = \mathbb{C}E$ (equivalently: $\mathcal{A}$ has no nontrivial *ideal*), the matrix $*$-algebra $\mathcal{A}$ is called *simple*.

We shall prove that every matrix $*$-algebra is the direct sum of simple matrix $*$-algebras:

$$\mathcal{A} = \bigoplus_{i=1}^{d} E_i \mathcal{A},$$

where the $E_i$ are the minimal idempotents of $\mathcal{C}_{\mathcal{A}}$. Every simple matrix $*$-algebra is $*$-isomorphic to a full matrix algebra. Together these facts imply Theorem 2.7.

To conclude this section, we will give an elementary and detailed proof of Theorem 2.7.

*Proof.* Let $\mathcal{B} \subseteq \mathcal{A}$ be an inclusionwise maximal commutative $*$-subalgebra of $\mathcal{A}$. Then any $A \in \mathcal{A}$ that commutes with every element of $\mathcal{B}$, is itself an element of $\mathcal{B}$. Indeed, if $A$ is normal, this follows from the maximality of $\mathcal{B}$ since $\mathcal{B} \cup \{A, A^*\}$ generates a commutative $*$-algebra containing $\mathcal{B}$. If $A$ is not normal, then $A + A^* \in \mathcal{B}$ by the previous argument and hence $A(A + A^*) = (A + A^*)A$, contradicting the fact that $A$ is not normal.

By replacing $\mathcal{A}$ by $U^*\mathcal{A}U$ for a suitable unitary matrix $U$, we may assume that $\mathcal{B}$ is in diagonal form as in Theorem 2.2. For any $A \in \mathcal{A}$ and $i, j = 0, \ldots, k$, denote by $A_{ij} \in \mathbb{C}^{|S_i| \times |S_j|}$ the restriction of $I_i A I_j$ to the rows in $S_i$ and columns in $S_j$. Let $A \in \mathcal{A}$ and $i, j \in \{0, \ldots, k\}$. We make the following observations:

(i) $A_{ii}$ is a multiple of the identity matrix and $A_{00} = 0$,
(ii) $A_{0i}$ and $A_{i0}$ are zero matrices,
(iii) $A_{ij}$ is either zero or a nonzero multiple of a (square) unitary matrix.

Item (i) follows since $I_i A I_i$ commutes with $I_0, \ldots, I_k$ and therefore belongs to $\mathcal{B}$. Hence $I_i A I_i$ is a multiple of $I_i$ and $I_0 A I_0 = 0$ since $I_0 \mathcal{B} I_0 = \{0\}$. Similarly, $I_0 A A^* I_0 = 0$, which implies that $I_0 A = 0$, showing (ii). For item (iii), suppose that $A_{ij}$ is nonzero and assume without loss of generality that $|S_i| \geq |S_j|$. Then by (i), $A_{ij} A_{ij}^* = \lambda I$ for some positive real $\lambda$, and therefore has rank $|S_i|$. This implies that $|S_j| = |S_i|$ and $\sqrt{\lambda} \cdot A_{ij}$ is unitary.

Observe that (ii) shows that $I_1 + \cdots + I_k$ is the unit of $\mathcal{A}$.

Define the relation $\sim$ on $\{1, \ldots, k\}$ by setting $i \sim j$ if and only if $I_i \mathcal{A} I_j \neq \{0\}$. This is an equivalence relation. Indeed, $\sim$ is reflexive by (i) and symmetric since $I_i \mathcal{A} I_j = (I_j \mathcal{A} I_i)^*$. Transitivity follows from (iii) since $I_h \mathcal{A} I_j \supseteq (I_h \mathcal{A} I_i)(I_i \mathcal{A} I_j)$ and the product of two unitary matrices is unitary.

Denote by $\{E_1, \ldots, E_d\} = \{\sum_{j \sim i} I_j : i = 1, \ldots, k\}$ the zero-one diagonal matrices induced by the equivalence relation. Since the center $\mathcal{C}_{\mathcal{A}}$ of $\mathcal{A}$ is contained in $\mathcal{B} = \mathbb{C}I_1 + \cdots + \mathbb{C}I_k$, it follows by construction that $E_1, \ldots, E_d$ span the center, and are its minimal idempotents. We find the following block structure of $\mathcal{A}$:

$$\mathcal{A} = \{0\} \oplus E_1 \mathcal{A} \oplus \cdots \oplus E_d \mathcal{A},$$

where the matrix $*$-algebras $E_i \mathcal{A}$ are simple. For the rest of the proof we may assume that $\mathcal{A}$ is simple ($d = 1$) and that $E_0 = 0$.

Since $\sim$ has only one equivalence class, for every matrix $A = (A_{ij})_{i,j=1}^k \in \mathcal{A}$, the 'blocks' $A_{ij}$ are square matrices of the same size. Furthermore, we can fix an $A \in \mathcal{A}$ for which all the $A_{ij}$ are unitary. For any $B \in \mathcal{A}$, we have $A_{1i}B_{ij}(A_{1j})^* = (AI_iBI_jA^*)_{11}$. By (i), it follows that $\{A_{1i}B_{ij}(A_{1j})^* : B \in \mathcal{B}\} = \mathbb{C}I$. Hence setting $U$ to be the unitary matrix $U := \mathrm{diag}(A_{11}, \ldots, A_{1k})$, we see that $U\mathcal{A}U^* = \{(a_{ij}I)_{i,j=1}^k : a_{i,j} \in \mathbb{C}\}$, which shows that $\mathcal{A}$ is $*$-isomorphic to $\mathbb{C}^{k\times k}$. □

### Relation to group representations

In the case that $\mathcal{A} = (\mathbb{C}^{n\times n})^G$, where $\pi : G \to U_n(\mathbb{C})$ is a unitary representation, the block diagonalization can be interpreted as follows. The diagonalization of the maximal commutative $*$-subalgebra $\mathcal{B}$, gives a decomposition $\mathbb{C}^n = V_1 \oplus \cdots \oplus V_k$ into irreducible submodules. Observe that $V_0 = \{0\}$ since $\mathcal{A}$ contains the identity matrix. The equivalence relation $\sim$ yields the isotypic components $\mathrm{Im}E_1, \ldots, \mathrm{Im}E_d$, where the sizes of the equivalence classes correspond to the block sizes $m_i$ in Theorem 2.7. Here Schur's lemma is reflected by the fact that $I_i\mathcal{A}I_j = \mathbb{C}I_i$ if $i \sim j$ and $\{0\}$ otherwise.

### Algorithmic aspects

If a matrix $*$-algebra $\mathcal{A}$ is given explicitly by a basis, then the above proof of Theorem 2.7 can be used to find a block diagonalization of $\mathcal{A}$ computationally. Indeed, it suffices to find an inclusion-wise maximal commutative $*$-subalgebra $\mathcal{B} \subseteq \mathcal{A}$ and compute a common system of eigenvectors for (basis) elements of $\mathcal{B}$. This can be done by standard linear algebra methods. For example finding a maximal commutative $*$-subalgebra of $\mathcal{A}$ can be done by starting with $\mathcal{B} = \langle A \rangle$, the $*$-subalgebra generated by an arbitrary Hermitian element in $\mathcal{A}$. As long as $\mathcal{B}$ is not maximal, there is a Hermitian element in $\mathcal{A} \setminus \mathcal{B}$ that commutes with all elements in (a basis of) $\mathcal{B}$, hence extending $\mathcal{B}$. Such an element can be found by solving a linear system in $O(\dim \mathcal{A})$ variables and $O(\dim \mathcal{B})$ constraints. In at most $\dim \mathcal{A}$ iterations, a maximal commutative $*$-subalgebra is found.

Practically more efficient is to find a "generic" Hermitian element $A \in \mathcal{A}$. Then the matrix $*$-algebra $\mathcal{B}$ generated by $A$ will be a maximal commutative $*$-subalgebra of $\mathcal{A}$ and diagonalizing the matrix $A$ also diagonalizes $\mathcal{B}$. Such a generic element can be found by taking a random Hermitian element from $\mathcal{A}$ (with respect to the basis), see Murota et. al. [74]. If a basis for the center of $\mathcal{A}$ is known a priori (or by solving a linear system in $O(\dim \mathcal{A})$ variables and equations), as an intermediate step the center could be diagonalized, followed by a block diagonalization of the simple components of $\mathcal{A}$, see Dobre, de Klerk, Pasechnik [34].

### 2.6 Regular $*$-representation

Let $\mathcal{A}$ be a matrix $*$-algebra of dimension $M$ and let $C_1, \ldots, C_M$ be an orthonormal basis of $\mathcal{A}$ with respect to the trace product $\langle \cdot, \cdot \rangle$. For fixed $A \in \mathcal{A}$, left-multiplication by $A$ defines a linear map $B \mapsto AB$ on $\mathcal{A}$. With respect to the orthonormal basis $C_1, \ldots, C_M$, this linear map is represented by the matrix $L(A) \in \mathbb{C}^{M \times M}$ given by

$$L(A)_{st} = \langle AC_t, C_s \rangle.$$

The map

$$L : \mathcal{A} \to \mathbb{C}^{M \times M}$$

is an injective $*$-homomorphism called the *regular $*$-representation of* $\mathcal{A}$. The fact that $L$ is linear and preserves matrix products is clear. Injectivity follows from the fact that $L(A) = 0$ implies $AA^* = 0$ and hence $A = 0$. Finally, the equations

$$L(A^*)_{st} = \langle A^* C_t, C_s \rangle = \langle C_t, AC_s \rangle = \overline{\langle AC_s, C_t \rangle} = \overline{L(A)_{ts}}$$

show that $L(A^*) = L(A)^*$. Because $L$ is linear, it is determined by the images $L(C_1), \ldots, L(C_M)$.

In many applications, for example in the case of a permutation action with the canonical basis $C_1, \ldots, C_M$ (Step $1\frac{1}{2}$ in the introduction), one only has an orthogonal basis which is not orthonormal. In that case, the map $L$ is given by

$$L(C_r)_{st} = \frac{\langle C_r C_t, C_s \rangle}{||C_s|| \, ||C_t||} = \frac{||C_s||}{||C_t||} p_{rt}^s,$$

where the $p_{rs}^t$ are the multiplication parameters defined by $C_r C_s = \sum_t p_{rs}^t C_t$. If we denote $\phi(C_r) = L(C_r)$, we obtain Theorem 1.3.

## 3 Invariant positive definite functions on compact spaces

Until now we considered only finite dimensional invariant semidefinite programs and the question: How can symmetry coming from the action of a finite group be exploited to simplify the semidefinite program? In several situations, some given in Section 6, one wants to work with infinite dimensional invariant semidefinite programs. In these situations using the symmetry coming from the action of an infinite, continuous group is a must if one wants to do explicit computations. In this section we introduce the cone of continuous, positive definite functions on a compact set $M$, as a natural generalization of the cone of positive semidefinite matrices. When a compact group $G$ acts continuously on $M$, we use the representation theory of $G$ to describe the subcone of $G$-invariant positive definite functions on $M$. This is the infinite dimensional analog of Step 2 (second version), given in the introduction, in the case of a permutation representation.

In general, this method evidences interesting links between the geometry of $M$, the representations of $G$ and the theory of orthogonal polynomials. We will review the spaces $M$, finite and infinite, where this method has been completely worked out.

### 3.1 Positive definite functions

Let $M$ be a compact space. We denote the space of continuous functions on $M$ taking complex values by $\mathcal{C}(M)$.

**Definition 3.1** *We say that $F \in \mathcal{C}(M^2)$ is* positive definite*, $F \succeq 0$, if $F(x,y) = \overline{F(y,x)}$ and, for all $n$, for all $n$-tuples $(x_1, \dots, x_n) \in M^n$ and vectors $(\alpha_1, \dots, \alpha_n) \in \mathbb{C}^n$,*

$$\sum_{i,j=1}^{n} \alpha_i F(x_i, x_j) \overline{\alpha_j} \geq 0.$$

*In other words, for all choices of finite subsets $\{x_1, \dots, x_n\}$ of $M$, the matrix*

$$\big(F(x_i, x_j)\big)_{1 \leq i,j \leq n}$$

*is Hermitian and positive semidefinite. The set of continuous positive definite functions on $M$ is denoted $\mathcal{C}(M^2)_{\succeq 0}$.*

In particular, if $M$ is a finite set with the discrete topology, the elements of $\mathcal{C}(M^2)_{\succeq 0}$ coincide with the Hermitian positive semidefinite matrices indexed by $M$. In general, $\mathcal{C}(M^2)_{\succeq 0}$ is a closed convex cone in $\mathcal{C}(M^2)$.

We now assume that a compact group $G$ *acts continuously on $M$*. Here we mean that, in addition to the properties of a group action (see introduction), the map $G \times M \to M$, with $(g,x) \mapsto gx$, is continuous. We also assume that $M$ is endowed with a *regular Borel measure* $\mu$ which is *$G$-invariant*.

> We recall here the basic notions on measures that will be needed. If $M$ is a topological space, a *Borel measure* is a measure defined on the $\sigma$-algebra generated by the open sets of $M$. A *regular Borel measure* is one which behaves well with respect to this topology, we refer to Rudin [83] for the precise definition.
> If $M$ is a finite set, the topology on $M$ is chosen to be the *discrete topology*, i.e. every subset of $M$ is an open set, and the measure $\mu$ on $M$ is the *counting measure* defined by $\mu(A) = |A|$.
> On a compact group $G$, there is a regular Borel measure $\lambda$ which is *left and right invariant*, meaning that $\lambda(gA) = \lambda(Ag) = \lambda(A)$ for all $g \in G$ and $A \subset G$ measurable, and such that the open sets have positive measure. This measure is unique up to a positive multiplicative constant and is called the *Haar measure on $G$* (see e.g. Conway [25, Theorem 11.4]). If $G$ is finite, the counting measure on $G$ provides a measure with these properties.

### 3.2 Unitary representations of compact groups

The definitions and properties of representations given in the introduction extend naturally to compact groups. We only have to modify the definition of a finite dimensional representation by asking that the homomorphism $\pi : G \to \mathrm{Gl}(V)$ is continuous.

In the proofs, integrating with the *Haar measure* $\lambda$ of $G$ replaces averaging on finite groups. For example, in the introduction, we have seen that every representation $V$ of a finite group $G$ is unitary. The argument was, starting from an arbitrary inner product $\langle u, v\rangle_0$ on $V$, to average it over all elements of $G$, in order to transform it into a $G$-invariant inner product. In the case of a more general compact group, we average over $G$ using the Haar measure (normalized so that $\lambda(G) = 1$), thus taking:

$$\langle u, v\rangle := \int_G \langle gu, gv\rangle d\lambda(g).$$

So Maschke's theorem holds: every finite dimensional $G$-space is the direct sum of irreducible $G$-subspaces. In contrast to the case of finite groups, the number of isomorphism classes of irreducible $G$-spaces is generally infinite. We fix a set $\mathcal{R} = \{R_k : k \geq 0\}$ of representatives, assuming for simplicity that this set is countable (it is the case e.g. if $G$ is a Lie group). Here $R_0$ is the trivial representation, i.e. $V = \mathbb{C}$ with $gz = z$ for all $g \in G$, $z \in \mathbb{C}$.

However, infinite dimensional representations arise naturally in the study of infinite compact groups. We shall be primarily concerned with one of them, namely the representation of $G$ defined by the space $\mathcal{C}(M)$ with the action of $G$ defined by $\pi(g)(f)(x) = f(g^{-1}x)$ (of course it has infinite dimension only if $M$ itself is not a finite set). It is a unitary representation for the standard inner product on $\mathcal{C}(M)$:

$$\langle f_1, f_2\rangle = \frac{1}{\mu(M)} \int_M f_1(x)\overline{f_2(x)} d\mu(x).$$

Moreover $\mathcal{C}(M)$ is a dense subspace of the *Hilbert space* $L^2(M)$ of measurable functions $f$ on $M$ such that $|f|^2$ is $\mu$-integrable (see e.g. Conway [25, Chapter 1] or Rudin [83, Chapter 4]). It turns out that the theory of finite dimensional representations of compact groups extends nicely to the unitary representations on Hilbert spaces:

**Theorem 3.2** *Every unitary representation of a compact group $G$ on a Hilbert space is the direct sum (in the sense of Hilbert spaces) of finite dimensional irreducible $G$-subspaces.*

The above theorem is in fact a consequence of the celebrated *Theorem of Peter and Weyl* (see e.g. Bump [21]), which states that the matrix coefficients of finite dimensional irreducible representations of a compact group $G$ span a subspace of $\mathcal{C}(G)$ which is dense for the topology of uniform convergence.

### 3.3 $G$-invariant positive definite functions on $M$

Now we come to our main concern, which is to describe the elements of $\mathcal{C}(M^2)^G_{\succeq 0}$, i.e. the $G$-invariant positive definite functions, in terms of the structure of the $G$-space $M$. In order to avoid technicalities with convergence issues, we shall systematically work in finite dimensional subspaces of $\mathcal{C}(M)$. So let $V$ be a $G$-subspace of $\mathcal{C}(M)$ of finite dimension. Let $V^{(2)}$ denote the subspace of $\mathcal{C}(M^2)$ spanned by elements of the form $f_1(x)\overline{f_2(y)}$, where $f_1$ and $f_2$ belong to $V$.

We take the following notations for an irreducible decomposition of $V$:

$$V = \bigoplus_{k \in I_V} \bigoplus_{i=1}^{m_k} H_{k,i}, \tag{8}$$

where for $1 \leq i \leq m_k$, the subspaces $H_{k,i}$ are pairwise orthogonal and $G$-isomorphic to $R_k$, and where $I_V$ is the finite set of indices $k \geq 0$ such that the multiplicities $m_k$ are not equal to 0. For all $k$, $i$, we choose an orthonormal basis $(e_{k,i,1}, \ldots, e_{k,i,d_k})$ of $H_{k,i}$, where $d_k = \dim(R_k)$, such that the complex numbers $\langle g e_{k,i,s}, e_{k,i,t} \rangle$ do not depend on $i$ (such a basis exists precisely because the $G$-isomorphism class of $H_{k,i}$ does not depend on $i$). From this data, we introduce the $m_k \times m_k$ matrices $E_k(x,y)$ with coefficients $E_{k,ij}(x,y)$:

$$E_{k,ij}(x,y) = \sum_{s=1}^{d_k} e_{k,i,s}(x)\overline{e_{k,j,s}(y)}. \tag{9}$$

Then $V^{(2)}$ obviously contains the elements $e_{k,i,s}(x)\overline{e_{k,j,s}(y)}$, which moreover form an orthonormal basis of this space.

**Theorem 3.3** *Let $F \in V^{(2)}$. Then $F$ is a $G$-invariant positive definite function if and only if there exist Hermitian positive semidefinite matrices $F_k$ such that*

$$F(x,y) = \sum_{k \in I_V} \langle F_k, \overline{E_k(x,y)} \rangle. \tag{10}$$

Before we give a sketch of the proof, several comments are in order:

1. If $M = [n]$, then we recover the parametrization of $G$-invariant positive semidefinite matrices given in (7).
2. The main point of the above theorem is to replace the property of a continuous function being positive definite with the property that the $m_k \times m_k$ matrices $F_k$ are positive semidefinite. In the introduction, we have already seen in the finite case $M = [n]$ that it allows to reduce the size of invariant semidefinite programs (following Step 2 (second version) of the introduction). In the general case, this theorem allows to replace a conic linear program involving the cone $\mathcal{C}(M^2)_{\succeq 0}$ with standard semidefinite programs (see Section 6 for examples where $M$ is the unit sphere of Euclidean space).

3. One might wonder to what extent the above matrices $E_k(x,y)$ depend on the choices involved in their definition. Indeed, one can prove that a different choice of orthonormal basis in $H_{k,i}$ does not affect $E_k(x,y)$. One could also start from another decomposition of the isotypic component $\mathcal{MI}_k$ of $V$. The new decomposition does not need to be orthogonal; the matrix $E_k(x,y)$ would then change to $AE_k(x,y)A^*$ for some $A \in \mathrm{Gl}_{m_k}(\mathbb{C})$. So, up to an action of $\mathrm{Gl}_{m_k}(\mathbb{C})$, the matrix $E_k(x,y)$ is canonically associated to $\mathcal{MI}_k$.
4. The statement implies that the matrices $E_k(x,y)$ themselves are $G$-invariant. In other words, they can be expressed in terms of the orbits $O(x,y)$ of the action of $G$ on $M^2$ in the form:

$$E_k(x,y) = Y_k(O(x,y)) \tag{11}$$

for some matrices $Y_k$. It is indeed the expression we are seeking for.

*Proof.* $F(x,y)$ is a linear combination of the $e_{k,i,s}(x)\overline{e_{l,j,t}(y)}$ since these elements form an orthonormal basis of $V^{(2)}$. In a first step, one shows that the $G$-invariance property $F(gx,gy) = F(x,y)$ results in an expression for $F$ of the form (10) for some matrices $F_k = (f_{k,ij})$; the proof involves the *orthogonality relations* between matrix coefficients of irreducible representations of $G$ (see e.g. Bump [21]). In a second step, $F \succeq 0$ is shown to be equivalent to $F_k \succeq 0$ for all $k \in I_V$. Indeed, for $\alpha(x) = \sum_{i=1}^{m_k} \alpha_i \overline{e_{k,i,s}(x)}$, we have

$$\sum_{i,j=1}^{m_k} \alpha_i f_{k,ij} \overline{\alpha_j} = \frac{1}{\mu(M)^2} \int_{M^2} \alpha(x) F(x,y) \overline{\alpha(y)} d\mu(x,y) \geq 0.$$

□

**Remark 3.4** *1. A straightforward generalization of the main structure theorem for matrix $*$-algebras to the $*$-algebra $\mathrm{End}^G(V)$ shows that*

$$\mathrm{End}^G(V) \simeq (V^{(2)})^G \simeq \bigoplus_{k \in I_V} \mathbb{C}^{m_k \times m_k}. \tag{12}$$

*An isomorphism from $(V^{(2)})^G$ to the direct sum of matrix algebras is constructed in the proof of Theorem 3.3, in an explicit way, from the decomposition (8). This isomorphism is given by the map sending $(F_k)_{k\in I_V} \in \oplus_{k \in I_V} \mathbb{C}^{m_k \times m_k}$ to $F \in (V^{(2)})^G$ defined by:*

$$F(x,y) = \sum_{k \in I_V} \langle F_k, \overline{E_k(x,y)} \rangle$$

*thus completing Step 2 (second version) of the introduction for compact groups.*

2. *Theorem 3.3 shows, jointly with Theorem 3.2, that any element of the cone* $\mathcal{C}(M^2)^G_{\succeq 0}$ *is a sum (in the sense of* $L^2$ *convergence) of the form* (10) *with possibly infinitely many terms. Moreover, under the additional assumption that* $G$ *acts transitively on* $M$*, Bochner [17] proves that the convergence holds in the stronger sense of* uniform convergence*, i.e. for the supremum norm.*

### 3.4 Examples: the commutative case

We mean here the case when $\mathrm{End}^G(\mathcal{C}(M))$ is a commutative algebra. From Theorem 3.3, this condition is equivalent to the commutativity of $\mathrm{End}^G(V)$ for all finite dimensional $G$-subspace of $\mathcal{C}(M)$. So, from the isomorphism (12), and since a matrix algebra $\mathbb{C}^{m\times m}$ is commutative if and only if $m = 0$ or $m = 1$, the commutative case corresponds to non vanishing multiplicities equal to 1 in the decomposition of $\mathcal{C}(M)$. The $G$-space $\mathcal{C}(M)$ is said to be multiplicity-free. Then the matrices $F_k$ in (10) have only one coefficient $f_k$ and $F \in \mathcal{C}(M^2)^G_{\succeq 0}$ is of the form

$$F(x,y) = \sum_{k\geq 0} f_k E_k(x,y), \quad f_k \geq 0. \tag{13}$$

We say that $M$ is *$G$-symmetric*, if $G$ acts *transitively on $M$* (i.e. there is only one orbit in $M$; equivalently $M$ is said to be a *homogeneous space*), and if moreover, for all $(x,y) \in M^2$, there exists $g \in G$ such that $gx = y$ and $gy = x$. In other words $(y,x)$ belongs to the $G$-orbit $O(x,y)$ of the pair $(x,y)$. This nonstandard terminology (sometimes it is also called a *generously transitive* group action) covers the case of *compact symmetric Riemannian manifolds* (see e.g. Berger [16]) as well as a large number of finite spaces. We provide many examples of such spaces below. The functions $Y_k$ (11) are often referred to as the *spherical functions* or *zonal spherical functions* of the homogeneous space $M$ (see e.g. Vilenkin, Klimyk [98]).

**Lemma 3.5** *If $M$ is $G$-symmetric then for all $V \subset \mathcal{C}(M)$, $(V^{(2)})^G$ is commutative.*

*Proof.* Since $F \in (V^{(2)})^G$ is $G$-invariant and $M$ is $G$-symmetric, $F(x,y) = F(y,x)$. The multiplication on the algebra $(V^{(2)})^G \simeq \mathrm{End}^G(M)$ is the convolution $F * F'$ of functions, given by:

$$(F * F')(x,y) = \int_M F(x,z)F'(z,y)d\mu(z).$$

A straightforward computation shows that $F * F' = F' * F$. □

An important subclass of symmetric spaces is provided by the *two-point homogeneous spaces.* These spaces are metric spaces, with distance $d(x,y)$ invariant by $G$, and moreover satisfy that two pairs of elements $(x,y)$ and $(x',y')$ belong to the same $G$-orbit if and only if $d(x,y) = d(x',y')$. They are obviously

$G$-symmetric since $d(x,y) = d(y,x)$. In other words, the distance function parametrizes the orbits of $G$ acting on pairs, thus the expression (13) of positive definite functions specializes to

$$F(x,y) = \sum_{k\geq 0} f_k Y_k(d(x,y)), \quad f_k \geq 0, \tag{14}$$

for some functions $Y_k$ in one variable. The functions $Y_k$ are explicitly known for many spaces, and are usually given by a certain family of *orthogonal polynomials* (see e.g. Szegö [91], Andrews, Askey, Roy [1]). The orthogonality property arises because the associated irreducible subspaces are pairwise orthogonal.

The complete list of real, compact, connected two-point homogeneous spaces was established by Wang [99]: the unit sphere of Euclidean space $S^{n-1} \subset \mathbb{R}^n$, which is treated in Section 6, the projective spaces over the fields of real, complex and quaternion numbers (denoted $P^{n-1}(K)$, with respectively $K = \mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$), and the projective plane over the octonions $P^2(\mathbb{O})$, are the only spaces having these properties. In each of these cases the functions $Y_k$ are given by *Jacobi polynomials* (see Table 1). These polynomials depend on two parameters $(\alpha, \beta)$ and are orthogonal for the measure $(1-t)^\alpha(1+t)^\beta dt$ on the interval $[-1,1]$. We denote by $P_k^{(\alpha,\beta)}(t)$ the Jacobi polynomial of degree $k$, normalized by the condition $P_k^{(\alpha,\beta)}(1) = 1$. When $\alpha = \beta = \lambda - 1/2$, these polynomials are equal (up to a nonnegative multiplicative factor) to the *Gegenbauer polynomials* $C_k^\lambda(t)$.

| Space | Group | $\alpha$ | $\beta$ |
|---|---|---|---|
| $S^{n-1}$ | $O_n(\mathbb{R})$ | $(n-3)/2$ | $(n-3)/2$ |
| $P^{n-1}(\mathbb{R})$ | $O_n(\mathbb{R})$ | $(n-3)/2$ | $-1/2$ |
| $P^{n-1}(\mathbb{C})$ | $U_n(\mathbb{C})$ | $n-2$ | $0$ |
| $P^{n-1}(\mathbb{H})$ | $U_n(\mathbb{H})$ | $2n-3$ | $1$ |
| $P^2(\mathbb{O})$ | $F_4(\mathbb{R})$ | $7$ | $3$ |

**Table 1.** The real compact two-point homogeneous spaces, their groups and their spherical functions $Y_k$.

The finite two-point homogeneous spaces are not completely classified, but several of them are relevant spaces for coding theory. The most prominent one is the Hamming space $\mathbf{q}^n$, associated to the Krawtchouk polynomials, which is discussed in Section 4. It is the space of $n$-tuples, over a finite set $\mathbf{q}$ of $q$ elements. An element of the Hamming space $\mathbf{q}^n$ is usually called a word, and $\mathbf{q}$ is called the alphabet. The closely related *Johnson spaces* are the subsets of binary words of fixed length and weight. The Hamming and Johnson spaces are endowed with the *Hamming distance*, counting the number of coordinates where to words disagree. The Johnson spaces have $q$-analogues, the *$q$-Johnson spaces*, the subsets of linear subspaces of $\mathbb{F}_q^n$ of fixed dimension. The distance between two subspaces of the same dimension is measured by the difference between this dimension and the dimension of their intersection. Other spaces are related to finite geometries and to spaces of matrices. In the later case, the distance between two matrices will be the rank of the difference of the matrices.

The spherical functions have been worked out, initially in the framework of association schemes with the work of Delsarte [29], and later in relation with harmonic analysis of finite classical groups with the work of Delsarte, Dunkl, Stanton and others. We refer to Conway, Sloane [26, Chapter 9] and to Stanton [88] for surveys on these topics. Table 2 displays the most important families of these spaces, together with their groups of isometries and the family of orthogonal polynomials to which their spherical functions belong; see Stanton [88] for additional information.

| Space | Group | Polynomial | Reference |
|---|---|---|---|
| Hamming space $\mathbf{q}^n$ | $S_q \wr S_n$ | Krawtchouk | [29] |
| Johnson space | $S_n$ | Hahn | [29],[31] |
| $q$-Johnson space | $\mathrm{Gl}_n(\mathbb{F}_q)$ | $q$-Hahn | [31] |
| *Maximal totally isotropic subspaces of dimension $k$, for a nonsingular bilinear form:* | | | |
| Symmetric | $\mathrm{SO}_{2k+1}(\mathbb{F}_q)$ | $q$-Krawtchouk | [88] |
| | $\mathrm{SO}_{2k}(\mathbb{F}_q)$ | $q$-Krawtchouk | [88] |
| | $\mathrm{SO}^-_{2k+2}(\mathbb{F}_q)$ | $q$-Krawtchouk | [88] |
| Symplectic | $\mathrm{Sp}_{2k}(\mathbb{F}_q)$ | $q$-Krawtchouk | [88] |
| Hermitian | $\mathrm{SU}_{2k}(\mathbb{F}_{q^2})$ | $q$-Krawtchouk | [88] |
| | $\mathrm{SU}_{2k+1}(\mathbb{F}_{q^2})$ | $q$-Krawtchouk | [88] |
| *Spaces of matrices:* | | | |
| $\mathbb{F}_q^{k\times n}$ | $\mathbb{F}_q^{k\times n}.(\mathrm{Gl}_k(\mathbb{F}_q)\times \mathrm{Gl}_n(\mathbb{F}_q))$ | Affine $q$-Krawtchouk | [32], [88] |
| $\mathrm{Skew}_n(\mathbb{F}_q)$ skew-symmetric | $\mathrm{Skew}_n(\mathbb{F}_q).\,\mathrm{Gl}_n(\mathbb{F}_q)$ | Affine $q$-Krawtchouk | [30], [88] |
| $\mathrm{Herm}_n(\mathbb{F}_{q^2})$ Hermitian | $\mathrm{Herm}_n(\mathbb{F}_{q^2}).\,\mathrm{Gl}_n(\mathbb{F}_{q^2})$ | Affine $q$-Krawtchouk | [88] |

**Table 2.** Some finite two-point homogeneous spaces, their groups and their spherical functions $Y_k$.

Symmetric spaces which are not two-point homogeneous give rise to functions $Y_k$ depending on several variables. A typical example is provided by the real Grassmann spaces, the spaces of $m$-dimensional subspaces of $\mathbb{R}^n$, $m \leq n/2$, under the action of the orthogonal group $O(\mathbb{R}^n)$. Then, the functions $Y_k$ are multivariate Jacobi polynomials, with $m$ variables representing the principal angles between two subspaces (see James, Constantine [49]). A similar situation occurs in coding theory when nonhomogeneous alphabets are considered, involving multivariate Krawtchouk polynomials. Table 3 shows some examples of these spaces with references to coding theory applications.

| Space | Group | Polynomial | Reference |
|---|---|---|---|
| Nonbinary Johnson | $S_{q-1} \wr S_n$ | Eberlein/Krawtchouk | [93] |
| Permutation group | $S_n^2$ | Characters of $S_n$ | [94] |
| Lee space | $D_q \wr S_n$ | Multivariate Krawtchouk | [3] |
| Ordered Hamming space | $(\mathbb{F}_q^k.B_k) \wr S_n$ | Multivariate Krawtchouck | [71], [14] |
| Real Grassmann space | $O_n(\mathbb{R})$ | Multivariate Jacobi | [49], [4] |
| Complex Grassmann space | $U_n(\mathbb{C})$ | Multivariate Jacobi | [49], [81], [82] |
| Unitary group | $U_n(\mathbb{C})^2$ | Schur | [27] |

**Table 3.** Some symmetric spaces, their groups and their spherical functions $Y_k$.

### 3.5 Examples: the noncommutative case

Only a few cases have been completely worked out. Among them, the Euclidean sphere $S^{n-1}$ under the action of the subgroup of the orthogonal group fixing one point is treated in Bachoc, Vallentin [6]. We come back to this case in Section 6. These results have been extended to the case of fixing many points in Musin [75]. The case of the binary Hamming space under the action of the symmetric group is treated in the framework of group representations in Vallentin [96], shedding a new light on the results of Schrijver [86]. A very similar case is given by the set of linear subspaces of the finite vector space $\mathbb{F}_q^n$, treated as a $q$-analogue of the binary Hamming space in Bachoc, Vallentin [8].

## 4 Block codes

In this section, we give an overview of symmetry reduction in semidefinite programming bounds for (error correcting) codes. We fix an alphabet $\mathbf{q} = \{0, 1, \ldots, q-1\}$ for some integer $q \geq 2$. The *Hamming space* $\mathbf{q}^n$ is equipped with a metric $d(\cdot,\cdot)$ called the *Hamming distance* which is given by

$$d(u, v) = |\{i : u_i \neq v_i\}| \text{ for all } u, v \in \mathbf{q}^n.$$

The isometry group of the Hamming space will be denoted $\mathrm{Aut}(q, n)$ and is a wreath product $S_q \wr S_n$ of symmetric groups. That is, the isometries are obtained by taking a permutation of the $n$ positions followed by independently permuting the symbols $0, \ldots, q-1$ at each of the $n$ positions. The group acts transitively on $\mathbf{q}^n$ and the orbits of pairs are given by the Hamming distance.

A subset $C \subseteq \mathbf{q}^n$ is called a *code* of length $n$. For a (nonempty) code $C$, $\min\{d(u, v) : u, v \in C \text{ distinct}\}$ is the *minimum distance* of $C$. An important quantity in coding theory is the maximum size $A_q(n, d)$ of a code of length $n$ and minimum distance at least $d$:

$$A_q(n, d) = \max\{|C| : C \subseteq \mathbf{q}^n \text{ has minimum distance at least } d\}.$$

These codes are often called *block codes* since they can be used to encode messages into fixed-length blocks of words from a code $C$. The most studied case is that of *binary codes*, that is $q = 2$ and $\mathbf{q} = \{0, 1\}$. In this case $\mathbf{q}$ is suppressed from the notation. An excellent reference work on coding theory is MacWilliams, Sloane [70].

Lower bounds for $A_q(n, d)$ are mostly obtained using explicit construction of codes. Our focus here is on upper bounds obtained using semidefinite programming.

### 4.1 Lovász' $\vartheta$ and Delsarte's linear programming bound

Let $\Gamma_q(n, d)$ be the graph on vertex set $\mathbf{q}^n$, connecting two words $u, v \in \mathbf{q}^n$ if $d(u, v) < d$. Then the codes of minimum distance at most $d$ correspond exactly to the independent sets in $\Gamma_q(n, d)$ and hence

$$A_q(n, d) = \alpha(\Gamma_q(n, d)),$$

where $\alpha$ denotes the independence number of a graph. The graph parameter $\vartheta'$, defined in Schrijver [86], is a slight modification of the Lovász theta number discussed in more detail in Section 6. It can be defined by the following semidefinite program

$$\vartheta'(\Gamma) = \max\{\langle X, J\rangle : \langle X, I\rangle = 1, X_{uv} = 0 \text{ if } uv \in E(\Gamma), X \geq 0, X \succeq 0\},$$

and it is a by-now classical lemma of Lovász [69] that for any graph $\Gamma$ the inequality $\alpha(\Gamma) \leq \vartheta'(\Gamma) \leq \vartheta(\Gamma)$ holds.

Hence $\vartheta'(\Gamma_q(n, d))$ is an (exponential size) semidefinite programming upper bound for $A_q(n, d)$. Using symmetry reduction, this program can be solved efficiently. For this, observe that the SDP is invariant under the action of $\mathrm{Aut}(q, n)$. Hence we may restrict $X$ to belong to the matrix $*$-algebra $\mathcal{A}$ of invariant matrices (the Bose-Mesner algebra of the Hamming scheme). The zero-one matrices $A_0, \ldots, A_n$ corresponding to the orbits of pairs:

$$(A_i)_{u,v} = \begin{cases} 1 & \text{if } d(u, v) = i, \\ 0 & \text{otherwise,} \end{cases}$$

form the canonical basis of $\mathcal{A}$. Writing $X = \frac{1}{q^n}(x_0 A_0 + \cdots + x_n A_n)$, we obtain an SDP in $n + 1$ variables:

$$\vartheta'(\Gamma_q(n, d)) = \max\Big\{ \sum_{i=0}^{n} x_i \tbinom{n}{i} : x_0 = 1, x_1 = \cdots = x_{d-1} = 0,$$
$$x_d, \ldots, x_n \geq 0,\ \sum_{i=0}^{n} x_i A_i \succeq 0 \Big\}.$$

The second step is to block diagonalize the algebra $\mathcal{A}$. In this case $\mathcal{A}$ is commutative, which means that the $A_i$ can be simultaneously diagonalized, reducing the positive semidefinite constraint to the linear constraints

$$\sum_{i=0}^{n} x_i P_i(j) \geq 0 \text{ for } j = 0, \ldots, n,$$

in the eigenvalues $P_i(j)$ of the matrices $A_i$. The $P_i$ are the *Krawtchouk polynomials* and are given by

$$P_i(x) = \sum_{k=0}^{i} (-1)^k \binom{x}{k} \binom{n-x}{i-k} (q-1)^{i-k}.$$

Thus the semidefinite program is reduced to a linear program, which is precisely the *Delsarte bound* [29]. This relation between $\vartheta'$ and the Delsarte bound was recognized independently in McEliece, Rodemich, Rumsey [72] and Schrijver [85]. In the more general setting of (commutative) association schemes, Delsarte's linear programming bound is obtained from a semidefinite program by symmetry reduction and diagonalizing the Bose-Mesner algebra.

### 4.2 Stronger bounds through triples

Delsarte's bound is based on the distance distribution of a code, that is, on the number of code word-pairs for each orbit of pairs. Using orbits of triples, the bound can be tightened as was shown in Schrijver [86]. We describe this method for the binary case ($q = 2$).

Denote by $\mathrm{Stab}(0, \mathrm{Aut}(2, n)) \subseteq \mathrm{Aut}(2, n)$ the stabilizer of 0. That is, $\mathrm{Stab}(0, \mathrm{Aut}(2, n))$ consists of just the permutations of the $n$ positions.

For any $u, v, w \in \{0, 1\}^n$, denote by $O(u, v, w)$ the orbit of the triple $(u, v, w)$ under the action of $\mathrm{Aut}(2, n)$. Observe that since $\mathrm{Aut}(2, n)$ acts transitively on $\{0, 1\}^n$, the orbits of triples are in bijection with the orbits of pairs under $\mathrm{Stab}(0, \mathrm{Aut}(2, n))$, since we may assume that $u$ is mapped to 0. There are $\binom{n+3}{3}$ such orbits, indexed by integers $0 \leq t \leq i, j \leq n$. Indeed, the orbit of $(v, w)$ under $\mathrm{Stab}(0, \mathrm{Aut}(2, n))$ is determined by the sizes $i, j$ and $t$ of respectively the supports of $v$ and $w$ and their intersection.

Let $C \subseteq \{0, 1\}^n$ be a code. We denote by $\lambda_{i,j}^t$ the number of triples $(u, v, w) \in C^3$ in the orbit indexed by $i, j$ and $t$. This is the analogue for triples of the distance distribution.

Denote by $M_C$ the zero-one $\{0, 1\}^n \times \{0, 1\}^n$ matrix defined by $(M_C)_{u,v} = 1$ if and only if $u, v \in C$. So $M_C$ is a rank 1 positive semidefinite matrix. Now consider the following two matrices

$$M' = \sum_{\sigma \in \mathrm{Aut}(2,n), 0 \in \sigma C} M_{\sigma C}, \qquad M'' = \sum_{\sigma \in \mathrm{Aut}(2,n), 0 \notin \sigma C} M_{\sigma C}.$$

Clearly, $M'$ and $M''$ are nonnegative and positive semidefinite. Furthermore, by construction $M'$ and $M''$ are invariant under the action of the stabilizer $\mathrm{Stab}(0, \mathrm{Aut}(2,n))$. Equivalently, any entry $M'_{uv}$ (or $M''_{uv}$) only depends on the orbit of $(0,u,v)$ under the action of $\mathrm{Aut}(2,n)$. In fact, $M'_{uv}$ equals (up to a factor depending on the orbit) the number of triples $(c,c',c'') \in C^3$ in that orbit.

The matrix $*$-algebra $\mathcal{A}_n$ of complex $\{0,1\}^n \times \{0,1\}^n$ matrices invariant under $\mathrm{Stab}(0, \mathrm{Aut}(2,n))$ has a basis of zero-one matrices corresponding to the orbits of pairs under the action of $\mathrm{Stab}(0, \mathrm{Aut}(2,n))$: $\{A^t_{i,j} : 0 \le t \le i,j \le n\}$. This algebra is called the *Terwilliger algebra* of the (binary) Hamming scheme. The facts about $M'$ and $M''$ above lead to a semidefinite programming bound for $A(n,d)$, in terms of variables $x^t_{i,j}$. Nonnegativity of $M', M''$ translates into nonnegativity of the $x^t_{i,j}$. Excluding positive distances smaller than $d$ translates into setting $x^t_{i,j} = 0$ for orbits containing two words at distance 1 through $d-1$. Semidefiniteness of $M'$ and $M''$ translate into

$$\sum_{i,j,t} x^t_{i,j} A^t_{i,j} \succeq 0, \qquad \sum_{i,j,t} (x^0_{i+j-t,0} - x^t_{i,j}) A^t_{i,j} \succeq 0.$$

There are some more constraints. In particular, for $u,v,w \in \{0,1\}^n$ the variables of the orbits of $(u,v,w)$ and $(w,u,v)$ (or any other reordering) must be equal. This further reduces the number of variables. For more details, see Schrijver [86].

Having reduced the SDP-variables by using the problem symmetry, the second step is to replace the positive semidefinite constraints by equivalent conditions on smaller matrices using a block diagonalization of $\mathcal{A}_n$. The block diagonalization of $\mathcal{A}_n$ is given explicitly in Schrijver [86] (see also Vallentin [96], Dunkl [36]). The number of blocks equals $1 + \lfloor \frac{n}{2} \rfloor$, with sizes $n+1-2k$ for $k = 0, \ldots, \lfloor \frac{n}{2} \rfloor$. Observe that the sum of the squares of these numbers indeed equals $\binom{n+3}{3}$, the dimension of $\mathcal{A}_n$.

For the non-binary case, an SDP upper bound for $A_q(n,d)$ can be defined similarly. The main differences are, that the orbits or triples are now indexed by four parameters $i,j,p,t$ satisfying $0 \le p \le t \le i,j$, $i+j-t \le n$. Hence the corresponding invariant algebra is a matrix $*$-algebra of dimension $\binom{n+4}{4}$. In that case the block diagonalization is a bit more involved having blocks indexed by the integers $a,k$ with $0 \le a \le k \le n+a-k$, and size $n+a+1-2k$. See Gijswijt, Schrijver, Tanaka [42] for details.

### 4.3 Hierarchies and $k$-tuples

#### Moment matrices

Let $V$ be a finite set and denote by $\mathcal{P}(V)$ its power set. For any $y : \mathcal{P}(V) \to \mathbb{C}$, denote by $M(y)$ the $\mathcal{P}(V) \times \mathcal{P}(V)$ matrix given by

$$[M(y)]_{S,T} = y_{S \cup T}.$$

Such a matrix is called a *combinatorial moment matrix* (see Laurent [66]).

Let $C \subseteq V$ and denote by $x = \chi^C \in \{0,1\}^V$ the zero-one incidence vector of $C$. The vector $x$ can be "lifted" to $y \in \{0,1\}^{\mathcal{P}(V)}$ by setting $y_S = \prod_{v \in S} x_v$ for all $S \in \mathcal{P}(V)$. Note that $y_\emptyset = 1$. Then the moment matrix $M(y)$ is positive semidefinite since $M(y) = yy^{\mathsf{T}}$. This observation has an important converse (see Laurent [66]).

**Theorem 1.** *Let $M(y)$ be a moment matrix with $y_\emptyset = 1$ and $M(y) \succeq 0$. Then $M(y)$ is a convex combination of moment matrices $M(y_1), \ldots, M(y_k)$, where $y_i$ is obtained from lifting a vector in $\{0,1\}^V$.*

Let $\Gamma = (V, E)$ be a graph on $V$. Theorem 1 shows that the independence number is given by

$$\begin{aligned}\alpha(\Gamma) = \max \Big\{ \sum_{v \in V} y_{\{v\}} : y_\emptyset = 1, y_S = 0 \text{ if } S \text{ contains an edge}, \\ M(y) \succeq 0 \Big\}.\end{aligned} \tag{15}$$

By replacing $M(y)$ by suitable principal submatrices, this (exponential size) semidefinite program can be relaxed to obtain more tractable upper bounds on $\alpha(\Gamma)$. For example, restricting rows and columns to sets of size at most 1 (and restricting $y$ to sets of size $\leq 2$), we obtain the Lovász $\vartheta$ number (adding nonnegativity of $y_S$ for $S$ of size 2 gives $\vartheta'$).

To describe a large class of useful submatrices, fix a nonnegative integer $s$ and a set $T \subseteq V$. Define $M_{s,T} = M_{s,T}(y)$ to be the matrix with rows and columns indexed by sets $S$ of size at most $s$, defined by

$$[M_{s,T}]_{S,S'} = y_{S \cup S' \cup T}. \tag{16}$$

So $M_{s,T}$ is a principal submatrix of $M(y)$, except that possibly some rows and columns are duplicated. For fixed $k$, we can restrict $y$ to subsets of size at most $k$ and replace $M(y) \succeq 0$ in (15) by conditions

$$M_{s,T} \succeq 0 \quad \text{for all } T \text{ of size at most } t,$$

where $2s+t \leq k$. This yields upper bounds on $\alpha(\Gamma)$ defined in terms of subsets of $V$ of size at most $k$. The constraints (16) can be strengthened further to

$$\tilde{M}_{s,T} = \big(M_{s,T' \cup T''}\big)_{T',T'' \subseteq T} \succeq 0 \quad \text{for all } T \text{ of size } t. \tag{17}$$

Here $\tilde{M}_{s,T}$ is a $2^{|T|} \times 2^{|T|}$ matrix of smaller matrices, which can be seen as a submatrix of $M(y)$ (again up to duplication of rows and columns). By the moment-structure of $\tilde{M}_{s,T}$, condition (17) is equivalent to

$$\sum_{R' \subseteq R} (-1)^{|R \setminus R'|} M_{s,R'} \succeq 0 \quad \text{for all } R \text{ of size at most } t.$$

Using SDP constraints of the form (16) and (17) with $2s + t \leq k$, we can obtain the (modified) theta number ($k = 2$), Schrijver's bound using triples ($k = 3$), or rather a strengthened version from Laurent [65], and the bound from Gijswijt, Mittelmann, Schrijver [41] ($k = 4$).

Several hierarchies of upper bounds have been proposed. The hierarchy introduced by Lasserre [64] (also see Laurent [65]) is obtained using (16) by fixing $t = 0$ and letting $s = 1, 2, \dots$ run over the positive integers. The hierarchy described in Gijswijt, Mittelmann, Schrijver [41] is obtained using (16) by letting $k = 1, 2, \dots$ run over the positive integers and restricting $s$ and $t$ by $2s+t \leq k$. Finally, the hierarchy defined in Gvozdenović, Laurent, Vallentin [47] fixes $s = 1$ and considers the constraints (17), where $t = 0, 1, \dots$ runs over the nonnegative integers. There, also computational results are obtained for Paley graphs in the case $t = 1, 2$.

A slight variation is obtained by restricting $y$ to sets of at least one element, deleting the row and column corresponding to the empty set from (submatrices of) $M(y)$. The normalization $y_\emptyset = 1$ is then replaced by $\sum_{v \in V} y_{\{v\}} = 1$ and the objective is replaced by $\sum_{u,v \in V} y_{\{u,v\}}$. Since $M_{1,\emptyset} \succeq 0$ implies that $(\sum_{v \in V} y_{\{v\}})^2 \leq y_\emptyset \sum_{u,v \in V} y_{\{u,v\}}$, this gives a relaxation. This variation was used in Schrijver [86] and Gijswijt, Schrijver, Tanaka [42].

**Symmetry reduction**

Application to the graph $\Gamma = \Gamma_q(n,d)$ with vertex set $V = \mathbf{q}^n$, gives semidefinite programming bounds for $A_q(n,d)$. A priory, this gives an SDP that has exponential size in $n$. However, using the symmetries of the Hamming space, it can be reduced to polynomial size as follows.

The action of $\mathrm{Aut}(q,n)$ on the Hamming space $V$, extends to an action on $\mathcal{P}(V)$ and hence on vectors and matrices indexed by $\mathcal{P}(V)$. Since the semidefinite program in (15) is invariant under $\mathrm{Aut}(q,n)$, we may restrict $y$ (and hence $M(y)$) to be invariant under $\mathrm{Aut}(q,n)$, without changing the optimum value. The same holds for any of the bounds obtained using principal submatrices. Thus $M(y)_{S,T}$ only depends on the orbit of $S \cup T$. In particular, for any fixed integer $k$, the matrices $M_{s,T}$ with $2s+|T| \leq k$ are described using variables corresponding to the orbits of sets of size at most $k$. Two matrices $M_{s,T}$ and $M_{s,T'}$ are equal (up to permutation of rows and columns) when $T$ and $T'$ are in the same orbit under $\mathrm{Aut}(q,n)$. Hence, the number of variables is bounded by the number of orbits of sets of size at most $k$ ($O(n^{2^{k-1}-1})$ in the binary case), and for each such orbit there are at most a constant number of distinct constraints of the form (16,17). This concludes the first step in the symmetry reduction.

The matrices $M_{s,T}$ are invariant under the subgroup of $\mathrm{Stab}(T, \mathrm{Aut}(q,n))$ that fixes each of the elements of $T$. The dimension of the matrix $*$-algebra of invariant matrices is polynomial in $n$ and hence the size of the matrices can be reduced to polynomial size using the regular $*$-representation.

By introducing more duplicate rows and columns and disregarding the row and column indexed by the empty set, we may index the matrices $M_{s,T}$ by *ordered* $s$-tuples and view $M_{s,T}$ as an $(\mathbf{q}^n)^s \times (\mathbf{q}^n)^s$ matrix. Let $\mathcal{A}_{s,T}$ be the matrix $*$-algebra of $(\mathbf{q}^n)^s \times (\mathbf{q}^n)^s$ matrices invariant under the the subgroup of $\mathrm{Stab}(T, \mathrm{Aut}(q,n))$. This matrix $*$-algebra can be block diagonalized using techniques from Gijswijt [40].

### 4.4 Generalizations

The semidefinite programs discussed in the previous sections, are based on the distribution of pairs, triples and $k$-tuples in a code. For each tuple (up to isometry) there is a variable that reflects the number of occurrences of that tuple in the code. Exclusion of pairs at distance $1, \ldots, d-1$, is modeled by setting variables for violating tuples to zero.

Bounds for other types of codes in the Hamming space can be obtained by setting the variables for excluded tuples to zero. This does not affect the underlying algebra of invariant matrices or the symmetry reduction. For triples in the binary Hamming space, this method was applied to orthogonality graphs in de Klerk, Pasechnik [58] and to pseudo-distances in Bachoc, Zémor [10]. Lower bounds on (nonbinary) covering codes have been obtained in Gijswijt [39] by introducing additional linear constraints.

Bounds on constant weight binary codes were given in Schrijver [86].

In the nonbinary Hamming space, the symbols $0, \ldots, q-1$ are all interchangeable. In the case of Lee-codes, the dihedral group $D_q$ acts on the alphabet, which leads to a smaller isometry group $D_q \wr S_n$. The Bose-Mesner algebra of the Lee-scheme is still commutative. The corresponding linear programming bound was studied in Astola [3]. To the best knowledge of the author, stronger bounds using triples have not been studied in this case.

## 5 Crossing numbers

We describe an application of the regular $*$-representation to obtain a lower bound on the crossing number of complete bipartite graphs. This was described in de Klerk, Pasechnik, Schrijver [59], and extends a method of de Klerk, et. al. [55].

The *crossing number* $\mathrm{cr}(\Gamma)$ of a graph $\Gamma$ is the minimum number of intersections of edges when $\Gamma$ is drawn in the plane such that all vertices are distinct. The *complete bipartite graph* $K_{m,n}$ is the graph with vertices $1, \ldots, m, u_1, \ldots, u_n$ and edges all pairs $iu_j$ for $i \in [m]$ and $j \in [n]$. (This notation will be convenient for our purposes.)

This relates to the problem raised by the paper of Zarankiewicz [101], asking if

$$\mathrm{cr}(K_{m,n}) \stackrel{?}{=} Z(m,n) = \lfloor \tfrac{1}{4}(m-1)^2 \rfloor \lfloor \tfrac{1}{4}(n-1)^2 \rfloor. \tag{18}$$

In fact, Zarankiewicz claimed to have a proof, which however was shown to be incorrect. In (18), $\leq$ follows from a direct construction. Equality was proved by Kleitman [53] if $\min\{m,n\} \leq 6$ and by Woodall [100] if $m \in \{7,8\}$ and $n \in \{7,8,9,10\}$.

Consider any $m,n$. Let $Z_m$ be the set of cyclic permutations of $[m]$ (that is, the permutations with precisely one orbit). For any drawing of $K_{m,n}$ in the plane and for any $u_i$, let $\gamma(u_i)$ be the cyclic permutation $(1,i_2,\ldots,i_m)$ such that the edges leaving $u_i$ in clockwise order, go to $1,i_2,\ldots,i_m$ respectively.

For $\sigma,\tau \in Z_m$, let $C_{\sigma,\tau}$ be equal to the minimum number of crossings when we draw $K_{m,2}$ in the plane such that $\gamma(u_1)=\sigma$ and $\gamma(u_2)=\tau$. This gives a matrix $C=(C_{\sigma,\tau})$ in $\mathbb{R}^{Z_m\times Z_m}$. Then the number $\alpha_m$ is defined by:

$$\alpha_m = \min\{\langle X,C\rangle : X \in \mathbb{R}_+^{Z_m\times Z_m}, X \succeq 0, \langle X,J\rangle = 1\}, \tag{19}$$

where $J$ is the all-one matrix in $\mathbb{R}^{Z_m\times Z_m}$.

Then $\alpha_m$ gives a lower bound on $\mathrm{cr}(K_{m,n})$, as was shown by de Klerk, et. al. [55].

**Theorem 5.1** *$cr(K_{m,n}) \geq \frac{1}{2}n^2\alpha_m - \frac{1}{2}n\lfloor\frac{1}{4}(m-1)^2\rfloor$ for all $m,n$.*

*Proof.* Consider a drawing of $K_{m,n}$ in the plane with $\mathrm{cr}(K_{m,n})$ crossings. For each cyclic permutation $\sigma$, let $d_\sigma$ be the number of vertices $u_i$ with $\gamma(u_i)=\sigma$. Consider $d$ as column vector in $\mathbb{R}^{Z_m}$, and define the matrix $X$ in $\mathbb{R}^{Z_m\times Z_m}$ by

$$X = n^{-2}dd^{\mathsf{T}}.$$

Then $X$ satisfies the constraints in (19), hence $\alpha_m \leq \langle X,C\rangle$. For $i,j = 1,\ldots,n$, let $\beta_{i,j}$ denote the number of crossings of the edges leaving $u_i$ with the edges leaving $u_j$. So if $i \neq j$, then $\beta_{i,j} \geq C_{\gamma(u_i),\gamma(u_j)}$. Hence

$$n^2\langle X,C\rangle = \langle dd^{\mathsf{T}}C\rangle) = d^{\mathsf{T}}Cd = \sum_{i,j=1}^{n} C_{\gamma(u_i),\gamma(u_j)}$$

$$\leq \sum_{\substack{i,j=1\\ i\neq j}}^{n} \beta_{i,j} + \sum_{i=1}^{n} C_{\gamma(u_i),\gamma(u_i)} = 2\mathrm{cr}(K_{m,n}) + n\lfloor\tfrac{1}{4}(m-1)^2\rfloor.$$

Therefore,

$$\mathrm{cr}(K_{m,n}) \geq \tfrac{1}{2}n^2\langle X,C\rangle - \tfrac{1}{2}n\lfloor\tfrac{1}{4}(m-1)^2\rfloor \geq \tfrac{1}{2}\alpha_m n^2 - \tfrac{1}{2}n\lfloor\tfrac{1}{4}(m-1)^2\rfloor. \quad \square$$

The semidefinite programming problem (19) defining $\alpha_m$ has order $(m-1)!$. Using the regular $*$-representation, we can reduce the size. Fix $m \in \mathbb{N}$. Let $G = S_m \times \{-1,+1\}$, and define $h : G \to S_{Z_m}$ by

$$h_{\pi,i}(\sigma) = \pi\sigma^i\pi^{-1}$$

for $\pi \in S_m$, $i \in \{-1, +1\}$, $\sigma \in Z_m$. So $G$ acts on $Z_m$. It is easy to see that $C$ and (trivially) $J$ are $G$-invariant.

Using the regular $*$-representation, $\alpha_m$ up to $m = 9$ was computed. For $m = 8$, $\alpha_8 = 5.8599856444\ldots$, implying

$$\mathrm{cr}(K_{8,n}) \geq 2.9299n^2 - 6n.$$

This implies for each fixed $m \geq 8$, with an averaging argument over all subgraphs $K_{8,n}$:

$$\lim_{n\to\infty} \frac{\mathrm{cr}(K_{m,n})}{Z(m,n)} \geq 0.8371 \frac{m}{m-1}.$$

Moreover, $\alpha_9 = 7.7352126\ldots$, implying

$$\mathrm{cr}(K_{9,n}) \geq 3.8676063n^2 - 8n,$$

and for each fixed $m \geq 9$:

$$\lim_{n\to\infty} \frac{\mathrm{cr}(K_{m,n})}{Z(m,n)} \geq 0.8594 \frac{m}{m-1}.$$

Thus we have an asymptotic approximation to Zarankiewicz's problem.

The orbits of the action of $G$ onto $Z_m \times Z_m$ can be identified as follows. Each orbit contains an element $(\sigma_0, \tau)$ with $\sigma_0 = (1, \ldots, m)$. So there are at most $(m-1)!$ orbits. Next, $(\sigma_0, \tau)$ and $(\sigma_0, \tau')$ belong to the same orbit if and only if $\tau' = \tau^g$ for some $g \in G$ that fixes $\sigma_0$. (There are only few of such $g$.) In this way, the orbits can be identified by computer, for $m$ not too large. The corresponding values $C_{\sigma,\tau}$ for each orbit $[\sigma, \tau]$, and the multiplication parameters, also can be found using elementary combinatorial algorithms. The computer does this all within a few minutes, for $m \leq 9$. But the resulting semidefinite programming problem took, in 2006, seven days. It was the largest SDP problem solved by then.

## 6 Spherical codes

Finding upper bounds for codes on the sphere is our next application. Here one deals with infinite-dimensional semidefinite programs which are invariant under the orthogonal group which is continuous and compact. So we are in the situation for which the techniques of Section 3 work.

We start with some definitions: The unit sphere $S^{n-1}$ of the Euclidean space $\mathbb{R}^n$ equipped with its standard inner product $x \cdot y = \sum_{i=1}^n x_i y_i$, is defined as usual by:

$$S^{n-1} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x \cdot x = 1\}.$$

It is a compact space, on which the orthogonal group $O_n(\mathbb{R})$ acts homogeneously. The stabilizer $\mathrm{Stab}(x_0, O_n(\mathbb{R}))$ of one point $x_0 \in S^{n-1}$ can be identified with the orthogonal group of the orthogonal complement $(\mathbb{R}x_0)^\perp$ of the line $\mathbb{R}x_0$ thus with $O_{n-1}(\mathbb{R})$, leading to an identification between the sphere and the quotient space $O_n(\mathbb{R})/O_{n-1}(\mathbb{R})$. The unit sphere is endowed with the standard $O_n(\mathbb{R})$-invariant Lebesgue measure $\mu$ with the normalization $\mu(S^{n-1}) = 1$. The *angular distance* $d_\theta(x, y)$ of $(x, y) \in (S^{n-1})^2$ is defined by

$$d_\theta(x, y) = \arccos(x \cdot y)$$

and is $O_n(\mathbb{R})$-invariant. Moreover, the metric space $(S^{n-1}, d_\theta)$ is *two-point homogeneous* (see Section 3.4) under $O_n(\mathbb{R})$.

The *minimal angular distance* $d_\theta(C)$ of a subset $C \subset S^{n-1}$ is by definition the minimal angular distance of pairs of distinct elements of $C$. It is a classical problem to determine the maximal size of $C$, subject to the condition that $d_\theta(C)$ is greater or equal to some given minimal value $\theta_{\min}$. This problem is the fundamental question of the theory of error correcting codes (see e.g. Conway, Sloane [26], Ericson, Zinoviev [37]). In this context, the subsets $C \subset S^{n-1}$ are referred to as *spherical codes*. In geometry, the case $\theta_{\min} = \pi/3$ corresponds to the famous *kissing number problem* which asks for the maximal number of spheres that can simultaneously touch a central sphere without overlapping, all spheres having the same radius (see e.g. Conway, Sloane [26]).

We introduce notations in a slightly more general framework. We say that $C \subset S^{n-1}$ *avoids* $\Omega \subset (S^{n-1})^2$ if, for all $(x, y) \in C^2$, $(x, y) \notin \Omega$. We define, for a measure $\lambda$,

$$A(S^{n-1}, \Omega, \lambda) = \sup\big\{\lambda(C) : C \subset S^{n-1} \text{ measurable, } C \text{ avoids } \Omega\big\}.$$

We have in mind the following situations of interest:

(i) $\Omega = \{(x, y) : d_\theta(x, y) \in ]0, \theta_{\min}[\}$ and $\lambda$ is the counting measure denoted $\mu_c$. Then, the $\Omega$-avoiding sets are exactly the spherical codes $C$ such that $d_\theta(C) \geq \theta_{\min}$.

(ii) $\Omega = \{(x, y) : d_\theta(x, y) = \theta\}$ for some value $\theta \neq 0$, and $\lambda = \mu$. Here we consider subsets avoiding only one value of distance, so these subsets can be infinite. This case has interesting connections with the famous problem of finding the *chromatic number of Euclidean space* (see Soifer [87], Bachoc et. al. [11], Oliveira, Vallentin [79], Oliveira [78])

(iii) $\Omega = \{(x, y) : d_\theta(x, y) \in ]0, \theta_{\min}[ \text{ or } (x, y) \notin \mathrm{Cap}(e, \phi)^2\}$, and $\lambda = \mu_c$. Here $\mathrm{Cap}(e, \phi)$ denotes the *spherical cap* with center $e$ and radius $\phi$:

$$\mathrm{Cap}(e, \phi) = \big\{x \in S^{n-1} : d_\theta(e, x) \leq \phi\big\}$$

and we are dealing with subsets of a spherical cap with given minimal angular distance.

The computation of $A(S^{n-1}, \Omega, \lambda)$ is a difficult and unsolved problem in general, so one aims at finding lower and upper bounds for this number. To that end, for finding upper bounds, we borrow ideas from combinatorial optimization. Indeed, if one thinks of the pair $(S^{n-1}, \Omega)$ being a graph with vertex set $S^{n-1}$ and edge set $\Omega$, then a set $C$ avoiding $\Omega$ corresponds to an *independent set*, and $A(S^{n-1}, \Omega, \lambda)$ to the *independence number* of this graph. In general, for a graph $\Gamma$ with vertex set $V$ and edge set $E$, an *independent set* is a subset $S$ of $V$ such that no pairs of elements of $S$ are connected by an edge, and the *independence number* $\alpha(\Gamma)$ of $\Gamma$ is the maximal number of elements of an independent set. The *Lovász theta number* $\vartheta(\Gamma)$ was introduced by Lovász [69]. It gives an upper bound of the independence number $\alpha(\Gamma)$ of a graph $\Gamma$, and it is the optimal solution of a semidefinite program. We review Lovász theta number in Section 6.1. Then, in Section 6.2, we introduce generalizations of this notion, in the form of conic linear programs, that provide upper bounds for $A(S^{n-1}, \Omega, \lambda)$. Using symmetry reduction, in the cases (i), (ii), (iii) above, it is possible to approximate these conic linear programs with semidefinite programs, that can be practically computed when the dimension $n$ is not too large. This step is explained in Section 6.5. It involves the description of the cones of $G$-invariant positive definite functions, for the groups $G = O_n(\mathbb{R})$ and $G = \mathrm{Stab}(x_0, O_n(\mathbb{R}))$. We provide this description, following the lines of Section 3, in Sections 6.3 and 6.4. Finally, in Section 6.6 we indicate further applications.

### 6.1 Lovász $\vartheta$

This number can be defined in many equivalent ways (see Lovász [69], Knuth [62]). We present here the most appropriate one in view of our purpose, which is the generalization to $S^{n-1}$.

**Definition 6.1** *The* theta number *of the graph $\Gamma = (V, E)$ with vertex set $V = \{1, 2, \ldots, n\}$ is defined by*

$$\vartheta(\Gamma) = \max \big\{ \langle X, J_n \rangle : X \succeq 0, \langle X, I_n \rangle = 1, X_{ij} = 0 \textit{ for all } (i,j) \in E \big\} \quad (20)$$

*where $I_n$ and $J_n$ denote respectively the identity and the all-one matrices of size $n$.*

The dual program gives another expression of $\vartheta(\Gamma)$ (there is no *duality gap* here because $X = I_n$ is a strictly feasible solution of (20) so the Slater condition is fulfilled):

$$\vartheta(\Gamma) = \min \big\{ t : X \succeq 0, X_{ii} = t - 1, X_{ij} = -1 \text{ for all } (i,j) \notin E \big\}. \quad (21)$$

We have already mentioned that this number provides an upper bound for the independence number $\alpha(\Gamma)$ of the graph $\Gamma$. It also provides a lower bound for the *chromatic number* $\chi(\overline{\Gamma})$ of the complementary graph $\overline{\Gamma}$ (the chromatic number of a graph is the minimal number of colors needed to color its vertices so that two connected vertices receive different colors); this is the content of the celebrated *Sandwich theorem* of Lovász [69]:

**Theorem 6.2**

$$\alpha(\Gamma) \le \vartheta(\Gamma) \le \chi(\overline{\Gamma}) \tag{22}$$

By modifying the definition of the theta number we get the strengthening $\alpha(\Gamma) \le \vartheta'(\Gamma) \le \vartheta(\Gamma)$, where

$$\begin{aligned}\vartheta'(\Gamma) = \max \big\{ \langle X, J_n \rangle : X \succeq 0,\ X \ge 0,\\ \langle X, I_n \rangle = 1,\ X_{ij} = 0 \text{ for all } (i,j) \in E \big\}.\end{aligned} \tag{23}$$

Again, this program has an equivalent dual form:

$$\begin{aligned}\vartheta'(\Gamma) = \min \big\{ t : X \succeq 0,\ X_{ii} \le t-1,\\ X_{ij} \le -1 \text{ for all } (i,j) \notin E \big\}.\end{aligned} \tag{24}$$

### 6.2 Generalizations of Lovász $\vartheta$ to the sphere

In order to obtain the wanted generalization, it is natural to replace in (23) the cone of positive semidefinite matrices indexed by the vertex set $V$ of the underlying graph, with the cone of continuous, positive definite functions on the sphere, defined in Section 3.1. In fact, there is a small difficulty here due to the fact that, in contrast with a finite dimensional Euclidean space, the space $\mathcal{C}(X)$ of continuous functions on an infinite compact space cannot be identified with its topological dual. Indeed, the *topological dual* $\mathcal{C}(X)^*$ of $\mathcal{C}(X)$ (i.e. the space of continuous linear forms on $\mathcal{C}(X)$) is the space $\mathcal{M}(X)$ of complex valued Borel regular measures on $X$ (see e.g. Rudin [83, Theorem 6.19], Conway [25, Appendix C]). Consequently, the two forms of $\vartheta'(\Gamma)$ given by (23) and (24) lead in the infinite case to two pairs of conic linear programs. As we shall see, the right choice between the two in order to obtain an appropriate bound for $A(S^{n-1}, \Omega, \lambda)$ depends on the set $\Omega$.

**Definition 6.3** *Let $\Omega^c = \{(x,y) : (x,y) \notin \Omega \text{ and } x \ne y\}$. Let*

$$\begin{aligned}\vartheta_1(S^{n-1}, \Omega) = \inf \big\{\ t : F \succeq 0,\ F(x,x) \le t-1,\\ F(x,y) \le -1 \text{ for all } (x,y) \in \Omega^c\ \big\},\end{aligned} \tag{25}$$

*and*

$$\begin{aligned}\vartheta_2(S^{n-1}, \Omega) = \sup \big\{\ \langle F, 1 \rangle : F \succeq 0,\ F \ge 0,\\ \langle F, \mathbf{1}_\Delta \rangle = 1,\\ F(x,y) = 0 \text{ for all } (x,y) \in \Omega\ \big\}.\end{aligned} \tag{26}$$

In the above programs, $F$ belongs to $\mathcal{C}((S^{n-1})^2)$. The notation $F \ge 0$ stands for "$F$ takes nonnegative values". The function taking the constant value 1 is denoted 1, so that $\langle F, 1 \rangle$ equals the integral of $F$ over $(S^{n-1})^2$. In contrast, with $\langle F, \mathbf{1}_\Delta \rangle$ we mean the integral of the one variable function $F(x,x)$ over $\Delta = \{(x,x) : x \in S^{n-1}\}$. Let us notice that, since $F$ is continuous in both programs, the sets $\Omega$ and $\Omega^c$ can be replaced by their topological closures $\overline{\Omega}$ and $\overline{\Omega^c}$.

A *positive semidefinite measure* on $(S^{n-1})^2$ is one that satisfies $\langle \lambda, f\rangle \geq 0$ for all $f \succeq 0$, where

$$\langle \lambda, f\rangle = \int_X f(x) d\lambda(x).$$

and this property is denoted $\lambda \succeq 0$. In a similar way a nonnegative measure $\lambda$ is denoted $\lambda \geq 0$.

**Theorem 6.4** *With the above notations and definitions, we have:*

1. *If $\overline{\Omega^c} \cap \Delta = \emptyset$, then the program dual to $\vartheta_1$ reads*

$$\begin{aligned}\vartheta_1^*(S^{n-1}, \Omega) = \sup \{\ \langle \lambda, 1\rangle : \ & \lambda \succeq 0,\ \ \lambda \geq 0,\\ & \lambda(\Delta) = 1,\\ & \mathrm{supp}(\lambda) \subset \overline{\Omega^c} \cup \Delta\ \}\end{aligned}$$

   *and*

$$A(S^{n-1}, \Omega, \mu_c) \leq \vartheta_1(S^{n-1}, \Omega) = \vartheta_1^*(S^{n-1}, \Omega). \tag{27}$$

2. *If $\overline{\Omega} \cap \Delta = \emptyset$, then the program dual to $\vartheta_2$ reads*

$$\begin{aligned}\vartheta_2^*(S^{n-1}, \Omega) = \inf \{\ t : \ & \lambda \succeq 0,\\ & \lambda \leq t\mu_\Delta - \mu^2 \ \textit{over}\ (S^{n-1})^2 \setminus \overline{\Omega}\ \},\end{aligned}$$

   *and*

$$A(S^{n-1}, \Omega, \mu) \leq \vartheta_2(S^{n-1}, \Omega) = \vartheta_2^*(S^{n-1}, \Omega). \tag{28}$$

*Proof.* The dual programs are computed in a standard way (see Duffin [35], Barvinok [15]). In order to prove that there is no duality gap, one can apply the criterion [15, Theorem 7.2].

For the inequality (27), let $C$ be a maximal spherical code avoiding $\Omega$. Then, the measure $\lambda = \delta_{C^2}/|C|$, where $\delta$ denotes the Dirac measure, is a feasible solution of $\vartheta_1^*$ with optimal value $|C| = A(S^{n-1}, \Omega, \mu_c)$.

For the inequality (28), let $(t, \lambda)$ be a feasible solution of $\vartheta_2^*$. Then, if $C$ avoids $\Omega$, $C^2 \subset (S^{n-1})^2 \setminus \overline{\Omega}$. Thus, $0 \leq \lambda(C^2) \leq t\mu(C) - \mu(C)^2$, leading to the wanted inequality $\mu(C) \leq t$. $\square$

**Example 6.5** *In the situations (i) and (ii) above, the set $\Omega$ fulfills the condition 1. of Theorem 6.4 while for (iii) we are in the case 2.*

### 6.3 Positive definite functions invariant under the full orthogonal group

It is a classical result of Schoenberg [84] that these functions, in the variables $(x, y) \in (S^{n-1})^2$, are exactly the nonnegative linear combinations of *Gegenbauer polynomials* (Section 3.4) evaluated at the inner product $x\cdot y$. We briefly review this classical result, following the lines of Section 3.

The space $\mathrm{Hom}_k^n$ of polynomials in $n$ variables $x_1, \ldots, x_n$ which are homogeneous of degree $k$ affords an action of the group $O_n(\mathbb{R})$ acting linearly on the $n$ variables. The Laplace operator $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ commutes with this action, thus the subspace $\mathrm{Harm}_k^n$ defined by:

$$\mathrm{Harm}_k^n = \{P \in \mathrm{Hom}_k^n : \Delta P = 0\}$$

is a representation of $O_n(\mathbb{R})$ which turns out to be irreducible (see e.g. Andrews, Askey, Roy [1]). Going from polynomials to polynomial functions on the sphere, we introduce $H_k^n$, the subspace of $\mathcal{C}(S^{n-1})$ arising from elements of $\mathrm{Harm}_k^n$. Then, for $d \geq k$, $H_k^n$ is a subspace of the space $V_d = \mathrm{Pol}_{\leq d}(S^{n-1})$ of polynomial functions on $S^{n-1}$ of degree up to $d$. We have:

**Theorem 6.6** *The irreducible $O_n(\mathbb{R})$-decomposition of $\mathrm{Pol}_{\leq d}(S^{n-1})$ is given by:*

$$\mathrm{Pol}_{\leq d}(S^{n-1}) = H_0^n \perp H_1^n \perp \cdots \perp H_d^n, \tag{29}$$

*where, for all $k \geq 0$, $H_k^n \simeq \mathrm{Harm}_k^n$ is of dimension $h_k^n = \binom{n+k-1}{k} - \binom{n+k-3}{k-2}$.*

*The function $Y_k(d_\theta(x, y))$ associated to $H_k^n$ following (14) equals:*

$$Y_k(d_\theta(x, y)) = h_k^n P_k^{((n-3)/2,(n-3)/2)}(x \cdot y). \tag{30}$$

*Proof.* For the sake of completeness we sketch a proof. Because $S^{n-1}$ is two-point homogeneous under $O_n(\mathbb{R})$, the algebra $(V_d)^{(2)}$ related to the finite dimensional $O_n(\mathbb{R})$-space $V_d$ is equal to the set of polynomial functions in the variable $t = x \cdot y$ of degree up to $d$, thus has dimension $d + 1$. On the other hand, it can be shown that the $d+1$ subspaces $H_k^n$, for $0 \leq k \leq d$, are non zero and pairwise distinct. Then, the statement (29) follows from the equality of the dimensions of $(V_d)^{(2)}$ and of $\mathrm{End}^{O_n(\mathbb{R})}(V_d)$ which are isomorphic algebras (12). It is worth to notice that we end up with a proof that the spaces $H_k^n$ are irreducible, without referring to the irreducibility of the spaces $\mathrm{Harm}_k^n$ (which can be proved in a similar way). The formula for $h_k^n$ follows by induction.

The functions $Y_k(d_{\theta(x,y)})$ associated to the decomposition (29) must be polynomials in the variable $x \cdot y$ of degree $k$; let us denote them temporary by $Q_k(t)$. A change of variables shows that, for a function $f(x) \in \mathcal{C}(S^{n-1})$ of the form $f(x) = \varphi(x \cdot y)$ for some $y \in S^{n-1}$,

$$\int_{S^{n-1}} f(x) d\mu(x) = c_n \int_{-1}^{1} \varphi(t)(1 - t^2)^{(n-3)/2} dt$$

for some constant $c_n$. Then, because the subspaces $H_k^n$ are pairwise orthogonal, the polynomials $Q_k$ must satisfy the orthogonality conditions

$$\int_{-1}^{1} Q_k(t) Q_l(t)(1 - t^2)^{(n-3)/2} dt = 0$$

for all $k \neq l$ thus they must be equal to the Gegenbauer polynomials up to a multiplicative factor. Integrating over $S^{n-1}$ the formula (9) when $x = y$ shows that $Y_k(0) = h_k^n$ thus computes this factor. □

**Corollary 6.7** *Let $P_k^n(t) = P_k^{((n-3)/2,(n-3)/2)}(t)$. Then, $F \in \mathcal{C}(S^{n-1})_{\succeq 0}^{O_n(\mathbb{R})}$ if and only if*

$$F(x,y) = \sum_{k \geq 0} f_k P_k^n(x \cdot y), \text{ with } f_k \geq 0 \text{ for all } k \geq 0. \tag{31}$$

### 6.4 Positive definite functions invariant under the stabilizer of one point

We fix an element $e \in S^{n-1}$ and let $G = \mathrm{Stab}(e, O_n(\mathbb{R}))$. Then, the orbit $O(x)$ of $x \in S^{n-1}$ under $G$ is the set

$$O(x) = \{y \in S^{n-1} : e \cdot y = e \cdot x\}.$$

The orbit $O((x,y))$ of $(x,y) \in S^{n-1}$ equals

$$O((x,y)) = \{(z,t) \in (S^{n-1})^2 : (e \cdot z, e \cdot t, z \cdot t) = (e \cdot x, e \cdot y, x \cdot y)\}.$$

In other words, the orbits of $G$ on $(S^{n-1})^2$ are parametrized by the triple of real numbers $(e \cdot x, e \cdot y, x \cdot y)$. So the $G$-invariant positive definite functions on $S^{n-1}$ are functions of the three variables $u = e \cdot x$, $v = e \cdot y$, $t = x \cdot y$. Their expression is computed in Bachoc, Vallentin [6].

**Theorem 6.8** *The irreducible decomposition of $V_d = \mathrm{Pol}_{\leq d}(S^{n-1})$ under the action of $G = \mathrm{Stab}(e, S^{n-1}) \simeq O_{n-1}(\mathbb{R})$ is given by:*

$$\mathrm{Pol}_{\leq d}(S^{n-1}) \simeq \bigoplus_{k=0}^{d} (\mathrm{Harm}_k^{n-1})^{d-k+1}. \tag{32}$$

*The coefficients of the matrix $Y_k = Y_k^n$ of size $d-k+1$ associated to $\mathrm{Harm}_k^{n-1}$ are equal to:*

$$Y_{k,ij}^n(u,v,t) = u^i v^j Q_k^{n-1}(u,v,t) \tag{33}$$

*where $0 \leq i, j \leq d-k$, and*

$$Q_k^{n-1}(u,v,t) = \big((1-u^2)(1-v^2)\big)^{k/2} P_k^{n-1}\Big(\frac{t-uv}{\sqrt{(1-u^2)(1-v^2)}}\Big)$$

*Proof.* We refer to Bachoc, Vallentin [6] for the details. In order to obtain the $G$-decomposition of the space $\mathrm{Pol}_{\leq d}(S^{n-1})$ we can start from the $O_n(\mathbb{R})$-decomposition (29). The $O_n(\mathbb{R})$-irreducible subspaces $H_k^n$ split into smaller subspaces following the $O_{n-1}(\mathbb{R})$-isomorphisms (see e.g. Vilenkin, Klimyk [98]):

$$\mathrm{Harm}_k^n \simeq \bigoplus_{i=0}^{k} \mathrm{Harm}_i^{n-1} \tag{34}$$

leading to (32). The expression (33) follows from the definition (9) of $E_{k,ij}(x,y)$ and from the construction of convenient basis $(e_{k,i,s})_s$. These basis arise from explicit embeddings of $H_k^{n-1}$ into $V_d$, defined as follows. For $x \in S^{n-1}$, let $x = ue + \sqrt{1-u^2}\zeta$ where $\zeta \in (\mathbb{R}e)^\perp$ and $\zeta \cdot \zeta = 1$. Let $\varphi_{k,i}$ send $f \in H_k^{n-1}$ to $\varphi_{k,i}(f)$ where

$$\varphi_{k,i}(f)(x) = u^i (1-u^2)^{k/2} f(\zeta).$$

Then, $\{e_{k,i,s} : 1 \leq s \leq h_k^{n-1}\}$ is taken to be the image by $\varphi_{k,i}$ of an orthonormal basis of $H_k^{n-1}$. $\square$

### 6.5 Reduction using symmetries

If the set $\Omega$ is invariant under a subgroup $G$ of $O_n(\mathbb{R})$, then the feasible sets of the conic linear programs (25) and (26) can be restricted to the $G$-invariant positive definite functions $F$. Indeed, symmetry reduction of finite dimensional semidefinite programs extends to infinite compact spaces, with the now familiar trick that replaces the finite average over the elements of a finite group by the integral for the Haar measure of the group. We apply this general principle to the programs (25) and (26) and we focus on the cases (i)–(iii) above.

**Case (i): $\boldsymbol{\Omega = \{(x,y) : d_\theta(x,y) \in ]0, \theta_{\min}[\}}$.**

The set $\Omega$ is invariant under $O_n(\mathbb{R})$. According to Theorem 6.4 we consider the program (25) where $F \succeq 0$ can be replaced by (31). The program $\vartheta_1$ becomes after a few simplifications:

$$\begin{aligned}\vartheta_1(S^{n-1}, \Omega) = \inf \Big\{ \ 1 + \sum_{k\geq 1} f_k : f_k \geq 0, \\ 1 + \sum_{k \geq 1} f_k P_k^n(t) \leq 0 \quad t \in [-1, s] \Big\},\end{aligned} \tag{35}$$

where $s = \cos(\theta_{\min})$. One recognises in (35) the so-called *Delsarte linear programming bound* for the maximal number of elements of a spherical code with minimal angular distance $\theta_{\min}$, see Delsarte, Goethals, Seidel [33], Kabatiansky, Levenshtein [51], Conway, Sloane [26, Chapter 9]. The optimal value of the linear program (35) is not known in general, although explicit feasible solutions leading to very good bounds and also to asymptotic bounds have been constructed (Kabatiansky, Levenshtein [51], Levenshtein [67], Odlyzko, Sloane [77]). Moreover, this infinite dimensional linear program can be efficiently approximated by semidefinite programs defined in the following way: in order to deal with only a finite number of variables $f_k$, one restricts to $k \leq d$ (it amounts to restrict to $G$-invariant positive definite functions of $V_d^{(2)}$). Then the polynomial $1 + \sum_{k=1}^d f_k P_k^n$ is required in (35) to be non-positive over a certain interval of real numbers. By the theorem of Lukács concerning non-negative polynomials (see e.g. Szegö [91, Chapter 1.21], this can be expressed as a sums of square condition, hence as a semidefinite program. Then, when $d$ varies, one obtains a sequence of semidefinite programs approaching $\vartheta_1(S^{n-1}, \Omega)$ from above.

### Case (ii): $\boldsymbol{\Omega = \{(x,y) : d_\theta(x,y) = \theta\}}$.

The set $\Omega$ is again invariant under $O_n(\mathbb{R})$ but now we deal with (26), which becomes:

$$\vartheta_2(S^{n-1}, \Omega) = \sup\Big\{ \ f_0 : \ f_k \geq 0, \sum_{k\geq 0} f_k = 1, \sum_{k\geq 0} f_k P_k^n(s) = 0\Big\} \tag{36}$$

where $s = \cos(\theta)$. This linear program has infinitely many variables but only two constraints. Its optimal value turns to be easy to determine (we refer to Bachoc et. al. [11] for a proof).

**Theorem 6.9** *Let $m(s)$ be the minimum of $P_k^n(s)$ for $k = 0, 1, 2, \ldots$. Then*

$$\vartheta_2(S^{n-1}, \Omega) = \frac{m(s)}{m(s) - 1}.$$

### Case (iii): $\boldsymbol{\Omega = \{(x,y) : d_\theta(x,y) \in ]0, \theta_{\min}[}$ or $\boldsymbol{(x,y) \notin \mathrm{Cap}(e,\phi)^2\}}$.

This set is invariant by the smaller group $G = \mathrm{Stab}(e, O_n(\mathbb{R}))$. Like in case (i), the program $\vartheta_1$ must be considered and in this program $F$ can be assumed to be $G$-invariant.

From Stone-Weierstrass theorem (see e.g. Conway [25, Theorem 8.1]), the elements of $\mathcal{C}((S^{n-1})^2)$ can be uniformly approximated by those of $V_d^{(2)}$. In addition, one can prove that the elements of $\mathcal{C}((S^{n-1})^2)_{\succeq 0}$ can be uniformly approximated by positive definite functions belonging to $V_d^{(2)}$. We introduce:

$$\begin{aligned}\vartheta_1^{(d)}(S^{n-1}, \Omega) = \inf\Big\{ \ t : \ & F \in \big(V_d^{(2)}\big)_{\succeq 0}, \\ & F(x,x) \leq t - 1, \\ & F(x,y) \leq -1 \text{ for all } (x,y) \in \Omega^c \ \Big\}.\end{aligned} \tag{37}$$

So, $\vartheta_1(S^{n-1}, \Omega) \leq \vartheta_1^{(d)}(S^{n-1}, \Omega)$, and the limiting value of $\vartheta_1^{(d)}(S^{n-1}, \Omega)$ when $d$ goes to infinity equals $\vartheta_1(S^{n-1}, \Omega)$. Since $\Omega$ and $V_d$ are invariant by $G$, one can moreover assume that $F$ is $G$-invariant. From Theorems 3.3 and 6.8, we have an expression for $F$:

$$F(x,y) = \sum_{k=0}^{d} \langle F_k, Y_k(u,v,t)\rangle, \quad F_k \succeq 0,$$

where the matrices $F_k$ are of size $d-k+1$. Replacing in $\vartheta_1^{(d)}$ leads to:

$$\begin{aligned}\vartheta_1^{(d)}(S^{n-1}, \Omega) = \inf \big\{ t : \ & F_k \succeq 0, \\ & \sum_{k=0}^{d} \langle F_k, Y_k(u,u,1)\rangle \leq t \qquad u \in [s', 1], \\ & \sum_{k=0}^{d} \langle F_k, Y_k(u,v,t))\rangle \leq -1 \quad s' \leq u \leq v \leq 1 \\ & \qquad\qquad\qquad\qquad\qquad -1 \leq t \leq s \big\}\end{aligned}$$

with $s = \cos(\theta)$ and $s' = \cos(\phi)$. The left hand sides of the inequalities are polynomials in three variables. Again, these constraints can be relaxed using sums of squares in order to boil down to true semidefinite programs. We refer to Bachoc, Vallentin [7] for the details, and for numerical computations of upper bounds for codes in spherical caps with given minimal angular distance.

### 6.6 Further applications

In Bachoc, Vallentin [6] it is shown how Delsarte linear programming bound [29] can be improved with semidefinite constraints arising from the matrices $Y_k^n$ (33). The idea is very much the same as for the Hamming space given in Schrijver [86] and explained in Section 4 : instead of considering constraints on pairs of points only, one exploits constraints on triples of points. More precisely, if $S_k^n(u,v,t)$ denotes the symmetrization of $Y_k^n(u,v,t)$ in the variables $(u,v,t)$, then the following semidefinite property holds for all spherical code $C$:

$$\sum_{(x,y,z)\in C^3} S_k^n(x\cdot y, y\cdot z, z\cdot x) \succeq 0. \tag{38}$$

From (38), it is possible to define a semidefinite program whose optimal value upper bounds the number of elements of a code with given minimal angular distance. In Bachoc, Vallentin [6], Mittelmann, Vallentin [73], new upper bounds for the kissing number have been obtained for the dimensions $n \leq 24$ with this method. We give next a simplified version of the semidefinite program used in [6]. Another useful version is given in Bachoc, Vallentin [9] for proving that the maximal angular distance of 10 points on $S^3$ is $\cos(1/6)$.

**Theorem 6.10** *The optimal value of the semidefinite program:*

$$\begin{aligned}\inf \Big\{ \ 1 + \langle F_0, J_{d+1}\rangle : \ & F_k \succeq 0 \\ & \sum_{k=0}^{d} \langle F_k, S_k^n(u,u,1)\rangle \leq -\frac{1}{3}, \quad -1 \leq u \leq s \\ & \sum_{k=0}^{d} \langle F_k, S_k^n(u,v,t)\rangle \leq 0, \qquad -1 \leq u,v,t \leq s \ \Big\}\end{aligned} \tag{39}$$

*is an upper bound for the number* $A(S^{n-1}, \Omega, \mu_c)$ *where* $\Omega = \{(x,y) \in (S^{n-1})^2 : s < x \cdot y < 1\}$*, i.e. for the maximal number of elements of a spherical code with minimal angular distance at least equal to* $\theta_{\min} = \arccos(s)$*.*

*Proof.* Let $(F_0, \ldots, F_k)$ a feasible solution of (39). Let

$$F(x,y,z) = \sum_{k=0}^{d} \langle F_k, S_k^n(x \cdot y, y \cdot z, z \cdot x)\rangle.$$

If $C$ is a spherical code, we consider $\Sigma = \sum_{(x,y,z) \in C^3} F(x,y,z)$. We have:

$$0 \le \Sigma = \sum_{x \in C} F(x,x,x) + \sum_{|\{x,y,z\}|=2} F(x,x,y) + \sum_{|\{x,y,z\}|=3} F(x,y,z)$$

where the inequality holds because of (38). Then, taking $S_0^n(1,1,1) = J_{d+1}$ and $S_k^n(1,1,1) = 0$ for $k \ge 1$ into account, we have $F(x,x,x) = \langle F_0, J_{d+1}\rangle$. If moreover $d_\theta(C) \ge \theta_{\min}$, we can apply the constraint inequalities of the program to the second and third terms of the right hand side. We obtain:

$$0 \le \Sigma \le \langle F_0, J_{d+1}\rangle |C| - |C|(|C|-1)$$

leading to the inequality $|C| \le 1 + \langle F_0, J_{d+1}\rangle$. □

## 7 Sums of squares

A fundamental task in polynomial optimization and in real algebraic geometry is to decide and certify whether a polynomial with real coefficients in $n$ indeterminates can be written as a *sum of squares*: *Given* $p \in \mathbb{R}[x_1, \ldots, x_n]$ *do there exist polynomials* $q_1, \ldots, q_m \in \mathbb{R}[x_1, \ldots, x_n]$ *so that*

$$p = q_1^2 + q_2^2 + \cdots + q_m^2 \ ?$$

This problem can be reformulated as a semidefinite feasibility problem: Let $z$ be a vector containing a basis of $\mathbb{R}[x_1, \ldots, x_n]_{\le d}$ the space of polynomials of degree at most $d$. For example, let $z$ be the vector containing the monomial basis

$$z = (1, x_1, x_2, \ldots, x_n, x_1^2, x_1 x_2, x_2^2, \ldots, x_n^d)$$

which has length $\binom{n+d}{d}$. A polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ of degree $2d$ is a sum of square if and only if there is a positive semidefinite matrix $X$ of size $\binom{n+d}{d} \times \binom{n+d}{d}$ so that the $\binom{n+2d}{2d}$ linear — linear in the entries of $X$ — equations

$$p(x_1, \ldots, x_n) = z^T X z$$

hold.

This semidefinite feasibility problem can be simplified if the polynomial $p$ has symmetries. The method has been worked out by Gatermann and Parrilo in [38]. In this section we give the main ideas of the method. For details and further we refer to the original article.

### 7.1 Basics from invariant theory

We start by explaining what we mean that a polynomial has symmetries. Again to simplify the presentation of the theory we consider the complex case only.

Let $G$ be a finite group acting on $\mathbb{C}^n$. This group action induces a group action on the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$ by

$$(gp)(x_1, \ldots, x_n) = p(g^{-1}(x_1, \ldots, x_n)),$$

and we say that a polynomial $p$ is $G$-invariant if $gp = p$ for all $g \in G$. The set $\mathbb{C}[x_1, \ldots, x_n]^G$ of all $G$-invariant polynomials is a ring, the *invariant ring*. By Hilbert's finiteness theorem it is generated by finitely many $G$-invariant polynomials. Even more is true: Since the invariant ring has the Cohen-Macaulay property it admits a Hironaka decomposition: There are $G$-invariant polynomials $\eta_i$, $\theta_j \in \mathbb{C}[x_1, \ldots, x_n]$ so that

$$\mathbb{C}[x_1, \ldots, x_n]^G = \bigoplus_{i=1}^{r} \eta_i \mathbb{C}[\theta_1, \ldots, \theta_s], \tag{40}$$

hence, every invariant polynomial can be *uniquely* written as a polynomial in the polynomials $\eta_i$ and $\theta_j$ where $\eta_i$ only occurs linearly. We refer to Sturmfels [90, Chapter 2.3] for the definitions and proofs; we only need the existence of a Hironaka decomposition here.

### 7.2 Sums of squares with symmetries

We consider the action of the finite group $G$ restricted to the $\binom{n+d}{d}$-dimensional vector space of complex polynomials of degree at most $d$. This defines a unitary representation

$$\pi : G \to \mathrm{Gl}(\mathbb{C}[x_1, \ldots, x_n]_{\leq d}).$$

From now on, by using the monomial basis, we see $\pi(g)$ as a regular matrix in $\mathbb{C}^{\binom{n+d}{d} \times \binom{n+d}{d}}$.

**Example 7.1** *For instance, the matrix $g^{-1} = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ acts on $\mathbb{C}^2$ and so on the polynomial $p = 1 + x_1 + x_2 + x_1^2 + x_1x_2 + x_2^2 \in \mathbb{C}[x_1, x_2]_{\leq 2}$ by*

$$\begin{aligned}
(gp)(x_1, x_2) &= p(x_1 + 2x_2, 3x_1 + 4x_2) \\
&= 1 + (x_1 + 2x_2) + (3x_1 + 4x_2) + (x_1 + 2x_2)^2 \\
&\quad + (x_1 + 2x_2)(3x_1 + 4x_2) + (3x_1 + 4x_2)^2 \\
&= 1 + (x_1 + 2x_2) + (3x_1 + 4x_2) + (x_1^2 + 4x_1x_2 + 4x_2^2) \\
&\quad + (3x_1^2 + 10x_1x_2 + 8x_2^2) + (9x_1^2 + 24x_1x_2 + 16x_2^2).
\end{aligned}$$

*and so defines the matrix*

$$\pi(g) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 & 9 \\ 0 & 0 & 0 & 4 & 10 & 24 \\ 0 & 0 & 0 & 4 & 8 & 16 \end{pmatrix}.$$

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial which is a sum of squares and which is $G$-invariant. Thus we have a positive semidefinite matrix $X \in \mathbb{R}^{\binom{n+d}{d} \times \binom{n+d}{d}}$ so that

$$p(x_1, \ldots, x_n) = z^{\mathsf{T}} X z = z^* X z,$$

and for every $g \in G$ we have

$$gp(x_1, \ldots, x_n) = (\pi(g)^* z)^* X (\pi(g)^* z) = z^* \pi(g) X \pi(g)^* z.$$

Hence, $X$ is $G$-invariant and lies in $\left(\mathbb{C}^{\binom{n+d}{d} \times \binom{n+d}{d}}\right)^G$, the commutant algebra of the matrix $*$-algebra spanned by the matrices $\pi(g)$ with $g \in G$. So by Theorem 2.7 there are numbers $D$, $m_1, \ldots, m_D$ and a $*$-isomorphism

$$\varphi : \left(\mathbb{C}^{\binom{n+d}{d} \times \binom{n+d}{d}}\right)^G \to \bigoplus_{k=1}^{D} \mathbb{C}^{m_k \times m_k}.$$

Hence, cf. Step 2 (second version) in Section 1.2, we can write the polynomial $p$ in the form

$$p(x_1, \ldots, x_n) = z^* \left( \sum_{k=1}^{D} \sum_{u,v=1}^{m_k} x_{k,uv} \varphi^{-1}(E_{k,uv}) \right) z$$

with $D$ positive semidefinite matrices

$$X_k = (x_{k,uv})_{1 \le u,v \le m_k}, \quad k = 1, \ldots, D.$$

We define $D$ matrices $E_1, \ldots, E_D \in (\mathbb{C}[x_1, \ldots, x_n]^G)^{m_k \times m_k}$ with $G$-invariant polynomial entries by

$$(E_k)_{uv} = \left(\varphi^{-1}(E_{k,uv})\right)_{\prod_i x_i^{\alpha_i}, \prod_i x_i^{\beta_i}} \prod_i x_i^{\alpha_i + \beta_i},$$

where we consider matrices in $\mathbb{C}^{\binom{n+d}{d} \times \binom{n+d}{d}}$ as matrices whose rows and columns are indexed by monomials $\prod_i x_i^{\alpha_i}$. Then, the polynomial $p$ has a representation of the form

$$p(x_1, \ldots, x_n) = \sum_{k=1}^{D} \langle X_k, E_k \rangle.$$

Since the entries of $E_k$ are $G$-invariant polynomials we can use a Hironaka decomposition to represent them in terms of the invariants $\eta_i$ and $\theta_j$. We summarize our discussion in the following theorem.

**Theorem 7.2** *Let $p$ be a $G$-invariant polynomial of degree $2d$ which is a sum of squares. Then there are numbers $D$, $m_1, \ldots, m_D$ so that $p$ has a representation of the form*

$$p(x_1, \ldots, x_n) = \sum_{k=1}^{D} \langle X_k, E_k \rangle,$$

*where $X_k \in \mathbb{C}^{m_k \times m_k}$ are positive semidefinite Hermitian matrices, and where*

$$E_k \in \left( \bigoplus_{i=1}^{r} \eta_i \mathbb{C}[\theta_1, \ldots, \theta_s] \right)^{m_k \times m_k}$$

*are matrices whose entries are polynomials (determined by a Hironaka decomposition of $\mathbb{C}[x_1, \ldots, x_n]^G$ and by the $*$-isomorphism $\varphi$).*

## 8 More applications

In the last years many results were obtained for semidefinite programs which are symmetric. This was done for a variety of problems and applications. In this final section we want to give a brief, and definitely not complete, guide to the extensive and growing literature.

### 8.1 Interior point algorithms

Kanno, Ohsaki, Murota, Katoh [52] consider structural properties of search directions in primal-dual interior-point methods for solving invariant semidefinite programs and apply this to truss optimization problems. de Klerk, Pasechnik [57] show how the dual scaling method can be implemented to exploit the particular data structure where the data matrices come from a low-dimensional matrix algebra.

### 8.2 Combinatorial optimization

Using symmetry in semidefinite programs has been used in combinatorial optimization for a variety of problems: quadratic assignment problem (de Klerk, Sotirov [61]), travelling salesman problem (de Klerk, Pasechnik, Sotirov [60]), graph coloring (Gvozdenović, Laurent [45], [46], Gvozdenović [44]), Lovász theta number (de Klerk, Newman, Pasechnik, Sotirov [56]).

### 8.3 Polynomial optimization

Jansson, Lasserre, Riener, Theobald [50] work out how constrained polynomial optimization problems behave which are invariant under the action of the symmetric group or the cyclic group. Among many other things, Cimprič, Kuhlmann, Scheiderer [24] extend the discussion of Gatermann, Parrilo [38] from finite groups to compact groups. Cimprič [23] transfers the method to compute minima of the spectra of differential operators. In [18] Bosse constructs symmetric polynomials which are non-negative but not sums of squares.

### 8.4 Low distortion geometric embedding problems

Linial, Magen, Naor [68] give lower bounds for low distortion embedding of graphs into Euclidean space depending on the girth. Vallentin [95] finds explicit optimal low distortion embeddings for several families of distance regular graphs. Both papers construct feasible solutions of semidefinite programs by symmetry reduction and by using the theory of orthogonal polynomials.

### 8.5 Miscellaneous

Bai, de Klerk, Pasechnik, Sotirov [12] exploit symmetry in truss topology optimization and Boyd, Diaconis, Parrilo, Xiao [19] in the analysis of fast mixing Markov chains on graphs.

### 8.6 Software

Pasechnik, Kini [80] develop a software package for the computer algebra system `GAP` for computing with the regular $*$-representation for matrix $*$-algebras coming from coherent configurations.

### 8.7 Surveys and lecture notes

Several surveys and lecture notes on symmetry in semidefinite programs with different aims were written in the last years. The lecture notes of Bachoc [5] especially discuss applications in coding theory and extend those of Vallentin [97] which focuses on aspects from harmonic analysis. The survey [54] of de Klerk discusses next to symmetry also the exploitation of other structural properties of semidefinite programs like low rank or sparsity.

## Acknowledgements

We thank the referee for the helpful suggestions and comments. The fourth author was supported by Vidi grant 639.032.917 from the Netherlands Organization for Scientific Research (NWO).

## References


1. G.E. Andrews, R. Askey, R. Roy, *Special functions*, Cambridge University Press, Cambridge, 1999.
2. R. Askey, *Orthogonal polynomials and special functions*, Society for Industrial and Applied Mathematics, Philadelphia, 1975.
3. J. Astola, *The Lee-scheme and bounds for Lee-codes*, Cybernet. Systems **13** (1982) 331–343.
4. C. Bachoc, *Linear programming bounds for codes in Grassmannian spaces*, IEEE Trans. Inf. Th. **52** (2006), 2111–2125.
5. C. Bachoc, *Semidefinite programming, harmonic analysis and coding theory*, arXiv:0909.4767 [cs.IT]
6. C. Bachoc, F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. **21** (2008), 909–924.
7. C. Bachoc, F. Vallentin, *Semidefinite programming, multivariate orthogonal polynomials, and codes in spherical caps*, special issue in the honor of Eichii Bannai, Europ. J. Comb. **30** (2009), 625–637.
8. C. Bachoc, F. Vallentin, *More semidefinite programming bounds (extended abstract)*, pages 129–132 in Proceedings "DMHF 2007: COE Conference on the Development of Dynamic Mathematics with High Functionality", October 2007, Fukuoka, Japan.
9. C. Bachoc, F. Vallentin, *Optimality and uniqueness of the (4,10,1/6) spherical code*, J. Comb. Theory Ser. A **116** (2009), 195–204.
10. C. Bachoc, G. Zémor, *Bounds for binary codes relative to pseudo-distances of k points*, Adv. Math. Commun. **4** (2010), 547–565.
11. C. Bachoc, G. Nebe, F.M. de Oliveira Filho, F. Vallentin, *Lower bounds for measurable chromatic numbers*, Geom. Funct. Anal. **19** (2009), 645–661.
12. Y. Bai, E. de Klerk, D.V. Pasechnik, R. Sotirov, *Exploiting group symmetry in truss topology optimization*, Optimization and Engineering **10** (2009), 331–349.
13. E. Bannai, T. Ito, *Algebraic combinatorics. I.*, The Benjamin/Cummings Publishing Co. Inc., Menlo Park, CA, 1984.
14. A. Barg, P. Purkayastha, *Bounds on ordered codes and orthogonal arrays*, Moscow Math. Journal **9**, 2009, 211–243.
15. A. Barvinok, *A course in convexity*, Graduate Studies in Mathematics **54**, American Mathematical Society, 2002.
16. M. Berger, *A Panoramic View of Riemannian Geometry*, Springer-Verlag, 2003.
17. S. Bochner, *Hilbert distances and positive definite functions*, Ann. of Math. **42** (1941), 647–656.
18. H. Bosse, *Symmetric, positive polynomials, which are not sums of squares*, Series: CWI. Probability, Networks and Algorithms [PNA], Nr. E0706, 2007.
19. S. Boyd, P. Diaconis, P.A. Parrilo, L. Xiao, *Symmetry analysis of reversible Markov chains*, Internet Mathematics **2** (2005).
20. A.E. Brouwer, A.M. Cohen and A. Neumaier, *Distance-regular graphs*, Springer-Verlag, Berlin, 1989.
21. D. Bump, *Lie Groups*, Graduate Text in Mathematics **225**, Springer-Verlag, 2004.
22. P. J. Cameron, *Coherent configurations, association schemes and permutation groups*, pages 55–71 in *Groups, combinatorics & geometry (Durham, 2001)*, World Sci. Publishing, River Edge, NJ, 2003.

23. J. Cimprič, *Estimating lowest eigenvalues of symmetric polynomial differential operators by semidefinite programming*, to appear in J. Math. Anal. Appl.
24. J. Cimprič, S. Kuhlmann, C. Scheiderer, *Sums of squares and moment problems in equivariant situations*, Trans. Amer. Math. Soc. **361** (2009), 735–765.
25. J.B. Conway, *A course in functional analysis*, Graduate Text in Mathematics **96**, Springer-Verlag, 2007.
26. J.H. Conway, N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, third edition, Springer-Verlag, New York, 1999.
27. J. Creignou, H. Diet, *Linear programming bounds for unitary codes*, Adv. Math. Commun. **4** (2010), 323–344.
28. K.R. Davidson, *C*-Algebras by Example*.
29. P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. (1973), vi+97.
30. P. Delsarte, J. M. Goethals, *Alternating bilinear forms over* $GF(q)$, J. Comb. Th. A **19** (1975), 26–50.
31. P. Delsarte, *Hahn polynomials, discrete harmonics and* $t$*-designs*, SIAM J. Appl. Math. **34** (1978), 157–166.
32. P. Delsarte, *Bilinear forms over a finite field, with applications to coding theory*, J. Comb. Th. A **25** (1978), 226-241.
33. P. Delsarte, J.M. Goethals, J.J. Seidel, *Spherical codes and designs*, Geom. Dedicata **6** (1977), 363–388.
34. E. de Klerk, C. Dobre, D.V. Pasechnik, *Numerical block diagonalization of matrix *-algebras with application to semidefinite programming*, preprint, 2009.
35. R.J. Duffin, *Infinite Programs*, in: Linear inequalities and related systems, (H.W. Kuhn, A.W. Tucker eds.), Princeton Univ. Press, 1956, 157–170.
36. C.F. Dunkl, *A Krawtchouk polynomial addition theorem and wreath product of symmetric groups*, Indiana Univ. Math. J. **25** (1976), 335–358.
37. T. Ericson, V. Zinoviev, *Codes on Euclidean spheres*, North-Holland, 2001.
38. K. Gatermann, P.A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Alg. **192** (2004), 95–128.
39. D.C. Gijswijt, *Matrix Algebras and Semidefinite Programming Techniques for Codes*, PhD thesis, University of Amsterdam, 2005.
40. D.C. Gijswijt, *Block diagonalization for algebra's associated with block codes*, arXiv:0910.4515 [math.OC]
41. D.C. Gijswijt, H.D. Mittelmann, A. Schrijver, *Semidefinite code bounds based on quadruple distances*, arXiv.math:1005.4959 [math.CO]
42. D.C. Gijswijt, A. Schrijver, H. Tanaka, *New upper bounds for nonbinary codes based on the Terwilliger algebra and semidefinite programming*, J. Comb. Theory Ser. A **113** (2006) 1719–1731.
43. M.X. Goemans, D.P. Williamson, *Approximation algorithms for MAX-3-CUT and other problems via complex semidefinite programming*, J. Comput. System Sci. **68** (2004) 442–470.
44. N. Gvozdenović, *Approximating the stability number and the chromatic number of a graph via semidefinite programming*, PhD Thesis, University of Amsterdam, 2008.
45. N. Gvozdenović, M. Laurent, *The operator* $\Psi$ *for the chromatic number of a graph*, SIAM J. Optim. **19** (2008), 572–591.
46. N. Gvozdenović, M. Laurent, *Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization*, SIAM J. Optim. **19** (2008), 592-615.

47. N. Gvozdenović, M. Laurent, F. Vallentin, *Block-diagonal semidefinite programming hierarchies for 0/1 programming*, Oper. Res. Lett. **37** (2009), 27–31.
48. R. A. Horn, C. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1990.
49. A.T. James, A.G. Constantine, *Generalized Jacobi polynomials as spherical functions of the Grassmann manifold*, Proc. London Math. Soc. **29** (1974), 174–192.
50. L. Jansson, J.B. Lasserre, C. Riener, T. Theobald, *Exploiting symmetries in SDP-relaxations for polynomial optimization*, Optimization Online, September 2006.
51. G.A. Kabatiansky, V.I. Levenshtein, *Bounds for packings on a sphere and in space*, Problems of Information Transmission **14** (1978), 1–17.
52. Y. Kanno, M. Ohsaki, K. Murota, N. Katoh, *Group symmetry in interior-point methods for semidefinite program*, Optimization and Engineering **2** (2001) 293–320.
53. D.J. Kleitman, *The crossing number of* $K_{5,n}$, J. Comb. Theory Ser. B **9** (1970) 315–323.
54. E. de Klerk, *Exploiting special structure in semidefinite programming: A survey of theory and applications*, European Journal of Operational Research **201** (2010), 1–10.
55. E. de Klerk, J. Maharry, D.V. Pasechnik, R.B. Richter, G. Salazar, *Improved bounds for the crossing numbers of* $K_{m,n}$ *and* $K_n$, SIAM J. Disc. Math. **20** (2006) 189–202.
56. E. de Klerk, M.W. Newman, D.V. Pasechnik, R. Sotirov, *On the Lovasz* $\theta$*-number of almost regular graphs with application to Erdős-Rényi graphs*, European J. Combin. **31** (2009), 879–888.
57. E. de Klerk, D.V. Pasechnik, *Solving SDP's in non-commutative algebras part I: the dual-scaling algorithm*, Discussion paper from Tilburg University, Center for economic research, 2005.
58. E. de Klerk, D.V. Pasechnik, *A note on the stability number of an orthogonality graph*, European J. Combin. **28** (2007) 1971–1979.
59. E. de Klerk, D.V. Pasechnik, A. Schrijver, *Reduction of symmetric semidefinite programs using the regular* $*$*-representation*, Math. Program., Ser. B **109** (2007), 613–624.
60. E. de Klerk, D.V. Pasechnik, R. Sotirov, *On Semidefinite Programming Relaxations of the Travelling Salesman Problem*, Discussion paper from Tilburg University, Center for economic research, 2007.
61. E. de Klerk, R. Sotirov, *Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem*, Math. Program. Ser. A **122** (2010), 225–246.
62. D.E. Knuth, *The sandwich theorem*, Electron. J. Combin. **1** (1994), 48 pp.
63. T.Y. Lam, *A first course in noncommutative rings*, Springer-Verlag, 1991.
64. J.B Lasserre, *An explicit exact SDP relaxation for nonlinear 0/1 programs.*, pages 293–302 in K. Aardal and A.M.H. Gerards, eds., *Lecture Notes in Computer Science* **2081**, 2001.
65. M. Laurent, *Strengthened semidefinite programming bounds for codes*, Math. Program. Ser. B **109** (2007) 239–261.
66. M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, pages 157–270 in *Emerging Applications of Algebraic Geometry*, Vol.

149 of IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivant (eds.), Springer-Verlag, 2009

67. V.I. Levenshtein, *Universal bounds for codes and designs*, in *Handbook of Coding Theory*, eds V. Pless and W. C. Huffmann, Amsterdam: Elsevier, 1998, 499-648.
68. N. Linial, A. Magen, A. Naor, *Girth and Euclidean Distortion*, Geom. Funct. Anal. **12** (2002), 380–394.
69. L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory **25** (1979), 1–5.
70. F.J. MacWilliams, N.J.A. Sloane, *The theory of error-correcting codes*, North-Holland Mathematical Library, Vol. 16. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977.
71. W.J. Martin, D.R. Stinson, *Association schemes for ordered orthogonal arrays and* $(T, M, S)$*-nets*, Canad. J. Math. **51** (1999), 326–346.
72. R.J. McEliece, E.R. Rodemich and H.C. Rumsey, Jr, *The Lovász bound and some generalizations*, J. Combinatorics, Inform. Syst. Sci. **3** (1978), 134–152.
73. H.D. Mittelmann, F. Vallentin, *High accuracy semidefinite programming bounds for kissing numbers*, Experiment. Math. **19** (2010), 174-178.
74. K. Murota, Y. Kanno, M. Kojima, S. Kojima, *A numerical algorithm for block-diagonal decomposition of matrix *-algebras, Part I: proposed approach and application to semidefinite programming*, to appear in Japan Journal of Industrial and Applied Mathematics.
75. O.R. Musin, *Multivariate positive definite functions on spheres*, arXiv:math/0701083 [math.MG]
76. A. Nemirovski, *Advances in convex optimization: Conic programming*, pages 413–444 in: *Proceedings of International Congress of Mathematicians, Madrid, August 22-30, 2006, Volume 1* (M. Sanz-Sol, J. Soria, J.L. Varona, J. Verdera, Eds.), European Mathematical Society Publishing House, 2007.
77. A.M. Odlyzko, N.J.A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Comb. Theory Ser. A **26** (1979), 210–214.
78. F.M. de Oliveira Filho, *New Bounds for Geometric Packing and Coloring via Harmonic Analysis and Optimization*, PhD thesis, University of Amsterdam, 2009.
79. F.M. de Oliveira Filho, F. Vallentin, *Fourier analysis, linear programming, and densities of distance avoiding sets in* $\mathbb{R}^n$, J. Eur. Math. Soc. (JEMS) **12** (2010), 1417–1428.
80. D.V. Pasechnik, K. Kini, *A GAP package for computation with coherent configurations*, pages 69–72 in: *Mathematical Software — ICMS 2010* (K. Fukuda, J. van der Hoeven, M. Joswig, N. Takayama, Eds.) LNCS 6327, 2010.
81. A. Roy, A. J. Scott, *Unitary designs and codes*, Des. Codes Cryptogr. **53** (2009), 13–31.
82. A. Roy, *Bounds for codes and designs in complex subspaces*, arXiv:0806.2317 [math.CO]
83. W. Rudin, *Real and Complex Analysis*, McGraw-Hill International Editions, 1987.
84. I.J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108.

85. A. Schrijver, *Association schemes and the Shannon capacity: Eberlein polynomials and the Erdős-Ko-Rado theorem*, Algebraic methods in graph theory Vol. I, II (Szeged, 1978), pp. 671–688, *Colloq. Math. Soc. János Bolyai* 25, North-Holland, Amsterdam-New York, 1981.
86. A. Schrijver, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory **51** (2005), 2859–2866.
87. A. Soifer, *The mathematical coloring book*, Springer-Verlag, 2008.
88. D. Stanton, *Orthogonal polynomials and Chevalley groups*, in Special functions: group theoretical aspects and applications, R.A. Askey, T.H. Koornwinder and W. Schempp (Eds.), Reidel Publishing Compagny, 1984.
89. J.F. Sturm, *Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones*, Optimization Methods and Software **11** (1999), 625–653.
90. B. Sturmfels, *Algorithms in Invariant Theory*, Springer-Verlag, 1993.
91. G. Szegö, *Orthogonal polynomials*, American Mathematical Society, 1939.
92. M. Takesaki, *Theory of Operator Algebras I.*, Encyclopaedia of Mathematical Sciences, 124. Operator Algebras and Non-commutative Geometry, 5. Springer-Verlag, Berlin, 2002.
93. H. Tarnanen, M. Aaltonen, J.-M. Goethals, *On the nonbinary Johnson scheme*, European J. Combin. **6** (1985), 279–285.
94. H. Tarnanen, *Upper bounds on permutation codes via linear programming*, European J. Combin. **20** (1999) 101–114.
95. F. Vallentin, *Optimal embeddings of distance transitive graphs into Euclidean spaces*, J. Comb. Theory Ser. B **98** (2008), 95–104.
96. F. Vallentin, *Symmetry in semidefinite programs*, Linear Algebra and Appl. 430 (2009), 360–369.
97. F. Vallentin, *Lecture notes: Semidefinite programs and harmonic analysis*, arXiv:0809.2017 [math.OC]
98. N.Ja. Vilenkin, A.U. Klimyk, *Representation of Lie Groups and Special Functions, Volume 2*, Kluwer Academic Publishers, 1993.
99. H.G. Wang, *Two-point homogeneous spaces*, Ann. of Math. **55** (1952), 177–191.
100. D.R. Woodall, *Cyclic-order graphs and Zarankiewicz's crossing-number conjecture*, J. Graph Theory **17** (1993) 657–671.
101. K. Zarankiewicz, *On a problem of P. Turán concerning graphs*, Fundamenta Mathematicae **41** (1954) 137–145.