The many faces of positivity to approximate structured optimization problems

Kuryatnikova, Olga

DOI:
10.26116/center-lis-1927

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

The many faces of positivity to approximate structured optimization problems

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof.dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op

vrijdag 20 september 2019 om 13.30 uur
door

Olga Kuryatnikova

geboren op 1 januari 1989 te Ryazan, Rusland

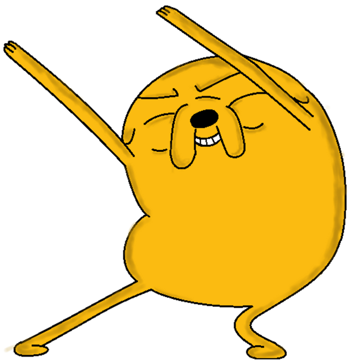
Promotiecommissie:

Promotores:	prof.dr.ir. Renata Sotirov
	dr. Juan C. Vera Lizcano
	dr. Luis F. Zuluaga
Overige leden:	prof.dr. Mirjam Dür
	prof.dr. Monique Laurent
	dr. Peter J.C. Dickinson
	dr. Fernando M. de Oliveira Filho

The many faces of positivity to approximate structured optimization problems

Dude, sucking at something is the first step towards being sort of good at something.

Jake the Dog, Adventure Time,
Season 1, Episode "His Hero"



Acknowledgements

The PhD years did not come easily to me, yet they were full of discoveries and inspiration due to the wonderful people who accompanied me in that journey.

I would like to thank my supervisor Juan Vera for so many things that a separate thesis would not be enough to describe them. I appreciate his motivating and patient academic supervision, demanding questions and keen research ideas. I am thankful for the discussions we had, both scientific and personal, for encouraging me to think outside the box and for listening to my opinion.

Next, I would like to thank Renata Sotirov for believing in my Research Master application and for supporting me both as an educational coordinator and as a supervisor during the final year of the PhD program. I enjoyed working with Renata, and I am grateful for her guidance and advice.

I am thankful to Luis Zuluaga for the research we did together and for giving me the opportunity to visit Lehigh University in May 2018. Besides generating ideas for this thesis, the visit opened a different way of working and looking at academia to me.

I would like to thank Mirjam Dür for hosting Juan Vera and me at Trier University in March 2018. Our discussions helped to improve this thesis, and I loved the warm atmosphere during the visit. Also, I am grateful to the organizers of the Oberwolfach Workshop on Copositivity and Complete Positivity 2017, Immanuel Bomze and Mirjam Dür, in particular, for letting me participate in that mind changing event.

Special thanks to Etienne de Klerk for making my commuting time to and from Rotterdam more valuable, for his help and advice.

I would also like to thank my PhD committee, Renata Sotirov, Juan Vera, Luis Zuluaga, Mirjam Dür, Monique Laurent, Peter Dickinson and Fernando de Oliveira Filho, for their time and effort spent reading this thesis and for their useful comments and suggestions.

I am grateful to all RM and PhD fellows I met in the last five years. In particular, to the group of students who made my life in the Netherelands fuller and brighter: Ahmadreza, Krzysztof, Marieke, Marleen, Roweno, Shobeir and Zeinab. I will never forget all the fun moments and conversations we had, and I appreciate your help and understanding. Also, talks about life and science with Ahmadreza, Bas, Hao, Pepijn and Riley were of great value to me.

Undoubtedly, my days would be harder without encouragement of my good old friends (forever young in our souls though!): Anastasia, Andrey, Anna, Elena, Irina, Kate and Lyuba.

Last but not least, I am thankful to my mother for her love and everything she gave me throughout the years.

Olga Kuryatnikova
Tilburg, September 2019

Contents

List of notation and acronyms

The chapters of this thesis contain distinct problems, and the notation and acronyms may differ per chapter. Below we present the list of notation and acronyms common for all chapters.

$\mathbb{R}^n$	n-dimensional Euclidean vector space
$\mathbb{R}^n_+$	cone of non-negative n-dimensional real vectors
$\mathbb{R}^n_{++}$	cone of strictly positive n-dimensional real vectors
$\mathbb{N}^n$	n-tuples of non-negative integers
$\mathbb{N}^n_+$	n-tuples of positive integers
$\mathbb{S}^n$	space of $n \times n$ real symmetric matrices
S^{n-1}	unit sphere in $\mathbb{R}^n$
$[n]$	$\{1, \ldots, n\}$
e	vector of all ones of an appropriate dimension
I_n	$n \times n$ identity matrix (we omit the subscript when the matrix size is clear)
J_n	$n \times n$ matrix of all ones (we omit the subscript when the matrix size is clear)
$C(V)$	space of real-valued continuous functions on the set V
$\mathcal{K}(V)$	space of symmetric real-valued continuous functions (kernels) on the set $V \times V$
$\mathcal{COP}(V)$	cone of copositive kernels on the set V
$\mathcal{PSD}(V)$	cone of positive definite kernels on the set V
$\mathcal{T}_d^V$	space of d-tensors on the set V
$\mathbb{S}^n_d$	space of symmetric d-tensors on the set $[n]$
$\mathbb{R}[x]$	ring of polynomials in variables x

$\mathbb{R}_d[x]$	ring of polynomials with degree not larger than d in variables x
$\mathbb{R}_{=d}[x]$	ring of polynomials with degree equal to d in variables x
$Sym(n)$	symmetric (or permutation) group on n elements
O_n	orthogonal group in dimension n
$\otimes$	Kronecker product
$\succeq$	$A \succeq B$ means that $A - B$ is a positive semidefinite matrix
$\succ$	$A \succ B$ means that $A - B$ is a positive definite matrix
IP	integer programming / program
SDP	semidefinite programming / program
LP	linear programming / program
SOCP	second-order cone programming / program
PSD	positive semidefinite (for matrices)
p.d.	positive definite (for kernels)
PO	polynomial optimization
SOS	sum-of-squares polynomials

CHAPTER 1

Introduction

In this thesis we obtain upper and lower bounds on several non-linear optimization problems using linear programming (LP) and semidefinite programming (SDP) relaxations. Throughout the thesis we use the notation SDP (resp. LP) do refer to both semidefinite (resp. linear) *programming* and *program*. The basic approach is to exploit the structure of a given problem, e.g., its combinatorial nature or inherent symmetry, to reformulate or relax the problem to the following general form:

$$\begin{aligned} \inf_f \ & L_0(f) \\ \text{s.t.}\ & L_1(f) = 0, \\ & f \in \mathcal{K}(V), \end{aligned} \tag{1.1}$$

where V is a set, L_0 and L_1 are affine operators, and $\mathcal{K}(V)$ denotes any of the following convex cones of continuous functions on V: the cone of entry-wise non-negative functions, the cone of positive (semi-) definite functions or the cone of copositive functions. Therefore problem (1.1) generalizes a classical *conic* problem where a convex function is minimized over the intersection of a convex cone and an affine subspace. Throughout the thesis, we deal with $V \subseteq \mathbb{R}^n$. Although some results can be extended to more general sets, we usually leave such extensions out of the scope of the thesis.

Different formulations of the initial problem would provide different relaxations. In this thesis we analyze conic formulations (1.1) since they exist for many optimization problems and allow to shift the hardness of the original problem to the conic constraint. As a result, the tools for dealing with the positive cones become available for the original problem.

Non-negativity, positive (semi-) definiteness and copositivity can be viewed as different faces of one notion – positivity. From here on we refer to the three cones in problem (1.1) as *the positive cones*. For more complex types of positive cones, there exist inner and/or outer approximations based on simpler positive cones, such as the cone of positive semidefinite matrices, the cone of non-negative matrices, or the second-order cone. In the sequel we treat the second-order cone as a special case of the positive semidefinite cone.

Assume we have a problem formulated as in (1.1) over a given positive cone. If optimization over this cone is not tractable, one can obtain LP or SDP relaxations of the original problem using approximations. The relaxations have the form (1.1) where the variables are matrices that belong to any of the mentioned cones, and the constraints are affine in the matrix variables. Problems of this form are also known as *linear matrix inequality (LMI)* problems.

LMI problems can be solved with the desired precision in polynomial time using, e.g., the interior point method by Nesterov and Nemirovski [156]. Hence one would like to obtain an LMI *reformulation* of a given problem. However, such reformulations may contain too many variables and constraints. Among others, this fact was pointed out in [28, 89, 209, 228]. Therefore it is common to reformulate the given problem to the form (1.1) and then use an LP or SDP *relaxation* to obtain LMI bounds on the optimal value if needed (see, for instance, LP relaxations in [209]).

1.1 The many faces of positivity

Three types of functions are essential in this thesis: matrices, kernels and polynomials. Let $\mathbb{S}^n$ be the cone of $n\times n$ *real symmetric matrices*. We generalize the notion of a symmetric matrix to the notion of a *kernel*. Denote the set of real-valued continuous functions on $V \subseteq \mathbb{R}^n$ by $C(V)$. The cone of kernels is the cone of symmetric real continuous functions on $V\times V$:

$$\mathcal{K}(V) = \{F \in C(V\times V) : F(x,y) = F(y,x),\ \forall\ x,y \in V\}.$$

We say that K is a *kernel on* V if $K \in \mathcal{K}(V)$. For any finite U of size n, $\mathcal{K}(U)$ is isomorphic to the cone of symmetric $n \times n$ matrices; we abuse the notation modulo this isomorphism, and thus we do not distinguish between kernels on finite sets and matrices in the rest of the thesis. Given $K \in \mathcal{K}(V)$ and $U \subseteq V$, we denote by $K^U \in \mathbb{S}^n$ the restriction of K to $U \times U$. For all $U \subseteq V$ and all $K \in \mathcal{K}(V)$ we have that $K^U \in \mathcal{K}(U)$.

For $n > 0$ we denote the set of n-variate polynomials with real coefficients by $\mathbb{R}[x] := \mathbb{R}[x_1,\dots,x_n]$. We denote by $\mathbb{R}_d[x]$ (respectively $\mathbb{R}_{=d}[x]$) the subset of $\mathbb{R}[x]$ of polynomials of degree not larger than (resp. equal to) d. The degree of a $p \in \mathbb{R}[x]$ is denoted by $\deg p$. For $d \geq 0$ we define $\mathbb{N}^n_d = \{\alpha \in \mathbb{N}^n : e^\intercal\alpha \leq d\}$. Given $h_1,\dots,h_m \in \mathbb{R}[x]$ and $\alpha \in \mathbb{N}^m_d$, we use $h^\alpha := \prod_{j=1}^m h_j^{\alpha_j}$. In particular $x^\alpha = \prod_{j=1}^m x_j^{\alpha_j}$. Also, we use the notation h to arrange the polynomials $h_1,\dots,h_m$ in an array; that is, $h := [h_1,\dots,h_m]^\intercal$. Finally, we use the names "homogeneous polynomials" and *"forms"* interchangeably. Notice that $\mathbb{S}^n$ is isomorphic to the set of forms of degree two.

1.1.1 Tensor representations

The set of homogeneous polynomials of degree two is isomorphic to $\mathbb{S}^n$, and similarly, homogeneous polynomials of higher degrees are connected to tensors. For a positive number $d \in \mathbb{N}$, a *tensor* of order $d \in \mathbb{N}$ over a set V (or a d-tensor over V) is a real-valued function on V^d. Denote by $\mathcal{T}_d^V$ the set of d-tensors on the set V. Denote by $Sym(d)$ the group of permutations on d elements.

Definition 1.1. *Let $\pi \in Sym(n), \sigma \in Sym(d)$.*

(a) We define the action π on any $v \in \mathbb{R}^n$ by $\pi v = [v_{\pi(1)}, \dots, v_{\pi(n)}]^\intercal$. That is, π permutes the entries of the vector.

(b) We define the right action of σ on any $M = [m_1, \dots, m_d] \in \mathbb{R}^{n\times d}$ by $M\sigma = [m_{\sigma(1)}, \dots, m_{\sigma(d)}]$. That is, the right action of σ permutes the columns of M. Now, let $M^\intercal = [m'_1, \dots, m'_n]$. We define the left action of π on M by $\pi M = [m'_{\pi(1)}, \dots, m'_{\pi(n)}]^\intercal$. That is, the left action of π permutes the rows of M.

Now, we define $\mathbb{S}_d^n$, the set of symmetric d-tensors on $V = [n]$, by

$$\mathbb{S}_d^n := \Big\{T \in \mathcal{T}_d^{[n]} : T(v_1, \dots, v_d) = T(v_{\pi(1)}, \dots, v_{\pi(d)}) \text{ for all } v_1, \dots, v_d \in V\Big)$$

For $T, S \in \mathcal{T}_d^{[n]}$, we define the tensor inner product of T and S by

$$\langle T, S\rangle := \sum_{v\in[n]^d} T(v)S(v). \tag{1.2}$$

There exists a connection between polynomials, matrices and kernels via tensors. Namely, for $d = 2$ and $V = [n]$, a tensor is an $n \times n$ matrix. In this thesis we abuse the notation and do not make a difference between $\mathbb{S}_2^n$ and $\mathbb{S}^n$. For $d = 2$ and $V \subseteq \mathbb{R}^n$, a symmetric and continuous tensor is a kernel. Finally, the connection between tensors and polynomials is as follows: with a tensor $T \in \mathcal{T}_d^{[n]}$, we associate a degree d form in n variables x:

$$T[x] := \sum_{v\in[n]^d} T(v) \prod_{i=1}^{d} x_{v_i} = \langle T, x^{\otimes d}\rangle, \tag{1.3}$$

where $\otimes$ denotes the Kronecker product, and

$$x^{\otimes d} = \underbrace{x \otimes x \otimes \cdots \otimes x}_{d}.$$

Moreover, for every homogeneous polynomial $p \in \mathbb{R}[x]$ there is a unique *symmetric tensor* for which the representation above is true.

Tensors, and symmetric tensors in particular, are widely used to obtain new results about non-negative polynomials (for instance, in [57]) or copositive matrices (see, e.g., [58]). In this thesis we use them not only for these purposes, but also to obtain new results about copositive kernels (see Chapter 2).

1.1.2 Positive (semi-) definite cones

Table 1.1 shows the positive cones used throughout the thesis. Next, we formally define these cones.

Table 1.1 – Positive cones considered in this thesis.

	positive (semi-) definite	non-negative	copositive
matrices	✓	✓	✓
kernels	✓	✓	✓
polynomials		✓	✓

A matrix $M \in \mathbb{S}^n$ is called *positive semidefinite* if $x^\intercal M x \geq 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. The matrix is called *positive definite* if $x^\intercal M x > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. An *SDP* is a linear optimization problem over the cone of positive semidefinite matrices. For more details on SDP, see the book [225].

We are also interested in the generalized idea of positive definiteness for kernels. Let $V \subseteq \mathbb{R}^n$. We follow the convention in the literature and call a kernel $K \in \mathcal{K}(V)$ *positive definite (p.d.)* if for any finite $U \subset V$ the matrix K^U is positive semidefinite. That is,

Definition 1.2. *$K \in \mathcal{K}(V)$ is positive definite (p.d.) if for any $u_1, \ldots, u_n \in V$ and $x \in \mathbb{R}^n$,*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} K(u_i, u_j) x_i x_j \geq 0.$$

We denote the cone of p.d. kernels on V by $\mathcal{PSD}(V)$. Notice that K is called positive definite when K^U is positive semidefinite and not necessarily positive definite. If $V \subset \mathbb{R}^n$ is a compact set, there exists another characterization of p.d. kernels.

Theorem 1.3 (Lemma 1 in Bochner [20])**.** *Let $V \subset \mathbb{R}^n$ be a compact set equipped with a finite measure μ strictly positive on open subsets. Then K is a p.d. kernel on V if and only if for any $g(x) \in C(V)$,*

$$\int_V \int_V K(x, y) g(x) g(y) d\mu(x) d\mu(y) \geq 0.$$

We denote by S^{n-1} the unit sphere in $\mathbb{R}^n$ and by O_n the orthogonal group in dimension n; that is, the group of $n \times n$ orthogonal matrices where the group operation is matrix multiplication.

Definition 1.4. *Let $P \in O_n$, $V \subseteq \mathbb{R}^n$ and $d > 0$.*

(a) We define the action of P on any $v \in \mathbb{R}^n$ by $v^P = Pv$. Similarly, we define the action of P on V by $V^P = \{v^P : v \in V\}$.

(b) We define the action of P on any $F \in C(V^d)$ as $F^P = F(v_1^P, \ldots, v_d^P)$ for all $v_1, \ldots, v_d \in V$.

(c) We define the left action of P on any $M \in \mathbb{R}^{n\times n}$ by $M^P = PM$ and the right action by ${}^P M = MP$. Similarly, we define the left action of P on any $\mathcal{M} \subseteq \mathbb{R}^{n\times n}$ by $\mathcal{M}^P = \{M^P : M \in \mathcal{M}\}$ and the right action by ${}^P\mathcal{M} = \{{}^P M : M \in \mathcal{M}\}$.

Based on Definition 1.4, we say that a function $F \in C(V^d)$ is *invariant under the action of O_n* if $F^P(v_1, \ldots, v_d) = F(v_1, \ldots, v_d)$ for all $P \in O_n$ and $v_1, \ldots, v_d \in V$. In this thesis we are particularly interested in p.d. kernels on S^{n-1} invariant under the action of O_n. General optimization problems over the cone of p.d. kernels are not efficiently solvable. However, p.d. kernels on S^{n-1} invariant under the action of O_n are well-studied by Schoenberg [201] and are frequently used in optimization [13, 43, 44].

1.1.3 Entry-wise non-negative cones

A function $f : V \to \mathbb{R}$ is called *entry-wise non-negative* if $f(v) \geq 0$ for all $v \in V$. The cone of entry-wise non-negative matrices is polyhedral, thus one can optimize over it in polynomial time. The cone of entry-wise non-negative (on a given set) polynomials is a complex object which has attracted a lot of research attention. Reznick [190] has written a good historical overview of these studies. In the sequel we call *entry-wise non-negative* polynomials simply *non-negative*.

Non-negative polynomials play an essential role in polynomial optimization (PO). PO problem is a problem of the following form: let $p, h_1, \ldots, h_m \in \mathbb{R}[x]$, we are interested in computing

$$\begin{aligned} &\inf_x \; p(x) \text{ s.t. } h_1(x) \geq 0, \ldots, h_m(x) \geq 0 \\ = &\sup_\lambda \; \lambda \text{ s.t. } p(x) - \lambda \geq 0 \text{ for all } x \text{ such that } h_1(x) \geq 0, \ldots, h_m(x) \geq 0, \end{aligned}$$

where the latter is the lower bound reformulation of the former. The lower bound formulation shows that solving a PO problem is equivalent to characterizing all polynomials non-negative on a feasible set of the problem. In general, optimization problems over non-negative polynomials are not efficiently solvable. One can approximate PO problems using approximations to the cone of non-negative polynomials on a given set.

1.1.4 Copositive cones

A matrix $M \in \mathbb{S}^n$ is called *copositive* if $x^\intercal M x \geq 0$ for all $x \in \mathbb{R}^n_+$. The matrix is called *strictly copositive* if $x^\intercal M x > 0$ for all $x \in \mathbb{R}^n_+$. A classical *copositive problem* is a linear optimization problem over the cone of copositive matrices. Now, a kernel $K \in \mathcal{K}(V)$ is copositive if for any finite $U \subset V$ the matrix K^U is copositive. That is,

Definition 1.5. *$K \in \mathcal{K}(V)$ is copositive if for any $u_1, \dots, u_n \in V$ and $x \in \mathbb{R}^n_+$,*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} K(u_i, u_j) x_i x_j \geq 0.$$

Copositive kernels were introduced by Dobre et al. [54], who also proposed an alternative definition in the spirit of Theorem 1.3.

Theorem 1.6 (Definition (2) in Dobre et al. [54] [20])**.** *Let $V \subset \mathbb{R}^n$ be a compact set equipped with a finite measure μ strictly positive on open subsets. Then K is a p.d. kernel on V if and only if for any $g(x) \in C(V)$ such that $g(x) \geq 0$ for all $x \in V$,*

$$\int_V \int_V K(x,y) g(x) g(y) d\mu(x) d\mu(y) \geq 0.$$

We denote the set of copositive kernels on V by $\mathcal{COP}(V)$. Notice that $\mathcal{PSD}(V) \subset \mathcal{COP}(V)$.

Finally, we consider the cone of copositive polynomials.

Definition 1.7. *$p \in \mathbb{R}[x]$ is copositive if $p(x) \geq 0$ for all $x \in \mathbb{R}^n_+$.*

Testing whether a given matrix is not copositive is NP-complete (see [148]), and therefore copositive problems are not efficiently solvable. The same conclusion applies to optimization problems over the cones of copositive kernels and polynomials as they contain the cone of copositive matrices (up to isomorphisms).

1.2 Approximations to copositive cones and the cone of non-negative polynomials

Some positive cones are hard to use in optimization while the other positive cones are more amenable for optimization. Often one can approximate the former using the latter. In this thesis we are especially interested in *inner approximations* to all copositive cones and the cone of non-negative polynomials; that is, in *subsets* of such cones. Hence in the sequel we present the most well-known inner approximations and provide references to existing outer approximations when possible.

1.2.1 Inner approximations to copositive cones

We first present a famous inner approximation to the cone of copositive polynomials since this approximation underlies some results for the cones of copositive matrices and kernels.

Theorem 1.8 (Pólya's Positivstellensatz [83]). *Let $p \in \mathbb{R}[x]$ be a homogeneous polynomial such that $p(x) > 0$ for all $x \in \mathbb{R}^n_+ \setminus \{0\}$. Then for some $r > 0$ all the coefficients of $(e^\intercal x)^r p(x)$ are non-negative.*

Given $r > 0$, a homogeneous polynomial for which the coefficients of $(e^\intercal x)^r p(x)$ are non-negative is clearly copositive. Pólya's Positivstellensatz implies that, when r goes to infinity, the set of homogeneous polynomials in $\mathbb{R}_d[x]$ for which the coefficients of $(e^\intercal x)^r p(x)$ are non-negative converges to the set of copositive homogeneous polynomials in $\mathbb{R}_d[x]$.

Pólya's Positivstellensatz is frequently used in literature to obtain results about non-negativity of polynomials and forms, such as [35, 171, 182, 206]. Outer approximations to the cones of non-negative polynomials on some sets, including the cone of copositive polynomials, was proposed by Lasserre [117].

Now, we move to the cone of copositive matrices. This is the most well-studied copositive cone. Extensive and structured information on this cone is provided in the theses by Dickinson [49] and Groetzner [73] and in the surveys by Bomze [23] and Dür [60]. There exist a variety of approximations to the cone of copositive matrices from the inside [26, 35, 83, 171, 173, 233] and from the outside [26, 120, 230].

Throughout the thesis we regularly use *sum-of-squares* polynomials, denoted by *SOS*.

Definition 1.9. *A polynomial $p \in \mathbb{R}_{2d}[x]$ is SOS if $p(x) = \sum_{i=1}^m q_i(x)^2$ for some $q_1, \ldots, q_m \in \mathbb{R}_d[x], m \in \mathbb{N}$.*

For a set V, we say that a sequence $(V_r)_{r \in \mathbb{N}_+}$ is *a hierarchy of subsets of* V, *or an inner hierarchy,* if $V_r \subseteq V_{r+1} \subseteq V$ for all $r \in \mathbb{N}_+$. One can define a hierarchy of *supersets* of V, or *an outer hierarchy,* analogously. Consider any $r > 0$. The following inner hierarchies for the cone of copositive matrices are frequently used in the literature and were introduced by de Klerk and Pasechnik [35], Peña et al. [173] and Parrilo [171], respectively. In the definitions of these hierarchies and later on e denotes the vector of all ones of an appropriate size.

$$\mathcal{C}_r^n := \Big\{ M \in \mathbb{S}^n : (e^\intercal x)^r (x^\intercal M x) \text{ has non-negative coefficients} \Big\}, \tag{1.4}$$

$$\mathcal{Q}_r^n := \Big\{ M \in \mathbb{S}^n : (e^\intercal x)^r (x^\intercal M x) = \sum_{e^\intercal \beta = r} x^\beta x^\intercal N_\beta x + \sum_{e^\intercal \beta = r} x^\beta x^\intercal S_\beta x, \tag{1.5}$$

$$N_\beta, S_\beta \in \mathbb{S}^n,\ N_\beta \geq 0 \text{ and } S_\beta \succeq 0 \text{ for all } \beta \in \mathbb{N}^n, e^\intercal \beta = r \Big\},$$

$$\mathcal{K}_r^n := \Big\{ M \in \mathbb{S}^n : \left(\sum_{i=1}^n x_i^2 \right)^r \sum_{i=1}^n \sum_{j=1}^n M_{ij} x_i^2 x_j^2 \text{ is SOS} \Big\}. \tag{1.6}$$

From the definitions one can immediately see that $\mathcal{C}_r^n \subseteq \mathcal{Q}_r^n \subseteq \mathcal{COP}([n])$ and $\mathcal{K}_r^n \subseteq \mathcal{COP}([n])$. Moreover,

$$\mathcal{C}_r^n \subseteq \mathcal{C}_{r+1}^n,\ \mathcal{Q}_r^n \subseteq \mathcal{Q}_{r+1}^n, \text{ and } \mathcal{K}_r^n \subseteq \mathcal{K}_{r+1}^n.$$

It is also known (see, e.g., [173]) that $\mathcal{Q}_r^n \subseteq \mathcal{K}_r^n$. This fact becomes clear from (1.7) further. All in all, we have

$$\mathcal{C}_r^n \subseteq \mathcal{Q}_r^n \subseteq \mathcal{K}_r^n \subseteq \mathcal{COP}([n]).$$

Finally, every strictly copositive matrix is contained in $\bigcup_r \mathcal{C}_r^n \subseteq \bigcup_r \mathcal{Q}_r^n \subseteq \bigcup_r \mathcal{K}_r^n$ [35, 173]. Hierarchies with the latter property are called *convergent*.

The convergence of $\mathcal{C}_r^n$ follows from Pólya's Positivstellensatz. Indeed, let $M \in \mathbb{S}^n$ be strictly copositive. Then, for a vector of variables $x = [x_1, \dots, x_n]^\intercal$, the form $x^\intercal M x$ is larger than zero on $\mathbb{R}_+^n \setminus \{0\}$. Hence, by Pólya's Positivstellensatz 1.8, there exists $r > 0$ such that $(e^\intercal x)^r x^\intercal M x$ has non-negative coefficients.

The convergence of $\mathcal{K}_r^n$ follows from Pólya's Positivstellensatz and some additional observations. Namely, let $M \in \mathbb{S}^n$ be strictly copositive and consider $r > 0$ such that all the coefficients of $q(x) := (e^\intercal x)^r x^\intercal M x$ are non-negative. Every $x \in \mathbb{R}_+$ can be written as $z^2,\ z \in \mathbb{R}$. By substituting x_i^2 for each $x_i,\ i \in [n]$ into $(e^\intercal x)^r x^\intercal M x$ and $q(x)$, we obtain the expression that is an SOS.

The convergence of $\mathcal{Q}_r^n$ follows from $\mathcal{C}_r^n \subseteq \mathcal{Q}_r^n$. Approximations $\mathcal{Q}_r^n$ stem from the result by Zuluaga et al. [233] that a homogeneous polynomial of degree $r+2$ is such that $p(x_1^2, \dots, x_n^2)$ is an SOS if and only if

$$p(x) = \sum_{\beta \in \mathbb{N}^n, e^\intercal \beta \leq r+2} x^\beta \sigma_\beta(x),\ \sigma_\beta \text{ is an SOS.} \tag{1.7}$$

Restricting ourselves to SOS of degrees zero and two in (1.7), we obtain the expressions for $\mathcal{Q}_r^n$. From (1.7) and (1.6) it is clear that $\mathcal{Q}_0^n = \mathcal{K}_0^n$ and $\mathcal{Q}_1^n = \mathcal{K}_1^n$.

As to infinite dimensional copositive programming, the most straightforward inner approximation to the cone of copositive kernels is the cone of positive definite kernels. In Chapter 2 we show how to generalize approximations $\mathcal{C}_r^n$ and $\mathcal{Q}_r^n$ for the case of copositive kernels. The generalized approximations $\mathcal{Q}_r^n$ include the set of p.d. kernels as a subset.

1.2.2 Inner approximations to the cone of non-negative polynomials

To approximate the cone of non-negative polynomials, it is common to use sums-of-squares (SOS) polynomials, which are clearly non-negative. Verifying whether a given polynomial is an SOS is equivalent to solving an SDP (see, for instance, [171]). SOS polynomials of fixed degrees form a proper cone (closed, convex, pointed, with nonempty interior). Recently, there have been successful attempts to apply the interior point method directly to this cone, without SDP reformulation [170]. However, this approach has not yet been broadly used in the literature.

SOS polynomials frequently occur in relation to seminal Schmüdgen's Positivstellensatz [199] and Putinar's Positivstellensatz [183].

Theorem 1.10 (Schmüdgen's Positivstellensatz [199])**.** *Let* $h_1, \dots, h_m \in \mathbb{R}[x]$ *be such that* $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, ..., h_m(x) \geq 0\}$ *is non-empty and compact, and assume that* $p(x) > 0$ *for all* $x \in S$*. Then there is* $r \geq 0$ *such that*

$$p = \sum_{\alpha \in \{0,1\}^m} \sigma_\alpha h^\alpha, \tag{1.8}$$

for some SOS polynomials σ_α *of degree* $r - \deg h^\alpha$ *for all* $\alpha \in \{0,1\}^m$*.*

For polynomials $h_1, \dots, h_m \in \mathbb{R}[x]$, we define their *quadratic module* as

$$\mathcal{QM}(h_1, \dots, h_m) = \{p \in \mathbb{R}[x] : p = \sigma_0 + \sum_{j=1}^{m} \sigma_j h_j,\ \sigma_j,\ j \in \{0, \dots, m\} \text{ are SOS.}\} \tag{1.9}$$

Definition 1.11. *Let* $h_1, \dots, h_m \in \mathbb{R}[x]$*. The quadratic module* $\mathcal{QM}(h_1, \dots, h_m)$ *is called* Archimedean *if there exists* $N > 0$ *such that* $N - \|x\|^2 \in \mathcal{QM}(h_1, \dots, h_m)$*.*

Theorem 1.12 (Putinar's Positivstellensatz [183])**.** *Let* $h_1, \dots, h_m(x) \in \mathbb{R}[x]$ *be such that* $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, ..., h_m(x) \geq 0\}$ *is non-empty and* $\mathcal{QM}(h_1, \dots, h_m)$ *is Archimedean, and assume that* $p(x) > 0$ *for all* $x \in S$*. Then there is* $r \geq 0$ *such that*

$$p = \sigma_0 + \sum_{j=1}^{m} \sigma_j h_j, \tag{1.10}$$

for some SOS polynomials σ_j *of degree* $r - \deg h_i$ *for all* $j \in \{0, \dots, m\}$*.*

Clearly, sets of all polynomials which have representations (1.8), (1.10) are subsets of polynomials non-negative on S. These and other representations that make the non-negativity of p on S evident are called *certificates of non-negativity of* p *on* S*.*

The degree of SOS to use in certificates (1.8) and (1.10) is usually unknown. By growing the degree, one obtains *a hierarchy* of approximations to the set of non-negative polynomials. Notice that under the assumptions of Schmüdgen's and Putinar's Positivstellensatzen, when the degree of SOS goes to infinity, the resulting approximations converge to the set of positive polynomials on S. This convergence is an important property since it allows to obtain convergent approximations to PO problems on S (see, for instance, [114]).

The size of the SDP corresponding to a general SOS certificate of non-negativity grows exponentially with the number of variables n and the number of polynomials m. To deal with this growth, much research attention is directed to certificates of non-negativity that are not based on SOS. A well-known result that does not use SOS is Handelman's Positivstellensatz.

Theorem 1.13 (Handelman's Positivstellensatz [82]). *Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and let $S = \{x : Ax \le b\}$ be a non-empty polytope. If $p(x) > 0$ for all $x \in S$, then*

$$p(x) = \sum_{\alpha\in\mathbb{N}^m} c_\alpha (b - Ax)^\alpha, \tag{1.11}$$

for some $c_\alpha \ge 0$ for all $\alpha \in \mathbb{N}^m$.

Representation (1.11) is linear in c_α, $\alpha \in \mathbb{N}^m$. Another certificate of this type was proposed by Dickinson and Povh [48]. We present this certificate for the case of compact sets.

Theorem 1.14 (Theorem 3.10. in [48]). *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$, and let $S = \{x \in \mathbb{R}^n_+ : h_1(x) \ge 0, \dots, h_m(x) \ge 0\}$ be non-empty. Denote $d_{max} = \max\{\deg h_1, \dots, \deg h_m, \deg p\}$. Assume that $h_1(x) = 1$ and $h_j(x) = M - e^\intercal x$ for some $M > 0$ and $j \in [m]$. If $p(x) > 0$ for all $x \in S$, then there exists $r \ge 0$ such that*

$$(1 + e^\intercal x)^{d_{\max} - \deg p + r}\, p(x) = \sum_{j=1}^{m} \sum_{\alpha^j \in \mathbb{N}^n_{d_{\max} - \deg h_j + r}} c_{\alpha^j} x^{\alpha^j} h_j(x), \tag{1.12}$$

where $c_{\alpha^j} \ge 0$ for all $\alpha^j \in \mathbb{N}^n_{d_{\max} - \deg h_j + r}$, $j \in [m]$.

One could also replace SOS in the certificates by alternative subsets of non-negative polynomials. Examples of such subsets are SOS constructed using subsets of monomials in $\mathbb{R}[x]$ [96, 115, 218, 221], scaled diagonally dominant sums-of-squares [2, 3], non-negative circuit polynomials [59, 91, 219] or hyperbolic polynomials [189, 197]. In Chapter 4 we derive a new certificate of non-negativity that is based not on SOS but on *copositive* polynomials.

1.3 Copositive reformulations of hard optimization problems

Among all formulations (1.1), we are especially interested in copositive ones since they exist for a large variety of problems. For instance, many discrete optimization problems can be written as classical copositive problems or duals of these problems [28]. Some examples are graph parameters, such as the independence number [35], the chromatic number [80] and the fractional chromatic number [180]. A copositive optimization problem is continuous and convex. This fact makes available tools from convex optimization, such as symmetry reductions, to discrete optimization.

Problems in discrete geometry can, too, be formulated as problem (1.1) over the cone of copositive kernels. Some examples are the stable set problem in topological packing graphs [54] and the measurable stable set problem in locally-independent graphs [43]. There are only two major techniques by Bachoc and Vallentin [12] and Delsarte et al. [44] that are numerically efficient for the spherical codes problem, which is an example of the stable set problem in topological packing graphs. Therefore new approaches to deal with problems in discrete geometry are of interest for the optimization community.

Finally, copositive polynomials appear in reformulations of general PO problems. For instance, several broad classes of quadratic problems have copositive formulations [14, 28, 29], and also optimization problems over sets defined by specific polynomial equalities [175]. Using copositivity allows applying the existing results for copositive polynomials to general PO problems. In Chapter 4 we show that a generic PO problem can be formulated as an optimization problem over the cone of copositive polynomials, which connects copositive programming to a variety of real-life problems with PO formulations.

1.4 Exploiting structure in optimization problems

We write general optimization problems using formulation (1.1) to represent (or relax) these possibly not convex problems as LMI problems. In this way, we can work with non-convex problems using the machinery from convex analysis.

1.4.1 Symmetry

One of the advantages of conic optimization is the possibility to exploit the symmetry of the problem efficiently. We say that a problem is symmetric if one can non-trivially map the set of its variables onto itself without changing the structure of the problem. For instance, one can permute or rotate the variables. Some typical symmetric problems are graph coloring, constructing binary and spherical codes, scheduling jobs on parallel machines. In this thesis we deal with each of these problems: we consider the maximum k-colorable subgraph problem in Chapter 5, the spherical codes problem in Chapters 2 and 3, and the problem of scheduling on selfish, not identical machines in Chapter 6. All these problems are non-convex and NP-hard.

Symmetry hampers the performance of enumeration algorithms for integer programs (IP), such as branch and bound or branch and cut. Symmetry implies many optimal solutions and many isomorphic subproblems in the enumeration tree. This fact leads to wasting the computational effort of enumeration algorithms [138]. The primary approach to tackling symmetry in an IP problem is to break this symmetry by fixing the values of some variables, perturbing the problem or adding valid non-symmetric inequalities to the problem. Most recent algorithms can recognize symmetrically equivalent solutions and either discard them or treat them differently [68, 165, 178]. These algorithms reduce the size of the enumeration tree but do not simplify the structure of the problem or reduce the number of variables in the problem.

On the other hand, symmetry helps to reduce the size of convex problems. Therefore symmetry is the main structural property we use in this thesis. If problem (1.1) is invariant under the action of a group, it is enough to optimize over the solutions to the problem invariant under the action of this group. This approach can be extremely efficient in convex programming, see [53], [70] or [37]. The space of invariant solutions usually has a lower dimension than the original space of variables, which results in dramatically reducing the size of the problem. We consider an example of this procedure in Chapter 5. Also, the space of invariant solutions can have an advantageous structure which allows for efficiently solvable approximations to the original problem. We use this fact in Chapters 2, 3 and 6. For instance, invariant positive definite kernels on the unit hypersphere [151, 201] have explicit characterizations which we use in Chapter 3.

1.4.2 Strongly positive polynomials

Chapter 4 deals with polynomials which are "strongly positive". To define this concept, let $p \in \mathbb{R}[x]$ be a polynomial of degree $\deg p$, and consider its highest degree homogeneous component $\tilde{p}(x)$ obtained by dropping from $p(x)$ all the terms whose total degree is less than $\deg p$. This component determines the behavior of p at infinity on unbounded sets. Namely, if $\tilde{p}(y) > 0$ for some $y \in \mathbb{R}^n$, then there is $k > 0$ such that $p(ky) > 0$ since the positive homogeneous component of the highest degree dominates the behavior of p for k large enough. However, if $\tilde{p}(y) = 0$, we do not know how the polynomial behaves when ky goes to infinity. This fact may complicate detecting whether p is non-negative on a given unbounded set.

Consider $p \in \mathbb{R}[x]$ and a set $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$. Let $\tilde{S} = \{x \in \mathbb{R}^n : \tilde{h}_1(x) \geq 0, \dots, \tilde{h}_m(x) \geq 0\}$. We say that p is strongly positive on S if $p(x) > 0$ for all $x \in S$, and $\tilde{p}(x) > 0$ for all $x \in \tilde{S} \setminus \{0\}$. Assumptions related to strong positivity have been used in the literature [48, 78, 79, 159, 185] to obtain valid approximations of polynomial optimization problems. In Chapter 4 we show that strongly positive polynomials have certificates of non-negativity based on copositive polynomials.

1.5 Overview of the thesis

The rest of this thesis consists of five self-contained chapters.

We begin in Chapter 2 where we study the kissing number problem using positive definite approximations to the cone of copositive kernels. The kissing number is the maximum number of non-overlapping unit hyper-spheres that can simultaneously touch another unit sphere, in n-dimensional space. It has been shown by Dobre et al. [54] that the maximum stable set problem in some infinite graphs, and the kissing number problem in particular, reduces to a minimization problem over the cone of copositive kernels. Optimizing over this infinite dimensional cone is not tractable, and approximations of this cone have been hardly considered in the literature. We propose two convergent hierarchies of subsets of copositive kernels, in terms of non-negative and positive definite kernels. Using these hierarchies, we construct upper bound relaxations to the kissing number problem.

To implement our bounds on kissing numbers, we extend the famous theorem of Schoenberg [201] that characterizes positive definite kernels on the unit sphere S^{n-1} invariant under the automorphisms of the sphere. This is done in Chapter 3. We obtain two generalizations of Schoenberg's theorem. The first one characterizes invariant (under the action of O_n) p.d. kernels on a product of S^{n-1} and a compact set which can depend on given parameters. Our second result characterizes invariant (under the action of O_n) continuous functions F on $(S^{n-1})^{r+2}$ such that $F(\cdot,\cdot,Z)$ is positive definite for every $Z \in (S^{n-1})^r$. When Z is fixed, this class reduces to the class of p.d. kernels invariant under the stabilizer of Z in the automorphism group of the sphere. For $r = 0$ and $r = 1$, these kernels have been used to obtain upper bounds on kissing numbers. We use our extension for $r > 1$ to implement the bounds for the kissing number problem from Chapter 2. The resulting bounds for $r \in \{0, 1, 2\}$ are fast-to-compute and lie between the currently existing LP and SDP bounds.

In Chapter 4 we show how to certify the non-negativity of polynomials using copositive polynomials. The certificates we obtain are valid for generic basic closed semi-algebraic sets and have a fixed small degree, while commonly used SOS certificates are guaranteed to exist only for compact sets and could have a large degree. The main benefit of our copositive certificates of non-negativity is their ability to translate results known exclusively for copositive polynomials to more general basic closed semialgebraic sets. In particular, we show how to use copositive polynomials to construct structured (e.g., sparse) certificates of non-negativity, even for unstructured semialgebraic sets. Last but not least, copositive certificates can be used to obtain hierarchies of tractable upper and lower bounds for PO problems.

Next, in Chapter 5 we consider the maximum k-colorable subgraph problem. For a given graph with n vertices, we look for the largest induced subgraph that can be colored in k colors such that no two adjacent vertices have the same color. This is a discrete optimization problem which admits a copositive reformulation and SDP relaxations. We propose several new SDP relaxations for this problem. The initial matrix size in the relaxations grows with n and k. We use the invariance of the problem under the color permutations to reduce the matrix size in the problem to order $(n+1)$, independently of k or the particular graph considered. Our relaxations show better numerical results than the existing SDP and IP-based relaxations for the majority of tested graphs.

In the final Chapter 6 we consider the problem of allocating tasks to unrelated parallel machines to minimize the time to complete all the tasks. The machines belong to agents who have to be paid, aim to maximize their utility and can lie about processing times of their machines. We are interested in the best approximation ratio R_n of a subclass of truthful mechanisms for n tasks on two machines. Using the symmetry of the problem, we propose a new continuous $min - max$ optimization model for finding R_n, as well as LP upper and lower bounds on R_n. The bounds are based on pointwise and piecewise approximations of cumulative distribution functions. Our method improves upon the existing bounds on R_n for small n. In particular, for $n = 2$ we show that $|R_2 - 1.505996| < 10^{-6}$.

1.6 Contributions to the literature

This thesis is based on the five research papers listed below. Each paper contains ideas and contributions from all its respective authors.

Chapter 2 O. Kuryatnikova and J. C. Vera. Positive semidefinite approximations to the cone of copositive kernels. 2018. Submitted. Extended abstract [109], ArXiv preprint 1812.00274 [110].

Chapter 3 O. Kuryatnikova and J. C. Vera. Generalizations of Schoenberg's theorem on positive definite kernels. 2019. Working paper. ArXiv preprint 1904.02538 [111].

Chapter 4 O. Kuryatnikova, J. C. Vera and L.F. Zuluaga. Copositive certificates of non-negativity for polynomials on unbounded sets. 2019. Submitted.

Chapter 5 O. Kuryatnikova, R. Sotirov and J. C. Vera. New SDP bounds on the maximum k-colorable subgraph problem. 2019. Working paper.

Chapter 6 O. Kuryatnikova and J. C. Vera. New bounds for truthful scheduling on two unrelated selfish machines. *Theory of Computing Systems*, 2019, [112], online first: https://doi.org/10.1007/s00224-019-09927-x.

CHAPTER 2

Positive semidefinite approximations to the cone of copositive kernels

2.1 Introduction

In this chapter we are interested in solution methods for infinite-dimensional copositive optimization, that is the optimization model obtained by replacing (finite-dimensional) copositive matrices with copositive kernels, which are their infinite-dimensional counterpart. Generalizing copositive optimization to infinite dimensions is inspired by successful infinite-dimensional generalizations of semidefinite programming (SDP). Such generalizations have proven useful in obtaining bounds for graph parameters in infinite graphs, by formulating an infinite-dimensional version of well-known SDP relaxations. In these relaxations PSD matrices are generalized to p.d. kernels. One of the applications of p.d. kernels is generalizing the famous Lovász ϑ-number [129] from finite graphs to certain types of infinite graphs. This fact has motivated some of the new results in packing problems in discrete geometry [38], the bounds on the measurable chromatic number [13] and the measurable stable set of infinite graphs [42]. In the finite case, some graph parameters for which Lovász ϑ-number provides a bound, such as the stable set or the chromatic number, can be formulated using copositive optimization. In the infinite case, the stable set problem in topological packing graphs [54] and the measurable stable set problem in locally-independent graphs [43] have been formulated using infinite-dimensional copositive optimization. We expect that in future more problems will be represented using infinite-dimensional copositive optimization and thus our results will be useful there.

Several methods have been proposed to approximately solve finite-dimensional copositive problems. The most usual approach is to approximate the copositive cone from the inside [35, 171, 173] or from the outside [120, 230]. Some researchers also exploit the structure of the problem and properties of the objective function [26, 227]. In the infinite-dimensional case, there are not many approximations for the cone of copositive kernels. The only known approach is to replace this cone by the better-studied cone of p.d. kernels, which is a subset of the cone of copositive kernels. When the kernels are defined on the unit sphere in $\mathbb{R}^n$, this results in a tractable relaxation of an infinite-dimensional copositive program using the characterization of p.d. kernels by Schoenberg [201], see Theorem 3.1.

One of the main contributions, In Section 2.3, is the definition of two converging inner hierarchies of subsets of the cone of copositive kernels on V for any compact $V \subset \mathbb{R}^n$. Our inner approximations generalize two existing inner hierarchies for copositive matrices (1.5) and (1.4), introduced in Chapter 1.The key element of our approach is to redefine the approximations using tensors. We also show that the new hierarchies provide converging upper bounds for the stable set problem when applied to the results by Dobre et al. [54]. Another contribution of this chapter is the application of the proposed hierarchies to construct *convergent* upper bounds on the kissing number (see Section 2.4).

De Laat in his PhD thesis [38] and in the related papers with Vallentin [39] and De Oliveira Filho [39, 40] provides a different type of p.d. kernel based approximation for the stable set problem on compact topological packing graphs. These approximations are not explicitly based on approximating the cone of copositive kernels but use a generalized version of the broadly used Lasserre's hierarchy [114]. The latter exploits Putinar's Positivstellensatz 1.12 on the polynomial optimization formulation of the stable set problem. Another well-known approximation by Bachoc and Vallentin [12] to a particular case of the stable set problem on compact topological packing graphs – the kissing number problem – is also based on the generalized Lasserre's hierarchy. The relation between our approximations and the approximations based on Lasserre's hierarchy is an interesting question for further research.

The outline of the chapter is as follows. In Section 2.2 we introduce the notation and provide more detail on copositive and positive definite kernels, as well as on tensors and tensor operators. In Section 2.3 we introduce generalized hierarchies (2.9) and (2.11), describe their main properties in Theorem 2.9 and show that they provide convergent upper bounds for the stable set problem by Dobre et al. [54] (Theorem 2.17). Finally, in Section 2.4, we show how to obtain hierarchies of convergent upper bounds for the kissing number problem using our results.

2.2 Tensor operators and their properties

To introduce our hierarchies of subsets for $\mathcal{COP}(V)$, we use the connection between tensors, kernels, matrices, and polynomials described in Section 1.1.1 of Chapter 1. We use tensor notation and terminology similar to Dong [58].

We begin by introducing two operators used to lift a given matrix to the space of symmetric tensors. The first operator is a lifting operator. For $r \geq 0$, we define the r-stack, $\mathbf{Stk^r} : \mathcal{T}_d^V \to \mathcal{T}_{d+r}^V$ as the operator that stacks r copies of a given tensor T

on each other; that is,

$$\mathbf{Stk^r}(T) := T \otimes e^{\otimes r}.$$

It follows that for all $T \in \mathcal{T}_d^V$, $u_1, \ldots, u_d, v_1, \ldots, v_r \in V$,

$$\mathbf{Stk^r}(T)(u_1, \ldots, u_d, v_1, \ldots, v_r) := T(u_1, \ldots, u_d). \tag{2.1}$$

Notice that $\mathbf{Stk}^0(T) := T$.

Remark 2.1. *In this chapter the notation u_i may refer to an entry of a vector, an element of a tuple of vectors, or a column of a matrix. We provide no additional information when the exact meaning of the notation is clear from the context.*

The second operator is the symmetrization operator $\sigma : \mathcal{T}_d^V \to \mathcal{T}_d^V$ which we define for any $T \in \mathcal{T}_d^V$ and $v_1, \ldots, v_d \in V$ by

$$\sigma(T)(v_1, \ldots, v_d) := \frac{1}{d!} \sum_{\pi \in Sym(d)} T(v_{\pi(1)}, \ldots, v_{\pi(d)}). \tag{2.2}$$

Lemma 2.2. *For $T \in \mathcal{T}_d^V$, $u_1, \ldots, u_d, v_1, \ldots, v_r \in V$ we have*

$$\sigma(\boldsymbol{Stk^r}(T))(u_1, \ldots, u_d, v_1, \ldots, v_r) := \tfrac{r!}{(r+d)!} \sum_{w_1, \ldots, w_d \in \{u_1, \ldots, u_d, v_1, \ldots, v_r\}} T(w_1, \ldots, w_d). \tag{2.3}$$

Proof. The result follows immediately from the definitions (2.1) and (2.2). □

Now, let $M \in \mathbb{S}^n$ and consider the polynomial $(e^\intercal x)^r(x^\intercal Mx)$. Notice that for any $v \in [n]^{r+2}$ and $\pi \in Sym(r+2)$ we have $x_{v_1} \cdots x_{v_{r+2}} = x_{v_{\pi(1)}} \cdots x_{v_{\pi(r+2)}}$. Recall that we abuse the notation and do not make a difference between $\mathbb{S}_2^n$ and $\mathbb{S}^n$, and thus we can apply tensor operators to M to obtain

$$\begin{aligned}(e^\intercal x)^r(x^\intercal Mx) &= \langle M \otimes e^{\otimes r}, x^{\otimes(r+2)} \rangle \\ &= \langle \sigma(M \otimes e^{\otimes r}), x^{\otimes(r+2)} \rangle = \langle \sigma(\mathbf{Stk^r}(M)), x^{\otimes(r+2)} \rangle.\end{aligned} \tag{2.4}$$

This implies that $\sigma(\mathbf{Stk^r}(T))$ is the *unique symmetric* tensor associated to $(e^\intercal x)^r(x^\intercal Mx)$. In this way we lift $M \in \mathbb{S}^n$ to the space $\mathbb{S}_{r+2}^n$.

Lemma 2.3. *Let $d > 0$, $r \geq 0$, $V \subset \mathbb{R}^n$, $T, S \in \mathcal{T}_d^V$. Then*

(a) $\sigma(T + S) = \sigma(T) + \sigma(S)$.

(b) $\boldsymbol{Stk^r}(T + S) = \boldsymbol{Stk^r}(T) + \boldsymbol{Stk^r}(S)$.

(c) $\boldsymbol{Stk}^{\mathbf{r+1}}(T) = \boldsymbol{Stk}^{\mathbf{1}}\left(\boldsymbol{Stk}^{\mathbf{r}}(T)\right).$

(d) If $\sigma(T) = \sigma(S)$*, then* $\sigma\left(\boldsymbol{Stk}^{\mathbf{r}}(T)\right) = \sigma\left(\boldsymbol{Stk}^{\mathbf{r}}(S)\right)$.

Proof. *a.*, *b.* and *c.* are straightforward. To prove *d.*, assume $\sigma(T) = \sigma(S)$. For $r = 0$,

$$\sigma\left(\mathbf{Stk}^{\mathbf{r}}(T)\right) = \sigma(T) = \sigma(S) = \sigma\left(\mathbf{Stk}^{\mathbf{r}}(S)\right).$$

Using *c.* and induction, it is enough now to prove the statement for $r = 1$. For any $v_1, \ldots, v_{d+1} \in V$,

$$\begin{aligned}
\sigma\left(\mathbf{Stk}^{\mathbf{1}}(T)\right)(v_1, \ldots, v_{d+1}) &\overset{(2.3)}{=} \frac{1}{(d+1)!} \sum_{w_1,\ldots,w_d \in \{v_1,\ldots,v_{d+1}\}} T(w_1, \ldots, w_d) \\
&= \frac{d!}{(d+1)!}\left(\sum_{k=1}^{d+1} \sigma(T)(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_{d+1})\right) \\
&= \frac{d!}{(d+1)!}\left(\sum_{k=1}^{d+1} \sigma(S)(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_{d+1})\right) \\
&= \sigma\left(\mathbf{Stk}^{\mathbf{1}}(S)\right)(v_1, \ldots, v_{d+1}).
\end{aligned}$$

□

Finally, we introduce the operator that projects high order tensors to lower order tensors and matrices, in particular. For any $d' \leq d$, and $v_1, \ldots, v_{d-d'} \in V$ we define the d'-slice, $\mathbf{Slc}^{\boldsymbol{v}} : \mathcal{T}_d^V \to \mathcal{T}_{d'}^V$ by

$$\mathbf{Slc}^{\boldsymbol{v}}(T)(u_1, \ldots, u_{d'}) := T(u_1, \ldots, u_{d'}, v_1, \ldots, v_{d-d'}) \quad \text{for all } T \in \mathcal{T}_d^V,\ u_1, \ldots, u_{d'} \in V. \tag{2.5}$$

That is, the slice is obtained by fixing the last indices of the tensor at $v_1, \ldots, v_{d-d'}$. For $T \in \mathcal{T}_d^{[n]}$ consider its associated polynomial $T[x]$ defined in (1.3). A monomial in $T[x]$ is denoted by $x_1^{\beta_1} \cdots x_n^{\beta} = x^{\beta}$ where $\beta \in \mathbb{N}^n$ and $\sum_{i=1}^n \beta_i = d$. Each x^{β} corresponds to $\prod_{i=1}^d x_{v_i}$ for some $v \in [n]^d$. Permuting the elements of v does not change the corresponding monomial. Let

$$\mathcal{P}_{\beta} = \{v \in [n]^d : \prod_{i=1}^{d} x_{v_i} = x^{\beta}\}. \tag{2.6}$$

Then the coefficient of the monomial x^{β} in $T[x]$ is $\sum_{v \in \mathcal{P}_{\beta}} T(v)$. We use several properties of $T[x]$ which are described using the following two lemmas:

Lemma 2.4. *Let* $n, d \in \mathbb{N}_+$ *and* $T \in \mathcal{T}_d^{[n]}$*. Let* x^{β} *be a monomial in* $T[x]$*. Then for any* $u, v \in \mathcal{P}_{\beta}$ *and* $\pi \in Sym(d)$

(a) $\pi v \in \mathcal{P}_\beta$.

(b) $\sigma(T)(v) = \sigma(T)(u)$.

Lemma 2.5. *Let $n, d \in \mathbb{N}_+$.*

(a) For $T \in \mathcal{T}^{[n]}_{d+2}$, $T[x] = \sum_{v \in [n]^d} \mathbf{Slc^v}(T)[x] \prod_{i=1}^d x_{v_i}$

(b) For any $r \in \mathbb{N}$ and $T \in \mathcal{T}^{[n]}_d$: $\mathbf{Stk^r}(T)[x] = (e^\intercal x)^r T[x]$

(c) For any $T, S \in \mathcal{T}^{[n]}_d$: $S[x] = T[x]$ if and only if $\sigma(S) = \sigma(T)$.

Proof. *a*. Let $T \in \mathcal{T}^{|n|}_{d+2}$. From (2.5),

$$T[x] = \sum_{u \in [n]^{d+2}} T(u) \prod_{i=1}^{d+2} x_{u_i} = \sum_{v \in [n]^d} \sum_{u_1, u_2 \in [n]} T(u_1, u_2, v) x_{u_1} x_{u_2} \prod_{i=1}^d x_{v_i}$$
$$= \sum_{v \in [n]^d} \mathbf{Slc^v}(T)[x] \prod_{i=1}^d x_{v_i}$$

b. Let $T \in \mathcal{T}^{[n]}_d$. First, consider the case $r = 0$:

$$\mathbf{Stk^r}(T)[x] = T[x] = (e^\intercal x)^0 T[x].$$

Now we show that $\mathbf{Stk^1}(T)[x] = (e^\intercal x) T[x]$ using *a*., then, by Lemma 2.3 *c*., the statement follows by induction.

$$\mathbf{Stk^1}(T)[x] = \sum_{u \in [n]^{d+1}} \mathbf{Stk^1}(T)(u) \prod_{i=1}^{d+1} x_{u_i}$$
$$= \sum_{u_{d+1} \in [n]} x_{u_{d+1}} \sum_{u_1, \ldots, u_d \in [n]} \mathbf{Stk^1}(T)(u) \prod_{i=1}^d x_{u_i} = (e^\intercal x) T[x].$$

c. Let $T, S \in \mathcal{T}^{[n]}_d$. For any $\beta \in \mathbb{N}^n$ such that $\sum_{i=1}^n \beta_i = d$, let $\mathcal{P}_\beta$ be defined as in (2.6). Then by Lemma 2.3 *a*. and Lemma 2.4, for any $v \in \mathcal{P}_\beta$

$$\sigma(T)(v) = \frac{1}{|\mathcal{P}_\beta|} \sum_{u \in \mathcal{P}_\beta} \sigma(T)(u) = \frac{1}{|\mathcal{P}_\beta|} \sigma \left(\sum_{u \in \mathcal{P}_\beta} T(u) \right) = \frac{1}{|\mathcal{P}_\beta|} \sum_{u \in \mathcal{P}_\beta} T(u).$$

Recall that $\sum_{u \in \mathcal{P}_\beta} T(u)$ is the coefficient of x^β, which concludes the proof. $\square$

2.3 Inner hierarchies for the cone of copositive kernels

In the sequel use the previously defined tensor operators together with the following two sets of tensors.

Definition 2.6. *A tensor $T \in \mathcal{T}_d^V$ is entry-wise non-negative if $T(v_1, \ldots, v_d) \geq 0$, for all $v_1, \ldots, v_d \in V$. We use $\mathcal{N}_d^V$ to denote the set of all entry-wise non-negative d-tensors on V.*

Definition 2.7. *A tensor $F \in \mathcal{T}_{d+2}^V$ is 2-p.d. on V if it is continuous, and for all $v_1, \ldots, v_d \in V$, $F(\cdot, \cdot, v_1, \ldots, v_d) \in \mathcal{PSD}(V)$.*

In this section we generalize the hierarchies $\mathcal{C}_r^n$ (1.4) and Q_r^n (1.5) from matrices to kernels. To provide an intuition for this generalization, we first write the hierarchy $\mathcal{C}_r^n$ in tensor form, based on (2.4):

$$\mathcal{C}_r^n = \left\{ M \in \mathbb{S}^n : (e^\intercal x)^r (x^\intercal M x) \text{ has non-negative coefficients} \right\} \tag{2.7}$$

$$= \left\{ M \in \mathbb{S}^n : \sigma(\mathbf{Stk^r}(M)) \in \mathcal{N}_{r+2}^{[n]} \right\}. \tag{2.8}$$

Based on the tensor reformulation (2.8), we introduce the following sets:

$$\mathcal{C}_r^V = \left\{ K \in \mathcal{K}(V) : \sigma\left(\mathbf{Stk^r}(K)\right) \in \mathcal{N}_{r+2}^V \right\}, \tag{2.9}$$

$$= \left\{ K \in \mathcal{K}(V) : \sum_{i,j \in [r+2]} K(v_i, v_j) \geq 0 \text{ for all } v_1, \ldots, v_{r+2} \in V \right\}, \tag{2.10}$$

$$Q_r^V = \left\{ K \in \mathcal{K}(V) : \sigma\left(\mathbf{Stk^r}(K)\right) - \sigma(S) \in \mathcal{N}_{r+2}^V, \right. \tag{2.11}$$

$$\left. \text{for some 2-p.d. } S \in \mathcal{T}_{r+2}^V \right\}.$$

$$= \left\{ K \in \mathcal{K}(V) : \sum_{i,j \in [r+2]} K(v_i, v_j) - \sigma(S)(v_1, \ldots, v_{r+2}) \geq 0 \right. \tag{2.12}$$

$$\left. \text{for all } v_1, \ldots, v_{r+2} \in V, \text{ and some 2-p.d. } S \in \mathcal{T}_{r+2}^V \right\},$$

where the equalities in (2.10) and (2.12) follow from Lemma 2.2. Proposition 2.8 shows that our constructions generalize the hierarchies $\mathcal{C}_r^n$ (1.4) and Q_r^n (1.5).

Proposition 2.8. *For any $r \in \mathbb{N}$,*

$$\mathcal{C}_r^n = \mathcal{C}_r^{[n]}, \ Q_r^n = Q_r^{[n]}.$$

Proof. First, $\mathcal{C}_r^n = \mathcal{C}_r^{[n]}$ follows immediately from (2.8). Next, define $\mathcal{P}_\beta$ as in (2.6) and let $M \in Q_r^{[n]}$. Define $N := \sigma\left(\mathbf{Stk^r}(K)\right) - \sigma\left(S\right)$ so that $N = \sigma(N)$. Then $(e^\intercal x)^r M[x] = N[x] + S[x]$ by Lemma 2.5. Using Lemma 2.5 again, the latter equality can be rewritten as

$$\begin{aligned}(e^\intercal x)^r M[x] &= \sum_{v\in[n]^r} \mathbf{Slc}^{\boldsymbol{v}}(N)[x]\prod_{i=1}^r x_{v_i} + \sum_{v\in[n]^r} \mathbf{Slc}^{\boldsymbol{v}}(S)[x]\prod_{i=1}^r x_{v_i}\\ &= \sum_{|\beta|=r} x^\beta \sum_{v\in\mathcal{P}_\beta} \mathbf{Slc}^{\boldsymbol{v}}(N)[x] + \sum_{|\beta|=r} x^\beta \sum_{v\in\mathcal{P}_\beta} \mathbf{Slc}^{\boldsymbol{v}}(S)[x]\\ &= \sum_{|\beta|=r} x^\beta \hat{N}_\beta[x] + \sum_{|\beta|=r} x^\beta \hat{S}_\beta[x]\end{aligned}$$

where $\hat{N}_\beta \geq 0$ and $\hat{S}_\beta$ is p.d. for all possible β as sums of non-negative and positive semidefinite kernels respectively. Thus $M \in Q_r^n$.

Now let $M \in Q_r^n$. By Lemma 2.5 this implies

$$\begin{aligned}(e^\intercal x)^r M[x] = \ &\mathbf{Stk^r}(M)[x] = \sum_{|\beta|=r} x^\beta N_\beta[x] + \sum_{|\beta|=r} x^\beta S_\beta[x]\\ &= \sum_{|\beta|=r} x^\beta \sum_{v\in\mathcal{P}_\beta} \frac{1}{|\mathcal{P}_\beta|} N_\beta[x] + \sum_{|\beta|=r} x^\beta \sum_{v\in\mathcal{P}_\beta} \frac{1}{|\mathcal{P}_\beta|} S_\beta[x].\end{aligned}$$

Define $\hat{N}$ and $\hat{S}$ as follows:

$$\begin{aligned}\hat{N} :\ \mathbf{Slc}^{\boldsymbol{v}}(\hat{N}) &= \frac{1}{|\mathcal{P}_\beta|} N_\beta \ \text{ for all } \beta\in\mathbb{N}^n,\ |\beta| = r \text{ and } v\in\mathcal{P}_\beta,\\ \hat{S} :\ \mathbf{Slc}^{\boldsymbol{v}}(\hat{S}) &= \frac{1}{|\mathcal{P}_\beta|} S_\beta \ \text{ for all } \beta\in\mathbb{N}^n,\ |\beta| = r \text{ and } v\in\mathcal{P}_\beta.\end{aligned}$$

Then

$$\mathbf{Stk^r}(M)[x] = \sum_{v\in[n]^r} \mathbf{Slc}^{\boldsymbol{v}}(\hat{N})[x]\prod_{i=1}^r x_{v_i} + \sum_{v\in[n]^r} \mathbf{Slc}^{\boldsymbol{v}}(\hat{S})[x]\prod_{i=1}^r x_{v_i} = \hat{N}[x] + \hat{S}[x],$$

where $\hat{N} \in \mathcal{N}_{r+2}^{[n]}$, $S \in \mathcal{T}_{r+2}^{[n]}$, 2-p.d. The last equality follows from Lemma 2.5 *a*. Hence, by the same lemma, $\sigma(\mathbf{Stk^r}(M)) = \sigma(\hat{N}) + \sigma(\hat{S})$, and thus $\sigma(\mathbf{Stk^r}(M)) - \sigma(\hat{S}) \in \mathcal{N}_{r+2}^{[n]}$. □

Up to a diference in notation, Proposition 2.8 was also obtained by Dong [58] using a different type of a proof.

For a set U contained in a vector space V, we denote by $\operatorname{cor} U$ the algebraic interior of U following the notation by Holmes [87]:

$$\begin{aligned}\operatorname{cor} U = \{&x\in U :\ \text{for all } y\in V \text{ there exists } \varepsilon_y > 0 \text{ such that}\\ &x + \varepsilon y \in U \text{ for all } \varepsilon\in[0,\varepsilon_y]\}.\end{aligned}$$

Notice that when V is a topological vector space, then $\operatorname{cor} U$ includes the interior of U (see. e.g., Chapter 15 in [87]). Moreover, $\operatorname{cor} U$ coincides with the interior of U

for non-empty convex sets in finite-dimensional spaces, such as the sets of copositive or positive semidefinite matrices (see, for instance, Chapter 17 in [87]). Algebraic interior is an important concept in convex optimization since it determines when two convex sets can be separated by a hyperplane, see Chapter 4 in [87] for more details.

Next theorem shows that properties of $\mathcal{C}_r^{[n]}, Q_r^{[n]}$ proven in [35] and [173] respectively can be generalized for compact V.

Theorem 2.9. *Let $V \subset \mathbb{R}^n$ be a compact set. Then,*

$$\mathcal{C}_0^V \subseteq \mathcal{C}_1^V \subseteq \cdots \subseteq \mathcal{COP}(V), \quad Q_0^V \subseteq Q_1^V \subseteq \cdots \subseteq \mathcal{COP}(V),$$

and $\operatorname{cor}\mathcal{COP}(V) \subseteq \bigcup_r \mathcal{C}_r^V \subseteq \bigcup_r Q_r^V$.

The key ingredient in the proof of Theorem 2.9 is the characterization of the algebraic interior of the copositive cone given in Proposition 2.10. When V is finite (i.e. for matrices), the algebraic interior of the copositive cone equals its interior and consists of those copositive matrices whose quadratic form is strictly positive on the standard simplex $\Delta^n := \{x \in \mathbb{R}^n : \; e^\intercal x = 1, \; x \geq 0\}$. This implies, by compactness of the simplex, that a matrix $M \in \operatorname{cor}\mathcal{COP}(V)$ if and only if there is $\epsilon > 0$ such that $x^T M x = \sum_{v\in V}\sum_{u\in V} M(v,u)x_i x_j \geq \epsilon$ for all $x \in \Delta^n$. Proposition 2.10 shows that for compact V the latter is true uniformly over all finite submatrices of a given p.d. kernel K.

Proposition 2.10. *Let $V \subset R^n$ be a compact set. Then*

$$\operatorname{cor}\mathcal{COP}(V) = \Big\{K \in \mathcal{K}(V) : \text{there is } \varepsilon > 0 \text{ such that for all } n > 0 \text{ and all } v_1,\ldots,v_n \in V,\; x \in \Delta^n \colon \sum_{i=1}^n \sum_{j=1}^n K(v_i,v_j)x_i x_j \geq \epsilon\Big\} \tag{2.13}$$

Proof. Let K, ε be such that (2.13) holds. Let $\hat{K} \in \mathcal{K}(V)$ be given. Since $\hat{K}$ is continuous and $V \times V$ is compact, $\hat{K}$ attains its maximum on $V \times V$. Let $\hat{K}^* = \max_{x,y\in V} \hat{K}(x,y)$. Then for any $v_1,\ldots,v_n \in V$ and $x \in \Delta^n$,

$$\begin{aligned}
&\sum_{i=1}^n \sum_{j=1}^n \left(K(v_i,v_j) - \frac{\varepsilon}{\hat{K}^*}\hat{K}(v_i,v_j)\right) x_i x_j \\
&\qquad \geq \varepsilon - \frac{\varepsilon}{\hat{K}^*}\left(\max_{i,j\in[n]} \hat{K}(v_i,v_j) \sum_{i=1}^n \sum_{j=1}^n x_i x_j\right) \\
&\qquad \geq 0.
\end{aligned}$$

Hence $K - \frac{\varepsilon}{\hat{K}^*}\hat{K} \in \mathcal{COP}(V)$ by definition, and thus $K \in \operatorname{cor}\mathcal{COP}(V)$. Now let $K \in \operatorname{cor}\mathcal{COP}(V)$, and denote by $\mathbb{J}$ the kernel on V that is always equal to one.

Then there is $\varepsilon > 0$ such that $K - \varepsilon\mathbb{J} \in \mathcal{COP}(V)$. Therefore for any choice of $v_1, \dots, v_n \in V$,

$$\begin{aligned}\min_{x\in\Delta^n} &\sum_{i=1}^{n}\sum_{j=1}^{n} K(v_i, v_j)x_i x_j \\ &= \min_{x\in\Delta^n} \sum_{i=1}^{n}\sum_{j=1}^{n} (K - \varepsilon\mathbb{J})(v_i, v_j)x_i x_j + \varepsilon \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \\ &\geq \varepsilon.\end{aligned}$$

□

To prove Theorem 2.9, we use several additional results. A result by Powers and Reznick [182] on the rate of convergence in Pólya's theorem (see, e.g., [83]), a characterization of $\mathcal{C}_r^V$ in terms of $\mathcal{C}_r^U$ for all finite $U \subset V$, and the fact that if $K \in \mathcal{Q}_r^V$, then for every finite $U \subset V$ we have $K^U \in \mathcal{Q}_r^U$.

Lemma 2.11. *Let $V \subset \mathbb{R}^n$ be a compact set, and let $r \in \mathbb{N}$. Then $K \in \mathcal{C}_r^V$ if and only if $K^U \in \mathcal{C}_r^U$ for every finite $U \subset V$.*

Proof. Let $U \subset V$ be finite. If $K \in \mathcal{C}_r^V$, then $\sigma(\mathbf{Stk^r}(K)) \in \mathcal{N}_{r+2}^V$, and thus $\sigma\left(\mathbf{Stk^r}(K^U)\right) \geq 0$, that is $K^U \in \mathcal{C}_r^U$. On the other hand, if $K^U \in \mathcal{C}_r^U$ for each finite $U \subset V$, then for any $v_1, \dots, v_{r+2} \in V$ we have $\sigma\left(\mathbf{Stk^r}(K^{\{v_1,\dots,v_{r+2}\}})\right)(v_1 \dots, v_{r+2}) = \sigma(\mathbf{Stk^r}(K))(v_1 \dots, v_{r+2}) \geq 0$. Hence $K \in \mathcal{C}_r^V$. □

Lemma 2.12. *Let $V \subset \mathbb{R}^n$ be a compact set, and let $r \in \mathbb{N}$. Then $K \in \mathcal{Q}_r^V$ implies that $K^U \in \mathcal{Q}_r^U$ for every finite $U \subset V$.*

Proof. Let $U \subset V$ be finite. If $K \in \mathcal{C}_r^V$, then there is a 2-p.d. function S such that $\sigma(\mathbf{Stk^r}(K)) - \sigma(S) \in \mathcal{N}_{r+2}^V$. Hence $\sigma\left(\mathbf{Stk^r}(K^U)\right) - \sigma(S^U) \in \mathcal{N}_{r+2}^U$, that is $K^U \in \mathcal{Q}_r^U$. □

We would like to emphasize the difference between the hierarchies $\mathcal{C}_r^V$ and $\mathcal{Q}_r^V$. Namely, while we can prove in Lemma 2.11 that if $K^U \in \mathcal{C}_r^U$ for every finite $U \subset V$, then $K \in \mathcal{C}_r^V$, we cannot show the analog of this statement for $\mathcal{Q}_r^V$. The first reason for this difference is that we work with non-symmetric tensors and use the symmetrization operator σ. Moreover, let $K^U \in \mathcal{Q}_r^U$ for every finite $U \subset V$. Let $U \subset V$ be finite. Then there exists a 2-p.d. tensor $S^U \in \mathcal{T}_{r+2}^U$ and a tensor $N^U \in \mathcal{N}_{r+2}^U$ such that $\sigma\left(\mathbf{Stk^r}(K^U)\right) - \sigma(S^U) = N^U$. However, we cannot claim that S^U or N^U change continuously with U.

The final result we need for the proof of Theorem 2.9 is as follows.

Theorem 2.13 (Powers and Reznick [182])**.** *Let $M \in \mathbb{S}^n$ be strictly copositive. Then the polynomial $(e^\intercal x)^r \sum_{i,j=1}^n M_{ij} x_i x_j$ has only positive coefficients if $r > \frac{L}{k} - 2$, where $L = \max_{ij} |M_{ij}|$ and $k = \min_{x \in \Delta^n} x^\intercal M x$.*

Notice that the result from Theorem 2.13 is *a certificate of copositivity* of M. That is, the expression that makes the copositivity of M evident. This certificate is a direct consequence of Polya's Positivstellensatz 1.8. Theorem 2.13 strengthens Polya's Positivstellensatz in the sense that it provides a bound on the number r.

Remark 2.14. *The bound on r in Theorem 2.13 does not depend on the size of M.*

Now we are ready to prove Theorem 2.9.

Proof of Theorem 2.9. Let $r \geq 0$ and let $K \in Q_r^V$. Then by (2.11) there exists a 2-p.d. $S \in \mathcal{T}_{r+2}^V$ such that $\sigma\left(\mathbf{Stk^r}(K)\right) \geq \sigma\left(S\right)$. First, we show that $K \in Q_{r+1}^V$. Define $N := \sigma\left(\mathbf{Stk^r}(K)\right) - \sigma\left(S\right)$ so that $N = \sigma(N)$. Then by *c.* and *d.* in Lemma 2.3,

$$\sigma\left(\mathbf{Stk^{r+1}}(K)\right) = \sigma\left(\mathbf{Stk}^1(N)\right) + \sigma\left(\mathbf{Stk}^1(S)\right).$$

From the definition (2.1) of the stack operator, $\mathbf{Stk}^1(S) \in \mathcal{T}_{r+3}^V$ is 2-p.d. and $\sigma\left(\mathbf{Stk}^1(N)\right) \in \mathcal{N}_{r+3}^V$. Thus $K \in Q_{r+1}^V$. Analogously, $\mathcal{C}_r^V \subseteq \mathcal{C}_{r+1}^V$.

The fact that $K \in \mathcal{COP}(V)$ follows by Lemma 2.12 and Proposition 2.8 since for any finite $U \subset V$ we have that $K^U \in Q_r^U \in \mathcal{COP}(U)$. As $\mathcal{C}_r^V \subseteq Q_r^V$ by construction of $\mathcal{C}_r^V$ (2.9) and Q_r^V (2.11), we also have $\mathcal{C}_r^V \subseteq \mathcal{COP}(V)$.

For the final part of the proof, let $K \in \operatorname{cor} \mathcal{COP}(V)$. Since K is continuous and $V \times V$ is compact, K attains its maximum and minimum values on $V \times V$.

Denote

$$L = \max_{x,y \in V} |K(x,y)|.$$

As $K \in \operatorname{cor} \mathcal{COP}(V)$, by Proposition 2.10 there is $\varepsilon > 0$ such that

$$\min_{x \in \Delta^{|U|}} x^\intercal K^U x \geq \varepsilon,$$

for all finite $U \subseteq V$.

Let $r > \frac{L}{\varepsilon} - 2$. By Theorem 2.13, $K^U \in \mathcal{C}_r^U$ for any $U \subseteq V$, which implies $K \in \mathcal{C}_r^V$ by Lemma 2.11. Thus, $\operatorname{cor} \mathcal{COP}(V) \subseteq \mathcal{C}_r^V \subseteq Q_r^V$ □

2.3.1 Approximating the stability number of infinite graphs

The stability number of a graph is the largest number of vertices such that no two of them are adjacent. In this subsection we introduce the *copositive* formulation of the stability number problem on infinite graphs by Dobre et al. [54]. We also prove that if the copositive formulation is strictly feasible, then the stability number can be approximated as closely as desired by replacing $\mathcal{COP}(V)$ with $\mathcal{C}_r^V$ or Q_r^V with r big enough.

Following de Laat and Vallentin [39], we define a compact topological packing graph as the graph where the vertex set is a compact *Hausdorff* topological space, and every finite clique is contained in an open clique. A topological space V is Hausdorff if every two distinct points in V have disjoint neighborhoods. In the sequel we use the property that a *compact* Hausdorff topological space is *normal*; that is, every two disjoint closed sets in it have disjoint open neighborhoods. The stability number of compact topological packing graphs is finite since every vertex of such graph is a clique and thus is contained in an open clique. This implies that any $U \subseteq V$ has a cover of open cliques. By compactness of U, this cover has a finite subcover. If U is infinite, then some members of U belong to the same clique in this subcover, and thus U cannot be a stable set.

The unit sphere S^{n-1} with the usual topology is a compact Hausdorff topological space. An example of a compact topological packing graph in this space is the graph $\mathcal{G}_n^\theta = (S^{n-1}, E^{\mathcal{G}_n^\theta})$ in which $(u,v) \in E^{\mathcal{G}_n^\theta}$ if and only if $u^\intercal v \in (\cos\theta, 1)$. That is, there is an edge between every two vertices when the angle between them is strictly smaller than θ. Notice that if U is a finite clique, by definition of $E^{\mathcal{G}_n^\theta}$ there is an open spherical cap that contains U and forms a clique in $E^{\mathcal{G}_n^\theta}$. Therefore $\mathcal{G}_n^\theta$ is a compact topological packing graph. An example of a graph that is not a compact topological packing graph is the graph $\mathcal{H}_n^\theta = (S^{n-1}, E^{\mathcal{H}_n^\theta})$ in which there is an edge between two vertices when the angle between them is equal to θ. Any open subset of S^{n-1} has points with a distance less than θ. Hence no open subset can be a clique in $\mathcal{H}_n^\theta$, and all cliques must be finite.

Theorem 2.15 (Theorem 1.2. from Dobre et al. [54]). *Let $G = (V, E)$ be a compact topological packing graph. Then the stability number of G equals*

$$\begin{aligned} \alpha(G) = \inf_{K \in \mathcal{K}(V),\ \lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & K(v,v) = \lambda - 1 && \textit{for all } v \in V, \\ & K(u,v) = -1 && \textit{for all } (u,v) \notin E,\ u \neq v, \\ & K \in \mathcal{COP}(V). \end{aligned} \tag{2.14}$$

An example of the stability number problem on a compact topological graph is the spherical codes problem. In the spherical codes problem, the number of points on the unit sphere in $\mathbb{R}^n$ for which the pairwise angular distance is not smaller than some value θ is maximized. This problem can be viewed as the stable set problem on the graph $\mathcal{G}_n^\theta = (S^{n-1}, E^{\mathcal{G}_n^\theta})$ introduced in the previous paragraph. A particular case of the spherical codes problem when $\theta = \frac{\pi}{3}$ is the kissing number problem, which we analyze in detail later in Section 2.4.

The equality sign in the constraint "$K(u,v) = -1$ for all $(u,v) \notin E,\ u \neq v$" of problem (2.14) and relaxations (2.16), (2.17) can be replaced by "$\leq$" without loss of generality. In the sequel we call two problems equivalent if from every feasible solution to one problem one can construct a feasible solution to the other with the same objective value.

Lemma 2.16. *Problem* (2.14) *is equivalent to the following problem:*

$$\begin{aligned} \alpha(G) = \inf_{K \in \mathcal{K}(V),\ \lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & K(v,v) = \lambda - 1 && \textit{for all } v \in V, \\ & K(u,v) \leq -1 && \textit{for all } (u,v) \notin E,\ u \neq v, \\ & K \in \mathcal{COP}(V). \end{aligned} \tag{2.15}$$

Proof. If the solution (K, λ) is feasible for (2.14), then it is clearly feasible for (2.15), and the objective values of the two problems coincide for this solution. Now, let (K, λ) be feasible for problem (2.15). Define $X := \{(v,v) : v \in V\}$ and $Y := \{(u,v) : u, v \in V,\ (u,v) \notin E,\ u \neq v\}$. Every $(v,v) \in X$ belongs to $C_v \times C_v$ where $C_v \subset V$ is an open clique. Since $C_v \times C_v \cap Y = \emptyset$, the sets X and Y are disjoint. As G is a compact topological packing graph, X and Y are closed. In particular, X is closed since V is closed, and Y is closed since its complement is open. Since V is Hausdorff and compact, it is normal. Hence by Urysohn's lemma (see, e.g., 15.6 in [222]), there is a continuous function $f : V \times V \to [0,1]$ such that $f(v,v) = 0$ for all $v \in V$ and $f(u,v) = 1$ for all $(u,v) \notin E,\ u \neq v$. By construction, if $(x,y) \in Y$, then $(y,x) \in Y$. Hence we can define a kernel

$$\hat{f}(x,y) = \max\Big\{ -\tfrac{1}{2}\Big(f(x,y) + f(y,x)\Big), K(x,y)\Big\} \text{ for all } x, y \in V.$$

We have $\hat{f} \geq K$, and thus $K \in \mathcal{COP}(V)$ implies $\hat{f} \in \mathcal{COP}(V)$. Since $\lambda \geq \alpha(G)$, we have that $\lambda \geq 1$ and $\hat{f}(v,v) = \lambda - 1$ for all $v \in V$. Therefore $(\hat{f}, \lambda)$ is feasible for problem (2.14), and the objectives of the two problems coincide for this solution. $\square$

Define the following relaxations to problem (2.15) (and thus to problem (2.14) by

Lemma 2.16):

$$\begin{aligned}\gamma_r(G) = \inf\ & \lambda \\ \text{s.t. } & K(v,v) = \lambda - 1 && \text{for all } v \in V, \\ & K(u,v) \le -1 && \text{for all } (u,v) \notin E,\ u \ne v, \\ & K \in \mathcal{C}_r^V. \end{aligned} \tag{2.16}$$

$$\begin{aligned}\nu_r(G) = \inf\ & \lambda \\ \text{s.t. } & K(v,v) = \lambda - 1 && \text{for all } v \in V, \\ & K(u,v) \le -1 && \text{for all } (u,v) \notin E,\ u \ne v, \\ & K \in Q_r^V. \end{aligned} \tag{2.17}$$

By Theorem 2.9, $\mathcal{C}_r^V \subseteq Q_r^V \subseteq \mathcal{COP}(V)$ for any r. Hence,

$$\alpha(G) \le \nu_r(G) \le \gamma_r(G). \tag{2.18}$$

In the next theorem and further in this chapter we say that problem (2.15) is *strictly feasible* if there exists feasible (K^+, λ^+) such that $K^+ \in \operatorname{cor} \mathcal{COP}(V)$.

Theorem 2.17. *Let $G = (V, E)$ be a compact topological packing graph. Assume problem* (2.15) *is strictly feasible. Then $\gamma_r(G) \downarrow \alpha(G)$ and $\nu_r(G) \downarrow \alpha(G)$.*

Proof. Let (K^+, λ^+) be a feasible solution to problem (2.15) such that $K^+ \in \operatorname{cor} \mathcal{COP}(V)$. From the definition of infimum, for any $n > 0$ there is (K_n, λ_n) feasible for (2.15) such that $\lambda_n \le \alpha(G) + \frac{1}{n}$. Let

$$K_n^+ = \tfrac{1}{n} K^+ + \left(1 - \tfrac{1}{n}\right) K_n \quad \text{and} \quad \lambda_n^+ = \tfrac{1}{n}\lambda^+ + \left(1 - \tfrac{1}{n}\right)\lambda_n.$$

By convexity of (2.15) the pair (K_n^+, λ_n^+) is feasible for problem (2.15). Since $K^+ \in \operatorname{cor} \mathcal{COP}(V)$, by Theorem 2.9 there is $m > 0$ such that $K_n^+ \in \mathcal{C}_r^V$ for all $r \ge m$. Hence,

$$\lim_{r \to \infty} \gamma_r(G) \le \lambda_n^+ = \tfrac{1}{n}\lambda^+ + \left(1 - \tfrac{1}{n}\right)\lambda_n \le \tfrac{1}{n}\lambda^+ + \left(1 - \tfrac{1}{n}\right)\left(\alpha(G) + \tfrac{1}{n}\right).$$

Taking the limit when $n \to \infty$ on both sides and using the second inequality in (2.18) we obtain

$$\alpha(G) \ge \lim_{r \to \infty} \gamma_r(G) \ge \lim_{r \to \infty} \nu_r(G).$$

The first inequality in (2.18) concludes the proof. □

Now, we show some graphs for which the conditions of Theorem 2.17 are satisfied. We use the following result by Motzkin and Straus [146]:

Theorem 2.18 (Motzkin and Straus [146]). *Let $G = (V, E)$ be a finite graph with adjacency matrix A and stability number α. Then*

$$\frac{1}{\alpha} = \min_{x \in \Delta^{|V|}} x^{\mathsf{T}}(A + I)x.$$

Proposition 2.19. *let $G = (V, E)$ be a compact topological packing graph that satisfies the following conditions:*

(a) every $v \in V$ has an open neighborhood whose closure is a clique;

(b) every closed clique is contained in an open clique.

Then problem (2.15) *is strictly feasible, and thus $\gamma_r(G) \downarrow \alpha(G)$ and $\nu_r(G) \downarrow \alpha(G)$.*

Proof. Since V is compact and every vertex is a clique, V can be covered by a finite number m of open cliques. Moreover, by (a) each of these cliques' closure is a clique too. Denote these closures by $C_1, \ldots, C_m$. By (b), each $C_j,\ j \in [m]$ is a subset of an open clique $B_j,\ j \in [m]$. Define

$$X := \bigcup_{j \in [m]} C_j \times C_j;\ Y := \bigcup_{j \in [m]} B_j \times B_j.$$

Then X is closed, Y is open, and $X \subset Y$. Moreover, $\{(v, v) : v \in V\} \subset X$. We have that $V \setminus Y$ and X are disjoint subsets of a compact Hausdorff space V. Hence, by Urysohn's lemma (see, e.g., 15.6 in [222]), there is a continuous function $g : V \times V \to [0, 1]$ such that $g(u, v) = 1$ for $(u, v) \in X$ and $g(u, v) = 0$ for $(u, v) \in V \setminus Y$. Let $\hat{G} = (V, \hat{E})$ be the subgraph of G such that $(u, v) \in \hat{E}$ if $(u, v) \in X$. Then $\hat{G}$ has a finite stability number since $C_1, \ldots, C_m$ are cliques in $\hat{G}$ that cover V. Moreover, $\alpha(\hat{G}) \geq \alpha(G)$ since every edge in $\hat{G}$ is an edge in G too.

Define a kernel $K(u, v) = 2\alpha(\hat{G})g(u, v) - 1$ for all $K \in V$. We claim that $(K, 2\alpha(\hat{G}))$ is strictly feasible for problem (2.15) with G. By definition, $K(u, u) = 2\alpha(\hat{G}) - 1$ and $K(u, v) = -1$ for $(u, v) \notin E$. To show that $K \in \operatorname{cor} \mathcal{COP}(V)$, fix $k > 0$ and $u_1, \ldots, u_k \in V$. Let $\hat{G}^u = (\{u_1, \ldots, u_k\}, E^u)$ be the graph defined by $(i, j) \in E^u$ if and only if $(u_i, u_j) \in X$. Notice that $\alpha(\hat{G}) \geq \alpha(\hat{G}^u)$. Let A^u be the adjacency matrix of $\hat{G}^u$. By definition of g we have

$$g(u_i, u_j) \geq A^u_{ij} + I_{ij} \text{ for all } i, j \leq k \tag{2.19}$$

Let $x \in \Delta^k$, then

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{j=1}^{k} K(u_i,u_j)x_ix_j &= 2\alpha(\hat{G})\sum_{i=1}^{k}\sum_{j=1}^{k} g(u_i,u_j)x_ix_j - \sum_{i=1}^{k}\sum_{j=1}^{k} x_ix_j \\
&\geq 2\alpha(\hat{G}) \min_{x\in\Delta^k} x^\intercal(A^u + I)x - 1 \qquad \text{(by (2.19))} \\
&= \frac{2\alpha(\hat{G})}{\alpha(\hat{G}^u)} - 1 \qquad \text{(by Theorem 2.18)} \\
&\geq \frac{2\alpha(\hat{G})}{\alpha(\hat{G})} - 1 \\
&= 1
\end{aligned}$$

Therefore $K \in \operatorname{cor}\mathcal{COP}(V)$ by Proposition 2.10. Hence $\gamma_r(G) \downarrow \alpha(G)$ and $\nu_r(G) \downarrow \alpha(G)$ by Theorem 2.17. □

The example from Proposition 2.19 shows that the algebraic interior of $\mathcal{COP}(S^{n-1})$ is non-empty. Notice that this is in contrast to $\operatorname{cor}\mathcal{PSD}(V)$, which turns out to be empty.

Proposition 2.20. *Let $V \subset \mathbb{R}^n$ be compact. Then* $\operatorname{cor}\mathcal{PSD}(V) = \emptyset$.

Proof. Let $K \in \mathcal{PSD}(V)$. By Mercer's theorem [142], there exists a sequence $(\lambda_i)_{i\in N_+}$, $\lambda_i \in \mathbb{R}_+$ for all $i \in N_+$, and a sequence of orthonormal functions $(e_i)_{i\in N_+}$, $e_i \in C(V)$ for all $i \in N_+$ such that for all $x, y \in V$

$$K(x,y) = \sum_{i\in\mathbb{N}_+} \lambda_i e_i(x)e_i(y).$$

This also implies, for any probability measure μ strictly positive on open subsets of V, that

$$\int_{x\in V} K(x,x)d\mu(x) = \sum_{i\in\mathbb{N}_+} \lambda_i < \infty,$$

and thus $\lim_{i\to\infty}\lambda_i = 0$. If $\lambda_i = 0$ for some $i \in \mathbb{N}_+$, it is straightforward that $K \notin \operatorname{cor}\mathcal{PSD}(V)$ by orthogonality of $(e_i)_{i\in N_+}$, using Theorem 1.3. Hence we assume that $\lambda_i > 0$ for all $i \in N_+$. Since $\lim_{i\to\infty}\lambda_i = 0$, there is an index set $J \subseteq \mathbb{N}_+$ and a subsequence $(\lambda_j)_{j\in J}$ of $(\lambda_i)_{i\in N_+}$ such that $\frac{\lambda_{j+1}}{\lambda_j} \leq \frac{1}{2}$ for all $j \in J$. Then $\frac{\sqrt{\lambda_{j+1}}}{\sqrt{\lambda_j}} \leq \frac{1}{\sqrt{2}} < 1$ for all $j \in J$, and hence the series $\sum_{j\in J}\lambda_j$ converges by the ratio test. For every $i \in N_+$ define $\hat{\lambda}_i = \sqrt{\lambda_i}$ if $i \in J$ and $\hat{\lambda}_i = 0$ otherwise. The convergence of $\sum_{j\in J}\lambda_j$ and the fact that $(e_i)_{i\in N_+}$ have norm one and are uniformly continuous imply that $\sum_{i\in\mathbb{N}_+} \hat{\lambda}_i e_i(x)e_i(y)$ is a uniformly convergent series of uniformly continuous functions. Hence $\hat{K}(x,y) = \sum_{i\in\mathbb{N}_+} \hat{\lambda}_i e_i(x)e_i(y)$ is a kernel. For all $x, y \in V$ we have

$$K(x,y) - \varepsilon\hat{K}(x,y) = \sum_{i\in\mathbb{N}_+} (\lambda_i - \varepsilon\hat{\lambda}_i)e_i(x)e_i(y).$$

Therefore for any $\varepsilon > 0$ there is $L > 0$ such that for all $i > L$ such that $i \in J$, we have $\frac{\lambda_i}{\hat{\lambda}_i} = \sqrt{\lambda_i} < \varepsilon$. Hence $\lambda_i - \varepsilon\hat{\lambda}_i < 0$ for all $i > L$ such that $i \in J$. Using Theorem 1.3 and orthogonality of $(e_i)_{i\in N_+}$ again, we see that $K - \varepsilon\hat{K} \notin \mathcal{PSD}(V)$ for all $\varepsilon > 0$. Therefore $K \notin \operatorname{cor}\mathcal{PSD}(V)$ and $\operatorname{cor}\mathcal{PSD}(V) = \emptyset$. □

2.4 Bounds on the kissing number problem

The kissing number is the largest number κ_n of non-overlapping unit spheres in $\mathbb{R}^n$ that can simultaneously touch other unit spheres. History of the problem is described in detail in, for instance, Musin [150]. The value of κ_n is known for $n = 1, 2, 3, 4, 8, 24$. Computing $\kappa_1 = 2$ and $\kappa_2 = 6$ is straightforward, but this is not the case for κ_n with $n > 2$. The question of whether $\kappa_3 = 12$ or $\kappa_3 = 13$ is attributed to the famous discussion between Isaac Newton and David Gregory in 1694, and the result $\kappa_3 = 12$ was proven only in 1953 by Schütte and van der Waerden [205]. The numbers $\kappa_8 = 240$ and $\kappa_{24} = 196560$ were found by Odlyzko and Sloane [164] and Levenshtein [125] in 1979. Finally, $\kappa_4 = 24$ was proven in 2003 by Musin [150]. A lot of research has been done to approximate the kissing numbers in other dimensions from above and below. In this chapter we are interested in obtaining upper bounds using the hierarchies proposed in Section 2.3.

In 1977, Delsarte, Goethals and Seidel [44] proposed an LP upper bound used later to obtain κ_8 and κ_{24}. To find this bound, one has to solve an infinite-dimensional LP, which can be approximated by a semidefinite program (SDP). The LP bound is not tight in general, for example, for κ_4 it cannot be less than 25, as was shown by Arestov and Babenko [7] in 1997. Musin [150] strengthened the LP bound for the four-dimensional case to obtain $\kappa_4 = 24$. In 2007, Pfender [177] strengthened the LP bound and obtained the best existing upper bounds for $n = 25$ and $n = 26$. In 2008, a new SDP upper bound was proposed by Bachoc and Vallentin [12]. Mittelman and Vallentin [143] used this approach to compute upper bounds for $5 \leq n \leq 23$. Later these results were strengthened for dimensions 9 to 23 by Machado and de Oliveira Filho [136] who exploited the symmetry of the SDP problem. Table 2.1 shows best-known bounds on the kissing numbers in some dimensions, bounds for other dimensions can be found in [143] and [136]. Finally, Musin [151] proposed a hierarchy generalizing the earlier mentioned linear and SDP approaches, but this hierarchy has not been implemented and is not proven to converge.

Table 2.1 – Best-known upper and lower bounds on the kissing number.

n	3	4	5	6	7	8	9	10	11	12	13	24	25	26
upper bound	12 [205]	24 [150]	44 [143]	78 [12]	134 [143]	240 [125, 164]	363 [136]	553 [136]	869 [136]	1356 [136]	2066 [136]	196560 [125, 164]	278083 [177]	396447 [177]
lower bound [154]	12	24	40	72	126	240	306	500	582	840	1154	196560	197040	198480
$\frac{ub-lb}{lb}$	0	0	10	0.08	0.06	0	0.19	0.11	0.49	0.61	0.79	0	0.41	0.997

The kissing number κ_n can be reformulated as the stability number on a graph whose vertex set is the unit sphere:

Definition 2.21. *$\mathcal{G}_n = (S^{n-1}, E)$ is the graph with the edge set*

$$E = \left\{(u, v) \in S^{n-1} \times S^{n-1} : u^\mathsf{T} v > \tfrac{1}{2}\right\}.$$

We have $\alpha(\mathcal{G}_n) = \kappa_n$, and thus the kissing number could be computed using (2.15), and approximations $\gamma(\mathcal{G}_n)$ (2.16), $\nu(\mathcal{G}_n)$ (2.17) could be used to find upper bounds on κ_n.

2.4.1 Convergence of the hierarchies for the kissing number

Let $n \in \mathbb{N}$, $n \geq 2$. In this subsection we show that the hierarchies proposed in Section 2.3 provide converging upper bounds on the kissing number κ_n. In view of Theorem 2.17, we only need to show that there exists a kernel $K \in \operatorname{cor}\mathcal{COP}(S^{n-1})$ feasible for problem (2.15) when $G = \mathcal{G}_n$.

Corollary 2.22. *For the graph $\mathcal{G}_n$ we have $\gamma_r(\mathcal{G}_n) \downarrow \kappa_n$ and $\nu_r(\mathcal{G}_n) \downarrow \kappa_n$.*

Proof. The result follows from Proposition 2.19 since $\mathcal{G}_n$ satisfies the conditions from this proposition. □

Our next goal is to compute upper bounds on κ_n solving problems (2.16) and (2.17), exploiting the symmetry of the graph $\mathcal{G}_n$. Proposition 2.23 below shows that, similarly to the finite case, γ_r is never tight for κ_n and γ_r provides only trivial bounds for small r. Thus we further restrict our attention to $\nu_r(\mathcal{G}_n)$.

Proposition 2.23. *Let $n > 0, r \geq 0$. Let $\gamma_r(\mathcal{G}_n)$ be the optimal value of problem* (2.16) *with $G = \mathcal{G}_n$. Then $\gamma_r(\mathcal{G}_n) > \kappa_n$ for all $r \in \mathbb{N}$, and if $\gamma_r(\mathcal{G}_n) < \infty$, then $r \geq \kappa_n - 1$.*

Proof. If problem (2.16) with $G = \mathcal{G}_n$ is infeasible, then $\gamma_r(\mathcal{G}_n) = \infty > \kappa_n$. So we assume feasibility. Let (K, λ) be a feasible solution for problem (2.16) with $G = \mathcal{G}_n$.

Let $U = \{u_1, \ldots, u_{\alpha(\mathcal{G}_n)}\} \subset S^{n-1}$ be a maximum stable set of $\mathcal{G}_n$, and denote by $\mathcal{G}_n^U$ the subgraph of $\mathcal{G}_n$ induced by U. We have then,

$$\gamma_r(\mathcal{G}_n^U) \leq \gamma_r(\mathcal{G}_n) < \infty. \tag{2.20}$$

Therefore by Theorem 4.2 from [35], $r \geq \alpha(\mathcal{G}_n^U) - 1 = \alpha(\mathcal{G}_n) - 1 = \kappa_n - 1$. Moreover, by Lemma 2.11 (K^U, λ) is a feasible solution to problem(2.16) with $G = \mathcal{G}_n^U$, and Corollary 2 in [173] implies

$$\kappa_n < \gamma_r(\mathcal{G}_n^U). \tag{2.21}$$

Using (2.20) we obtain $\kappa_n < \gamma_r(\mathcal{G}_n)$. □

Remark 2.24. *In the proof of Proposition 2.23 we use the results from [35] and [173]. These papers consider problem* (2.16) *where the inequality sign in the constraint "$K(u, v) \leq -1$ for all $(u, v) \notin E,\ u \neq v$" is replaced by the equality sign. Using either sign results in equivalent problems. This follows from the fact that the sum of a copositive kernel and a non-negative kernel is copositive, applying the approach from the proof of Lemma 2.16 to problems* (2.16) *and* (2.17).

2.4.2 Using the symmetry of the sphere to simplify the problem

To implement the bound $\nu_r(\mathcal{G}_n)$, we exploit convexity of problem (2.17) and invariance of $\mathcal{G}_n$ under O_n similarly to [37, 53, 70]. Hence we only need to characterize the subset of $Q_r^{S^{n-1}}$ invariant under the action of O_n, which we denote by $\left(Q_r^{S^{n-1}}\right)^{O_n}$.

Let μ_H be *the Haar measure* on O_n. Considering O_n as a topological subspace of $R^{n\times n}$ makes O_n a topological group; that is, taking products and inverses are continuous operations. Let Σ^{O_n} be the Borel σ-algebra on O_n. Then we define μ_H as the unique Borel probability measure on O_n such that $\mu_H(U) = \mu_H(U^P) = \mu_H({}^P U)$ for all $P \in O_n$ and $U \in \Sigma^{O_n}$ (the notation U^P is explained in Definition 1.4).

Proposition 2.25. *Problem* (2.17) *for $\mathcal{G}_n$ is equivalent to the following problem:*

$$\begin{aligned} \nu_r(\mathcal{G}_n) = \inf_{K,\lambda\in\mathbb{R}}\ & \lambda \\ \text{s.t.}\ & K(x, x) = \lambda - 1, && \text{for all } x \in S^{n-1}, \\ & K(x, y) \leq -1, && \text{for all } x, y \text{ with } x^\intercal y \in [-1, \tfrac{1}{2}], \\ & K \in \left(Q_r^{S^{n-1}}\right)^{O_n}. \end{aligned} \tag{2.22}$$

Proof. For any $d > 0$ and continuous function $F : \mathbb{R}^d \to \mathbb{R}$, define

$$F^{O_n}(x_1, \ldots, x_d) := \int_{P\in O_n} F(Px_1, \ldots, Px_d) d\mu_H(P).$$

We claim that if K is feasible for problem (2.17), then so is K^{O_n}, and the objective values for both kernels are the same. Therefore, by convexity of problem (2.17), we can reduce ourselves to the solutions that are invariant under the action of O_n, which leads us to problem (2.22).

Next, we prove the claim. Let (K, λ) be feasible for problem (2.17). Clearly, $K^{O_n}(x,x) = K(x,x) = \lambda - 1$ for all $x \in S^{n-1}$ and $K(x,y) \leq -1$ for all $x, y \in S^{n-1}$ such that $x^{\intercal}y \in [-1, \frac{1}{2}]$. It is left to show that $K^{O_n} \in Q_r^{S^{n-1}}$. We have $\sigma\left(\mathbf{Stk^r}(K)\right) - \sigma(S) \in \mathcal{N}_{r+2}^{V}$ for some $S \in \mathcal{T}_{r+2}^{V}$ that is 2-p.d. Therefore

$$\begin{aligned}
\sigma\left(\mathbf{Stk^r}(K^{O_n})\right)(x_1, \ldots, x_{r+2}) &= \int_{P \in O_n} \sigma\left(\mathbf{Stk^r}(K)\right)(Px_1, \ldots, Px_{r+2}) d\mu_H(P) \\
&= \left(\sigma\left(\mathbf{Stk^r}(K)\right)\right)^{O_n}(x_1, \ldots, x_{r+2}) \\
&\geq \left(\sigma(S)\right)^{O_n}(x_1, \ldots, x_{r+2}) \\
&= \sigma\left(S^{O_n}\right)(x_1, \ldots, x_{r+2}) \\
&= \sigma(\hat{S})(x_1, \ldots, x_{r+2}),
\end{aligned}$$

where $\hat{S} = S^{O_n}$ is 2 p.d. since for every $y_1, \ldots, y_r \in S^{n-1}$ we have

$$\hat{S}(\cdot, \cdot, y_1 \ldots, y_r) = S^{O_n}(\cdot, \cdot, y_1 \ldots, y_r) = \int_{P \in O_n} S(\cdot, \cdot, Py_1 \ldots, Py_r) d\mu_H(P),$$

which is p.d. since $S(\cdot, \cdot, x_1 \ldots, x_r) \in \mathcal{PSD}\left(S^{n-1}\right)$ for any $x_1, \ldots, x_r \in S^{n-1}$. □

The condition $K \in \left(Q_r^{S^{n-1}}\right)^{O_n}$ cannot be implemented directly and requires further simplification, which we do in two steps. First, we reduce the number of variables in the problem. Namely, instead of $n(r+2)-$variate functions we further work with $\binom{r+2}{2}-$variate functions. Each argument of such a function corresponds to an inner product between a pair of variables $x, y, z_1, \ldots, z_r \in S^{n-1}$. As a result, the number of variables does not grow with the dimension of the sphere. Second, we characterize $\left(Q_r^{S^{n-1}}\right)^{O_n}$ in terms of these new variables. The second step is challenging, so we perform it in the next Chapter 3 of the thesis.

Remark 2.26. *From here on in this chapter and in Chapter 3 we use the notation* $X = [x_1, \ldots, x_d] \in (S^{n-1})^d$ *to denote the matrix* X *which has the vectors* $x_1, \ldots, x_d$ *as its columns.*

We say that $X = [x_1, \ldots, x_d] \in (S^{n-1})^d$ if $x_1 \in S^{n-1}, \ldots, x_d \in S^{n-1}$. Let

$$O_n(X) = \left\{Z = (z_1, \ldots, z_d) \in (S^{n-1})^d : \; X^{\intercal}X = Z^{\intercal}Z\right\}$$

be the orbit of X under the action of O_n. The set of all orbits $\mathcal{O} = \{O_n(X) : X \in (S^{n-1})^d\}$ is in natural correspondence with the *elliptope*:

$$\begin{aligned}
\mathcal{E}^d :&= \left\{Y \in \mathbb{S}^d : \text{there is } X{=}(x_1, \ldots, x_d) \in (S^{n-1})^d \text{ such that } Y_{ij}{=}x_i^{\intercal}x_j\right\} \\
&= \left\{Y \in \mathbb{S}^d : Y_{ii} = 1 \;\text{ for all } i \in [d], \; Y \succeq 0\right\}.
\end{aligned} \tag{2.23}$$

For any $F \in C\left((S^{n-1})^d\right)$ invariant under the action of O_n there exists $\phi_F : \mathcal{E}^d \to \mathbb{R}$, such that

$$F(X) = \phi_F(X^\intercal X), \text{ for all } X \in (S^{n-1})^d. \tag{2.24}$$

That is, F depends on the inner products of $x_1^\intercal x_2, \ldots, x_{d-1}^\intercal x_d$ only.

Proposition 2.27. *Given $d \leq n$, let $F \in C\left((S^{n-1})^d\right)$ be invariant under the action of O_n and let $\phi_F : \mathcal{E}^d \to \mathbb{R}$ be such that $F(X) = \phi_F(X^\intercal X)$, for all $X \in (S^{n-1})^d$, then*

(a) F is non-negative if and only if ϕ_F is non-negative.

(b) F is continuous if and only if ϕ_F is continuous.

Proof. Part a is straightforward as co-domains of ϕ_F and F coincide. Next we prove statement b. Let ϕ_F be continuous. Since T is continuous, F is continuous as a composition of continuous functions. Now let F be continuous. We want to show that ϕ_F is continuous. Consider a sequence $(Y^k)_{k\in\mathbb{N}_+}$ in $\mathcal{E}^d$ such that $\lim_{k\to\infty} Y^k = Y$ and the corresponding sequence $\left(\phi_F(Y^k)\right)_{k\in\mathbb{N}_+}$. Since $\mathcal{E}^d$ is compact, $Y \in \mathcal{E}^d$. Also, for all $k \in \mathbb{N}^+$ there is $X^k \in (S^{n-1})^d$ such that $[X^k]^\intercal[X^k] = Y^k$. Since co-domains of ϕ_F and F coincide, the co-domain of ϕ_F is compact as F is continuous and $(S^{n-1})^d$ is compact. Therefore $\left(\phi_F(Y^k)\right)_{k\in\mathbb{N}_+}$ is a bounded sequence. Now, we restrict ourselves to an arbitrary convergent subsequence of this sequence. Then, up to a convergent subsequence of $(X^k)_{k\in\mathbb{N}_+}$, we have

$$\lim_{k\to\infty} \phi_F(Y^k) = \lim_{k\to\infty} F(X^k) = F(\lim_{k\to\infty} X^k) = \phi_F(Y).$$

□

Proposition 2.27 allows replacing kernels on S^{n-1} in problem (2.22) by continuous functions $\phi : \mathcal{E}^{r+2} \to \mathbb{R}$.

Lemma 2.28. *Problem* (2.22) *is equivalent to the following problem:*

$$\nu_r(\mathcal{G}_n) = \inf_{\phi, F} \quad \phi(1) + 1 \tag{2.25}$$

$$\begin{aligned} \text{s.t.}\quad & \phi(u) \leq -1, \ \textit{for all } u \in [-1, \tfrac{1}{2}], \\ & \phi(x_1^\intercal x_2) + \cdots + \phi(x_{r+1}^\intercal x_{r+2}) \geq \sigma\left(S(x_1, \ldots, x_{r+2})\right), \qquad (2.26)\\ & \textit{for all } x_1, \ldots x_{r+2} \in S^{n-1}, \\ & \phi \in C([-1,1]), \ S \in C\left((S^{n-1})^{r+2}\right)^{O_n} \\ & S \ \textit{is 2-p.d.} \end{aligned}$$

Proof. Let (K, λ) be feasible for problem (2.22). By (2.24), using Proposition 2.27 and $\mathcal{E}^1 = [-1, 1]$, there exists $\phi \in C([-1, 1])$ such that $K(x, y) = \phi(x^\intercal y)$ for all $x, y \in S^{n-1}$. Moreover, $\phi(u) \leq -1$ for all $u \in [-1, \frac{1}{2}]$. Finally, the constraint (2.26) is satisfied by (2.12). Hence ϕ is feasible for problem (2.25) with the objective value $\phi(1) + 1 = \phi(x^\intercal x) + 1 = K(x, x) + 1 = \lambda$, for all $x \in S^{n-1}$.

Now, let $\phi \in C([-1, 1])$ be feasible for problem (2.25). Define $K(x, y) := \phi(x^\intercal y)$ for all $x, y \in S^{n-1}$ and $\lambda := \phi(1) + 1$. Then K is a kernel by Proposition 2.27, $K(x, x) + 1 = \lambda$ for all $x \in S^{n-1}$, and $K(x, y) \leq -1$ for all $x, y \in S^{n-1}$ such that $x^\intercal y \in [-1, \frac{1}{2}]$. Hence (K, λ) is feasible for (2.22) using the definition (2.12) of $Q_r^{S^{n-1}}$. Moreover, the objective value for K equals $\phi(1) + 1$. □

To obtain numerical results, our next step is to characterize 2-p.d. tensors on $(S^{n-1})^{r+2}$ invariant under the action of O_n. We perform this step in the next Chapter 3 of the thesis.

Acknowledgements

We would like to thank Renata Sotirov for her useful comments on an early version of the material presented in this chapter, Miriam Dür and Claudia Adams for their remarks on the sufficiency condition in Proposition 2.10, and Fernando de Oliveira Filho for his advice on improving this chapter and for pointing out a generalization of our initial argument in Proposition 2.19.

CHAPTER 3

Generalizations of Schoenberg's theorem on positive definite kernels

3.1 Introduction

The seminal theorem of I.J. Schoenberg [201] characterizes positive definite kernels on the unit sphere S^{n-1} invariant under the automorphisms of the sphere. Recall that O_n is the orthogonal group in dimension n, which is the automorphism group of S^{n-1}.

Theorem 3.1 (Schoenberg [201])**.** *Let $n \geq 2$. The kernel $K \in \mathcal{K}(S^{n-1})$ is invariant under the action of O_n and p.d. if and only if there exists $c_i \geq 0$ for $i = 0, 1, \dots$ such that*

$$K(x, y) = \sum_{i \in \mathbb{N}} c_i P_i^{\frac{n}{2}-1}(x^\intercal y), \tag{3.1}$$

where the series converges absolutely uniformly. Also, the coefficients of the expansion (3.1) *are given by*

$$c_k = \frac{1}{p_{n/2-1,k}} \int_{x,y \in S^{n-1}} K(x, y)\, P_k^{\frac{n}{2}-1}(x^\intercal y) d\omega_n(x) d\omega_n(y), \tag{3.2}$$

where ω_n is the standard measure on the unit sphere S^{n-1} in $\mathbb{R}^n$, and

$$p_{\frac{n}{2}-1,k} := \int_{x,y \in S^{n-1}} \left| P_k^{\frac{n}{2}-1}(x^\intercal y) \right|^2 d\omega_n(x) d\omega_n(y).$$

We discuss details about Gegenbauer polynomials in Section 3.2.

In this chapter we provide two generalizations of the above result to positive definite kernels on fiber bundles. Theorem 3.7 is based on the following idea: for a set V, characterize p.d. kernels on $V \times S^{n-1}$ which are invariant under the action of O_n when the arguments from V are fixed. There exist several extensions of Schoenberg's theorem in this direction. First, Guella et al. [77] characterized F in the case when

$V = S^{m-1}$ for $m = 1, 2, \ldots, \infty$. Next, the situation where V is a compact group was considered in Berg and Porcu [17]. Finally, Guella and Menegatto [76] generalized Schoenberg's theorem for the case where V is a general set without any assumptions about its algebraic structure or topology. Inspired by these results, we let the set $V \subseteq \mathbb{R}^n$ depend on the choice of parameters from another given set B. We describe the whole family of kernels generated by all possible choices of parameters in B. To work in this setting, we use the notion of a fiber bundle, see Section 3.2.2 for more details.

Next, Theorem 3.8 characterizes the class of continuous functions $F(x, y, Z)$ on $(S^{n-1})^{r+2}$ invariant under O_n such that $F(\cdot, \cdot, Z)$ is a p.d. kernel for every $Z \in (S^{n-1})^r$.

Our work is inspired by the connection of p.d. kernels to some combinatorial problems on infinite compact graphs. The well known linear programming upper bound for the kissing number problem by Delsarte et al. [44] can be obtained using Schoenberg's theorem 3.1. The extension of Schoenberg's theorem by Bachoc and Vallentin [12] is used to obtain the strongest existing SDP upper bounds on the kissing number [12, 136, 143]. We use the results in this chapter to implement the hierarchy of upper bounds ν_r (2.25) on the kissing number problem from Chapter 2. As a result, we obtain alternative SDP bounds for this problem.

The kissing number problem is a particular instance of the more general spherical codes problem. Both problems are described in Chapter 2. Schoenberg's theorem has been used to obtain bounds on spherical codes [11, 177], as well as bounds for other problems from coding theory and discrete geometry, such as binary codes [203], sphere packings [39, 40, 101], distance avoiding sets [43], measurable chromatic number [13], one-sided kissing number [149]. The extension of Schoenberg's theorem by Musin [151] has been used to obtain bounds for the maximum number of equiangular lines in $\mathbb{R}^n$ [41].

Schoenberg's theorem has been generalized in several ways. First, let $r \geq 0$, and pick r distinct points in (S^{n-1}). Consider p.d. kernels invariant under the automorphisms of the sphere fixing those points, that is the stabilizer of those points in O_n. Schoenberg's theorem describes the case $r = 0$, when no points are fixed. Next, Bachoc and Vallentin [12] characterized the case when K is a polynomial and $r = 1$ point is fixed. Finally, Musin [151] characterized the case $r \leq n - 2$. In this chapter we extend this idea even further in Theorem 3.8. Namely, we consider the class of continuous functions $F(x, y, Z)$ on $(S^{n-1})^{r+2}$ such that $F(\cdot, \cdot, Z)$ is a p.d. kernel for every $Z \in (S^{n-1})^r$. According to Definition 2.7 in Chapter 2 we call such functions *2-p.d. functions*. There is a close connection between our result and [151], as for any fixed $Z \in (S^{n-1})^r$ we have that $F(\cdot, \cdot, Z)$ is a p.d. kernel invariant under the stabilizer of Z. Thus Musin's result characterizes $F(\cdot, \cdot, Z)$ for each fixed Z. However, it does not fully characterize F as the dependence on Z is not explicit in [151] since Z is assumed to be constant.

The approach in this chapter differs from the approach by Musin [151]. Musin [151] uses modified Gegenbauer polynomials and the corresponding modification of the classical addition theorem for Gegenbauer polynomials [105]. On the contrary, we reduce the class of considered functions to the case where Schoenberg's theorem applies. To prove our results, we generalize the notion of a p.d. kernel from a kernel on a set to a kernel on a fiber bundle. A fiber bundle is a natural object that relates to 2-p.d. functions since every such function can be viewed as a family of kernels parametrized by $Z \in (S^{n-1})^r$. Our result can be also viewed as a continuation of another known extension of Schoenberg's theorem due to Bochner [20] who generalized the theorem for group invariant p.d. kernels on compact topological spaces.

By increasing r in ν_r (2.25), we obtain a hierarchy of upper bounds on the kissing number. We implement ν_r for $r \in \{0, 1, 2\}$ using the characterization of 2-p.d. functions on $(S^{n-1})^{r+2}$. The bound for level zero of the hierarchy coincides with Delsarte, Goethals and Seidel [44] linear programming (LP) bound for the spherical codes problem. The bound for level one is similar to the semidefinite programming (SDP) bound for the spherical codes problem by Bachoc and Vallentin [12] but numerically weaker (see Section 3.7).

We analyze the relation between our bounds and the bound in [12] in Section 3.7. The results are not conclusive, so further research is needed to understand the connection better. Our bound for level one of the hierarchy is numerically weaker than the bound in [12]. However, in contrast to the approach in [12], we provide not one upper bound but a sequence of convergent upper bounds on the kissing number. Higher levels of our hierarchies could provide stronger bounds.

The outline of the chapter is as follows. In Section 3.2 we present the basic notation and concepts used throughout the chapter and motivate our study. In particular, Subsection 3.2.2 introduces fiber bundles and kernels on them. In Section 3.3 we present our main Theorems 3.7 and 3.8. Section 3.4 contains major proofs. Further observations and ideas about future research are considered in Section (3.5). In Section 3.6, we describe the application of our results to the kissing number problem and show computational results. In Section 3.7 we compare our optimization problems with some well known existing upper bounds on the kissing number. All computations in this chapter are done in MATLAB R2017a with Yalmip [128] on a computer with the processor Intel® Core™ i5-3210M CPU @ 2.5 GHz and 7.7 GiB of RAM. SDP programs are solved with MOSEK Version 8.0.0.64.

3.2 Preliminaries and motivation

Throughout the chapter we use the following properties of p.d. kernels due to Schoenberg [200].

Theorem 3.2 (Schoenberg [200]). *Let $V \subset \mathbb{R}^n$ be compact, then*

(a) A sum of finitely many p.d. kernels on V is a p.d. kernel.

(b) An entry-wise product of finitely many p.d. kernels on V is a p.d. kernel.

(c) A continuous function which is a limit of a sequence of p.d. kernels on V is a p.d. kernel.

Schoenberg [201] characterized p.d. kernels on the unit sphere invariant under the action of O_n in terms of Gegenbauer, or ultraspherical, polynomials (see Theorem 3.1). Gegenbauer polynomials $P_k^\alpha(t) : [-1,1] \to \mathbb{R}$ of order α and degree k are inductively defined for any $\alpha > -\frac{1}{2}$ and $k \geq 0$ (see, e.g., Chapter 6.4 in [5]), as $P_0^\alpha(t) = 1, P_1^\alpha(t) = 2\alpha t$ and for $k > 1$

$$dP_k^\alpha(t) = 2t(k+\alpha-1)P_{k-1}^\alpha(t) - (k+2\alpha-2)P_{k-2}^\alpha(t) \tag{3.3}$$

For $\alpha > 0$ we have $P_i^\alpha(t) \leq P_i^\alpha(1)$ and $P_i^\alpha(1) > 0$, which implies that $\sum_{i\in\mathbb{N}} c_i$ in expansion (3.1) converges.

For each fixed order α, Gegenbauer polynomials form an orthogonal basis, with respect to the weight function $(1-t^2)^{\alpha-\frac{1}{2}}$, for univariate polynomials on the interval $[-1,1]$. Thus, every univariate polynomial $g(t)$ of degree k can be represented via its Gegenbauer polynomial expansion

$$g(t) = \sum_{i=0}^{k} c_i P_i^\alpha(t), \ c_i \in \mathbb{R}. \tag{3.4}$$

For $\alpha = \frac{n}{2} - 1$, the weight function $(1-t^2)^{\alpha-\frac{1}{2}}$ can be associated with ω_n. Namely, for any $f \in C([-1,1])$ we obtain, by switching to polar coordinates,

$$\int_{x,y\in S^{n-1}} f(x^\intercal y) d\omega_n(x) d\omega_n(y) = \omega_n(S^{n-1})\omega_{n-1}(S^{n-2}) \int_{-1}^{1} f(t)(1-t^2)^{\frac{n-3}{2}} dt,$$

where $\omega_n(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$ is the surface area of S^{n-1}. Based on this observation, in the sequel we use the following notation

$$\begin{aligned} p_{\frac{n}{2}-1,k} &:= \int_{x,y\in S^{n-1}} \left| P_k^{\frac{n}{2}-1}(x^\intercal y) \right|^2 d\omega_n(x) d\omega_n(y) \\ &= \omega_n(S^{n-1})\omega_{n-1}(S^{n-2}) \int_{-1}^{1} \left| P_k^{\frac{n}{2}-1}(t) \right|^2 (1-t^2)^{\frac{n-3}{2}} dt. \end{aligned}$$

Recall that $Z = [z_1, \ldots, z_r] \in (S^{n-1})^r$ denotes the matrix Z which has the vectors $z_1, \ldots, z_r$ as its columns (see Remark 2.26). Let

$$\mathrm{Stab}_{O_n}(Z) = \{P \in O_n : PZ = Z, \text{ i.e., } Pz_i = z_i foralli \in [r]\}$$

be the stabilizer of Z in O_n. Musin [151] extended Schoenberg's theorem by characterizing p.d. kernels on the unit sphere invariant under the action of $\mathrm{Stab}_{O_n}(Z)$. To present this extension theorem, denote by $\mathcal{R}(Z)$ the range of Z, let $\Pi_Z = Z(Z^\intercal Z)^{-1}Z^T$ be the orthogonal projection onto $\mathcal{R}(Z)$, and let $\Pi_Z^\perp = I - \Pi_Z$, where I is the identity matrix, be the orthogonal projection onto $\mathcal{R}(Z)^\perp$.

Theorem 3.3 (Musin [151])**.** *Let $n \geq 2$ and $n \geq r+2$, and let $Z \in (S^{n-1})^r$ be of rank r. The kernel $K \in \mathcal{K}(S^{n-1})$ is invariant under the action of* $\mathrm{Stab}_{O_n}(Z)$ *and p.d. if and only if there exist p.d. kernels c_i on $\{x \in \mathbb{R}^r : z = Z^\intercal y,\ y \in S^{n-1}\}$ for $i = 0, 1, \ldots$ such that*

$$K(x, y) = \sum_{i \in \mathbb{N}} c_i(Z^\intercal x, Z^\intercal y) P_i^{\frac{n-r}{2}-1}\left(\frac{\left(\Pi_Z^\perp x\right)^\intercal \Pi_Z^\perp y}{\|\Pi_Z^\perp x\|\|\Pi_Z^\perp y\|}\right). \tag{3.5}$$

When $r = 1$ and K is a polynomial, the result in Theorem 3.3 follows from the decomposition by Bachoc and Vallentin [12], who use classical results on spherical harmonics, see, e.g., Chapter 9 in [5].

Although this fact is not stated explicitly, in Theorem 3.3 functions c_i can differ for different choices of Z. More precisely, for each orbit $Z^{O_n} = \{PZ : P \in O_n\}$ of $Z \in (S^{n-1})^r$ we have a different $c_i^{Z^{O_n}}$. Since Z is fixed, this dependence of c_i on the orbit of Z is implicit in Theorem 3.3. We generalize Theorem 3.3 taking this dependence into account in order to characterize 2-p.d. functions on the unit sphere invariant under the action of O_n.

We generalize Theorem 3.3 using the following observation. Let $Z \in (S^{n-1})^r$ be of rank r. For any $x \in \mathbb{R}^n$ we can write $x = \Pi_Z x + \Pi_Z^\perp x$, where $\Pi_Z x \in \mathcal{R}(Z)$ and $\Pi_Z^\perp x \in \mathcal{R}(Z)^\perp$. Now, $\mathrm{Stab}_{O_n}(Z)$ fixes $\Pi_Z S^{n-1}$ and acts transitively on $\Pi_Z^\perp S^{n-1}$. Therefore for every $x \in S^{n-1}$ one can separate the fixed component $\Pi_Z x$ and exploit the symmetry of the varying component $\Pi_Z^\perp x$. Notice that $\Pi_Z^\perp x$ is isomorphic to the unit sphere in $\mathbb{R}^{n-r}$. Moreover, every action of $\mathrm{Stab}_{O_n}(Z)$ on $\Pi_Z^\perp x$ is associated with an action of O_n on S^{n-r-1}. Hence, one can use Schoenberg's theorem on S^{n-r-1} to characterize $K(x, y)$ when the components of x and y that belong to $\Pi_Z x$ are fixed. For every fixed Z that would provide expansion (3.5). Our goal is to show what happens when we let Z vary, as we can obtain a different expansion (3.5) for different Z. To be able to work with all these expansions simultaneously, we consider a generalization of p.d. kernels on a set to p.d. kernels on a fiber bundle (see Subsection 3.2.2 for precise definitions).

3.2.1 Motivation

Theorems 3.1 and 3.3 were used to obtain new upper bounds on the spherical codes problem. Recall that in this problem, the number $A(n,\theta)$ of points on S^{n-1} is maximized, for which the pairwise angular distance is not smaller than some value θ. Schoenberg's theorem (Theorem 3.1) leads to the linear programming upper bound for the spherical codes problem by Delsarte et al. [44], and Musin's theorem (Theorem 3.3) when $r = 1$ leads to the semi-definite programming bounds by Bachoc and Vallentin [12].

Our findings in this chapter are motivated by problem (2.25) in the previous Chapter 2, which provides an upper bound on the kissing number problem. Let $V = S^{n-1}$, $r \geq 1$, and let F be a 2-p.d. function (see Definition 2.7) on the unit sphere invariant under the action of O_n. Implementing problem (2.25) requires a characterization of F. It is clear that for every fixed $Z \in (S^{n-1})^r$, $F(\cdot,\cdot,Z)$ is invariant under $\mathrm{Stab}_{O_n}(Z)$, and thus has the form as in Theorem 3.3. However, in the context of problem (2.25), Z can vary. The question is then how to modify Theorem 3.3 to make explicit the dependence on Z.

3.2.2 Fiber bundles and kernels on fiber bundles

We think on kernels parameterized by a set of parameters B. Here not only the kernel depends on the parameters from B, but also the domain of the kernel might depend on the choice of the parameters from B. Kernels on fiber bundles are not usually used in optimization, but they have been studies in physics (see, e.g., [56]), and they can be naturally associated to 2-p.d. functions (see Remark 3.5).

A fiber bundle is a map $f : A \to B$, where A is called the total space and B is called the base space [90]. For each $b \in B$, $A_b := f^{-1}(b) \subset A$ is called the fiber over b. We think of A_b as representing a set "parameterized by" b. Our definition of a bundle is quite unrestrictive. In particular, we do not ask the fibers to be homeomorphic. We assume the following about $f : A \to B$. First, we restrict ourselves to the case where A and B are Euclidean sets with natural topology. Next, as we work with continuous functions over fiber bundles, we require that *local cross sections* exist in $f : A \to B$. That is, there exists a *continuous* map $g : U \to A$, where $U \subseteq B$ is an open set and $f(g(u)) = u$ for all $u \in U$. Finally, in some cases we require that the map f is *proper*; that is, preimages of compact sets under f are compact.

Definition 3.4. *As examples, we define the following bundles which we frequently use in the sequel.*

1. *Given $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^m$, let $\pi_{A,B} : A \times B \to B$ be the projection bundle defined by $\pi_{A,B}(a,b) := b$.*

2. *Given bundles* $f_1 : A_1 \to B$ *and* $f_2 : A_2 \to B$, *define*

$$A_1 \times_B A_2 := \{(a_1, a_2) \in A_1 \times A_2 : f_1(a_1) = f_2(a_2)\}.$$

We define the fiber product bundle $f_1 \times_B f_2 : A_1 \times_B A_2 \to B$ *as*

$$f_1 \times_B f_2(a_1, a_2) := f_1(a_1).$$

The fiber product is also called the Whitney sum. It has the property that for every $b \in B$, $(A_1 \times_B A_2)_b = (A_1)_b \times (A_2)_b$.

3. *Given a bundle* $f : A \to B$ *and* $U \in \mathbb{R}^k$, *let* ${}_{U\times}f : U \times A \to B$ *be the bundle such that* ${}_{U\times}f(u, a) := f(a)$. *In the case* $U = S^{n-1}$, *we call* ${}_{S^{n-1}\times}f$ *a cylinder.*

Now, we introduce the notion of a kernel on a bundle. The idea is that for each $b \in B$ we have a kernel on A_b, and the dependence on b is continuous. Given a bundle $f : A \to B$, we define a kernel on f to be a continuous map $K : A \times_B A \to \mathbb{R}$ such that $K_b : A_b \times A_b \to \mathbb{R}$ is a kernel for each $b \in B$. We say K is p.d. on f if K_b is p.d. for each $b \in B$.

We say that a bundle $f : U \to B$ is a subbundle of $g : A \to B$ if $U \subseteq A$, and $f = g|_U$. We call $f : U \to B$ a projection subbundle if it is a subbundle of some projection bundle.

Remark 3.5. *Given any projection bundle* $\pi_{A,B} : A \times B \to B$, *we have* $\pi_{A,B} \times_B \pi_{A,B} \cong \pi_{A\times A,B}$, *and thus every p.d. kernel on a subbundle* $f : U \to B$ *of* $\pi_{A,B}$ *is in correspondence with a continuous map* $K : U \times U \times B \to \mathbb{R}$ *such that for each* $b \in B$, $K(\cdot, \cdot, b) = K_b$ *is a p.d. kernel on* U_b. *In the sequel we abuse the notation and make no difference between a kernel on projection subbundle* $f : U \to B$ *and its corresponding continuous map.*

Our last definition is the action of a group on a bundle. Given bundle $f : A \to B$ and group G, for G to act on f means that G acts both on A and on B, and both actions are consistent with f. That is, for all $g \in G$ and $a \in A$, $f(a)^g = f(a^g)$. We denote the orbit of $a \in A$ under G by $a^G := \{a^g : g \in G\}$, and let $\mathcal{O}_G(A) = \left\{a^G : a \in A\right\}$ be the set of orbits of A. We define b^G and $\mathcal{O}_G(B)$ analogously.

When G acts on a bundle $f : A \to B$, it is natural to define the G-orbit bundle of f as $\mathcal{O}_G(f) : \mathcal{O}_G(A) \to \mathcal{O}_G(B)$ such that $\mathcal{O}_G(f)(a^G) = f(a)^G$. Notice that in the G-orbit bundle, for any $b^G \in \mathcal{O}_G(B)$ we have $[\mathcal{O}_G(A)]_{b^G} = \mathcal{O}_G(A_b)$.

Now, we propose an extension of the group acting on B to a group acting on projection bundle $\pi_{A,B}$. Assume G acts on B. We define the vertical action of G on $\pi_{A,B}$ by fixing the elements of A; that is, for all $a \in A$, $b \in B$ and $g \in G$ we define the action of G on $A \times B$ by $(a, b)^g := (a, b^g)$. Notice that this action and the action of G on B are consistent with $\pi_{A,B}$, and thus define an action on $\pi_{A,B}$. Moreover, $\mathcal{O}_G(\pi_{A,B}) = \pi_{A,\mathcal{O}_G(B)}$. For any projection subbundle $f : U \to B$, we say that G acts vertically on f if the action of G is the restriction of the vertical action on the corresponding projection bundle. Notice that this is the case only if for any $b \in B$ and $g \in G$ we have $U_b = U_{b^g}$.

In general, looking at $\mathcal{O}_G(f)$ is not enough to characterize p.d. kernels on f invariant under the action of G as kernels are bivariate functions and thus one should look at 2-orbits, instead of 1-orbits. One exception is the case of vertical actions, as the following straightforward proposition shows. From here on our groups of interest are either O_n or its subgroups. Hence we formulate the proposition for O_n although it could be generalized to a wider class of groups.

Proposition 3.6. *Let $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$. Assume that O_n acts on B and endows $\mathcal{O}_{O_n}(B)$ with the usual topology. Let $f : A \to B$ be a projection subbundle such that O_n acts vertically on f. Let K be a kernel on f invariant under the vertical action of O_n. Define the function K^{O_n} as*

$$K^{O_n}(a_1, a_2, b^{O_n}) := K(a_1, a_2, b) \text{ for all } b \in B,\ a_1, a_2 \in A_b.$$

Then K^{O_n} is a kernel on $\mathcal{O}_{O_n}(f)$, and K is p.d. if and only if K^{O_n} is p.d.

Proof. The result follows from the definition of the vertical action and the O_n-orbit bundle. □

This work is motivated by 2-p.d. functions introduced in Chapter 2. Remark 3.5 explains that 2-p.d. functions on A are p.d. kernels on π_{A,A^r} for some given r. In chapter Chapter 2 we are interested is in $(r+2)$-variate 2-p.d. functions on S^{n-1} invariant under the natural action of O_n on $(S^{n-1})^{r+2}$. In the language of kernels on fiber bundles, those are p.d. kernels on $\pi_{S^{n-1},(S^{n-1})^r}$ invariant under the natural action of O_n on $\pi_{S^{n-1},(S^{n-1})^r}$. In the rest of this chapter we characterize such kernels.

3.3 Main results

Next we present two generalizations of Schoenberg's theorem (Theorem 3.1). Given a bundle $f : A \to B$, we define the horizontal action of O_n on the cylinder ${}_{S^{n-1}\times}f$ by $(x,a)^P = (Px, a)$ and $b^P = b$, for each $x \in S^{n-1}$, $a \in A$, $b \in B$ and $P \in O_n$. In our first Theorem we characterize the p.d. kernels on cylinders, invariant under the horizontal action of O_n. In particular, it allows to characterize invariant (under the action of O_n) p.d. kernels on a product of S^{n-1} and a compact set.

Theorem 3.7. *Let $n \geq 2$, $r > 0, m > 0$, and let $A \in \mathbb{R}^r, B \in \mathbb{R}^m$. Let $f : A \to B$ be a bundle with proper f. Then a kernel K on the cylinder ${}_{S^{n-1}\times}f$ is p.d. and* *invariant under the horizontal action of O_n if and only if there are p.d. kernels c_i on f, $i = 0, 1, \dots$ such that for all $b \in B$, $u_1, u_2 \in S^{n-1}$, and $a_1, a_2 \in A_b$*

$$K_b\left(\left[\begin{smallmatrix} u_1 \\ a_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} u_2 \\ a_2 \end{smallmatrix}\right]\right) = \sum_{i \in \mathbb{N}} (c_i)_b(a_1, a_2) P_i^{\frac{n}{2}-1}(u_1^\intercal u_2). \tag{3.6}$$

Also, the coefficients of the expansion are given by

$$(c_k)_b(a_1, a_2) = \frac{1}{p_{n/2-1,k}} \int_{u_1, u_2 \in S^{n-1}} K_b\left(\left[\begin{smallmatrix} u_1 \\ a_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} u_2 \\ a_2 \end{smallmatrix}\right]\right) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2) d\omega_n(u_1) d\omega_n(u_2). \tag{3.7}$$

To show the intuition behind Theorem 3.7, let $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^n$ be compact sets, and consider a function $G \in C\Big(B \times (A \times S^{n-1})^2\Big)$. By fixing elements $b \in B$ and $a_1, a_2 \in A$, we obtain the function $G(b, a_1, a_2, \cdot, \cdot) \in C(S^{n-1} \times S^{n-1})$. Assume that $G(b, a_1, a_2, \cdot, \cdot)$ is a kernel invariant under the action of O_n. Then G satisfies the conditions in Theorem 3.7, and therefore it has expansion (3.6). However, Theorem 3.7 allows considering a more general setting: using fiber bundles makes it possible for the set A to depend on some parameters from the set B. Namely, for every $b \in B$ we may have $A_b \subset \mathbb{R}^n$, and these sets may vary for different b. Theorem 3.7 allows us to work with all these sets simultaneously. It says that even if A depends on B, the function G still has representation (3.6).

Notice that $\pi_{S^{n-1},(S^{n-1})^r}$ is isomorphic to the cylinder ${}_{S^{n-1}\times}\mathrm{id}\,(S^{n-1})^r$, where $\mathrm{id}\,(S^{n-1})^r$ is the identity bundle on $(S^{n-1})^r$. Consider the (natural) action of O_n on $\pi_{S^{n-1},(S^{n-1})^r}$. This action is not horizontal, and thus Theorem 3.7 does not apply. Our second theorem describes p.d. kernels on the bundle $\pi_{S^{n-1},(S^{n-1})^r}$ which are invariant under the action of O_n described in Section 1.1 of Chapter 1. Given $r > 0$, define

$$\mathcal{S} = \{Z \in (S^{n-1})^r : \operatorname{rank} Z = r\}. \tag{3.8}$$

Notice that set $\mathcal{S}$ is dense in $(S^{n-1})^r$. Also, define

$$\mathcal{E}^r = \{Y \in \mathbb{S}^r : Y \succeq 0,\ Y_{ii} = 1 \text{ for all } i \in \{1, \dots, r\}\}.$$

Theorem 3.8. *Let $r \geq 0$ and $n \geq r + 2$. A kernel K on $\pi_{S^{n-1},(S^{n-1})^r}$ is p.d. $\Big($in other words, a function $K \in C\Big((S^{n-1})^{r+1}\Big)$ is 2-p.d.$\Big)$ and invariant under the action of O_n if and only if there are p.d. kernels c_i, $i = 0, 1, \dots$ on the projection subbundle $f : \Big\{\left[\begin{smallmatrix} 1 & y^\intercal \\ y & Y \end{smallmatrix}\right] \in \mathcal{E}^{r+1} : Y \succ 0\Big\} \to \{Y \in \mathcal{E}^r : Y \succ 0\}$, such that for all $x, y \in S^{n-1}$ and $Z \in \mathcal{S}$,*

$$K_Z(x, y) = \sum_{i \in \mathbb{N}} (c_i)_{Z^\intercal Z}(Z^\intercal x, Z^\intercal y) P_i^{\frac{n-r}{2}-1}\left(\frac{\left(\Pi_Z^\perp x\right)^\intercal \Pi_Z^\perp y}{\|\Pi_Z^\perp x\| \|\Pi_Z^\perp y\|}\right). \tag{3.9}$$

Theorem 3.8 characterizes 2-p.d. functions on $(S^{n-1})^{r+2}$. Let F be a 2-p.d. function on $(S^{n-1})^{r+2}$. When we fix $Z \in S^{n-1}$, we obtain $F(\cdot, \cdot, Z) \in \mathcal{PSD}\,(S^{n-1})$ invariant under $\operatorname{Stab}_{O_n}(Z)$. This kernel is characterized in Theorem 3.3 and has expansion (3.5).

What we show in Theorem 3.8 is that letting Z vary implies letting c_i depend on $Z^\intercal Z$ in expansion (3.5). Notice that Theorem 3.3 follows from Theorem 3.8, as given any $Z \in (S^{n-2})^r$ and any p.d. kernel K on S^{n-1} invariant under the action of $\mathrm{Stab}_{O_n}(Z)$, there is a p.d. kernel $\hat{K}$ on $\pi_{S^{n-1},(S^{n-1})^r}$ invariant under the action of O_n such that $\hat{K}_Z(x,y) = K(x,y)$.

3.4 Proofs of main theorems

In this section we present the proofs of Theorems 3.7 and 3.8.

3.4.1 Proof of Theorem 3.7

Let K be a kernel on ${}_{S^{n-1}\times}f$. The "only if" part of the statement follows from Theorems 3.2 and 3.1. To prove the converse, let K be p.d. and invariant under the horizontal action of O_n. Let $b \in B$ and $a_1, a_2 \in A_b$. The kernel

$$G_b^{a_1,a_2}(u_1,u_2) = K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right) + K_b\left(\left[\begin{smallmatrix}u_1\\a_2\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_1\end{smallmatrix}\right]\right) + K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_1\end{smallmatrix}\right]\right) + K_b\left(\left[\begin{smallmatrix}u_1\\a_2\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right)$$

is p.d. on S^{n-1} and invariant under O_n. From Schoenberg's theorem (Theorem 3.1) we have

$$G_b^{a_1,a_2}(u_1,u_2) = \sum_{k\geq 0} (d_i)_b(a_1,a_2) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2), \tag{3.10}$$

where the $(d_i)_b(a_1,a_2)$ are non-negative, and the series (3.10) converges absolutely uniformly.

As O_n acts transitively on S^{n-1}, we have

$$K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right) = K_b\left(\left[\begin{smallmatrix}u_2\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_1\\a_2\end{smallmatrix}\right]\right) = K_b\left(\left[\begin{smallmatrix}u_1\\a_2\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_1\end{smallmatrix}\right]\right),$$

thus,

$$\begin{aligned} G_b^{a_1,a_2}(u_1,u_2) &= 2K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right) + K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_1\end{smallmatrix}\right]\right) + K_b\left(\left[\begin{smallmatrix}u_1\\a_2\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right) \\ G_b^{a_i,a_i}(u_1,u_2) &= 4K_b\left(\left[\begin{smallmatrix}u_1\\a_i\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_i\end{smallmatrix}\right]\right) \qquad (i=1,2). \end{aligned}$$

Defining $(c_k)_b(a_1,a_2) = (d_k)_b(a_1,a_2) - \frac{1}{4}(d_k)_b(a_1,a_1) - \frac{1}{4}(d_k)_b(a_2,a_2)$ and using (3.10), we obtain

$$\begin{aligned} K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right) &= \tfrac{1}{2}\left(G_b^{a_1,a_2}(u_1,u_2) - \tfrac{1}{4}G_b^{a_1,a_1}(u_1,u_2) - \tfrac{1}{4}G_b^{a_2,a_2}(u_1,u_2)\right) \\ &= \tfrac{1}{2}\sum_{k\in\mathbb{N}} (c_k)_b(a_1,a_2) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2). \end{aligned} \tag{3.11}$$

Remark 3.9. *Notice that Schoenberg's theorem can not be applied to* $K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right)$ *directly, as this is not necessarily a p.d. kernel for all* $b \in B$ *and* $a_1, a_2 \in A_b$. *Intuitively the reason for this is that this function does not correspond to a "principal submatrix" of* K_b *when* $a_1 \neq a_2$.

Next, we argue that c_k's are p.d. kernels on f. Fix $k \geq 0$, then we claim that

Claim 3.10. *For every* $b \in B$, $a_1, a_2 \in A_b$,

$$(c_k)_b(a_1, a_2) = \tfrac{1}{p_{n/2-1,k}} \int_{u_1,u_2 \in S^{n-1}} K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2) d\omega_n(u_1) d\omega_n(u_2)$$

Fix $b \in B$. Claim 3.10 and the continuity of K imply that $(c_k)_b$ is continuous. From (3.7) it follows that $(c_k)_b(a_1, a_2) = (c_k)_b(a_2, a_1)$ for all $a_1, a_2 \in A_b$. Hence $(c_k)_b$ is a kernel.

From our assumptions, A_b is compact. We use Theorem 1.3 to show that $(c_k)_b$ is p.d. on $S^{n-1} \times A_b$. Let $h \in C(A_b)$ be given, then

$$\begin{aligned}
&\int_{a_1,a_2 \in A_b} (c_k)_b(a_1, a_2) h(a_1) h(a_2) d\mu(a_1) d\mu(a_2) \\
&\quad = \int_{\substack{a_1,a_2 \in A_b \\ u_1,u_2 \in S^{n-1}}} K_b\left(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right]\right) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2) h(a_1) h(a_2) d\omega_n(u_1) \ldots d\mu(a_2) \\
&\quad \geq 0.
\end{aligned}$$

The inequality holds since K_b and $P_k^{\frac{n}{2}-1}$ are p.d. on $S^{n-1} \times A_b$, and from Theorem 3.2.b their product is p.d. too.

To finish, we prove Claim 3.10. We know that $\left\{P_i^{\frac{n}{2}-1}\right\}_{i \in \mathbb{N}}$ is a sequence of orthogonal polynomials on S^{n-1} under ω_n. Fix $k \geq 0$, $b \in B$ and $a_1, a_2 \in A_b$. From Schoenberg's theorem 3.1 and the definition of $(c_i)_b(a_1, a_2)$ in 3.11, we have that $\sum_{i \geq 0} (c_i)_b(a_1, a_2) P_i^{\frac{n}{2}-1}(u_1^\intercal u_2)$ converges absolutely uniformly on $S^{n-1} \times S^{n-1}$. Hence, as P_k is continuous on $[-1, 1]$ and therefore bounded, the series

$$\sum_{i \geq 0} (c_i)_b(a_1, a_2) P_i^{\frac{n}{2}-1}(u_1^\intercal u_2) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2)$$

converges absolutely uniformly too. Therefore

$$\begin{aligned}
&\int_{u_1,u_2 \in S^{n-1}} K_b\Big(\left[\begin{smallmatrix}u_1\\a_1\end{smallmatrix}\right],\left[\begin{smallmatrix}u_2\\a_2\end{smallmatrix}\right], b\Big) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2) d\omega_n(u_1) d\omega_n(u_2) \\
&\quad = \int_{u_1,u_2 \in S^{n-1}} \sum_{i \geq 0} (c_i)_b(a_1, a_2) P_i^{\frac{n}{2}-1}(u_1^\intercal u_2) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2) d\omega_n(u_1) d\omega_n(u_2) \\
&\quad = \sum_{i \geq 0} (c_i)_b(a_1, a_2) \int_{u_1,u_2 \in S^{n-1}} P_i^{\frac{n}{2}-1}(u_1^\intercal u_2) P_k^{\frac{n}{2}-1}(u_1^\intercal u_2) d\omega_n(u_1) d\omega_n(u_2) \\
&\quad = \; p_{n/2-1,k} (c_k)_b(a_1, a_2).
\end{aligned}$$

3.4.2 Proof of Theorem 3.8

The idea of the proof is to relate kernels on $\pi_{S^{n-1},(S^{n-1})^r}$ to kernels on cylinders over S^{n-r-1} and apply Theorem 3.7. We do this via continuous transformations using Lemmas 3.11 and 3.12 below.

Lemma 3.11. *Let $f : C \to B$ be a bundle with proper f. Let $T : A \to C$ be a continuous function. Let K be a kernel on f. Define L by $L_b(a_1, a_2) = K_b(T(a_1), T(a_2))$ for all $b \in B$, $a_1, a_2 \in A_b$.*

1. *L is a kernel on $f \circ T$.*
2. *If K is p.d. on f, then L is p.d. on $f \circ T$.*
3. *If T is surjective and L is p.d. on $f \circ T$, then K is p.d. on f.*

Proof. As T and K are continuous, L is continuous by definition. Also, for any $b \in B$ and any $a_1, a_2 \in A_b$, we have $L_b(a_1, a_2) = K_b(T(a_1), T(a_2)) = K_b(T(a_2), T(a_1)) = L_b(a_2, a_1)$. Thus L is a kernel on $f \circ T$. Now assume K is p.d. on f. For any $k > 0$, any $b \in B$, and any $a_1, a_2, \ldots, a_k \in A_b$, we have that $\left[L_b(a_i,a_j) \right]_{1 \le i,j \le k} = \left[K_b(T(a_i),T(a_j)) \right]_{1 \le i,j \le k} \succeq 0$. Thus L is p.d. on $f \circ T$. Now assume T is surjective and L is p.d. on $f \circ T$. Then given $k > 0$, any $b \in B$, and any $c_1, c_2, \ldots, c_k \in C_b$ there are $a_1, a_2, \ldots, a_k \in A_b$ such that $c_i = T(a_i)$. Thus the matrix $\left[K_b(c_i,c_j) \right]_{1 \le i,j \le k} = \left[L_b(a_i,a_j) \right]_{1 \le i,j \le k}$ is p.d.. Therefore K is p.d. on f. □

Let $Z \in (S^{n-1})^r$. We have $\dim(\mathcal{R}(Z)^\perp) = n - \dim(\mathcal{R}(Z)) = n - \operatorname{rank} Z$, and therefore $\mathcal{R}(Z)^\perp$ is isomorphic to $\mathbb{R}^{n-\operatorname{rank} Z}$. Namely, there is an isomorphism ϕ_Z between $R^{n-\operatorname{rank} Z-1}$ and $\mathcal{R}(Z)^\perp$. Analogously, $\mathcal{R}(Z)$ has dimension $\operatorname{rank} Z$ and thus is isomorphic to $\mathbb{R}^{\operatorname{rank} Z}$, and there is an isomorphism γ_Z between $R^{\operatorname{rank} Z}$ and $\mathcal{R}(Z)$. In the sequel we restrict most of our arguments to $\mathcal{S}$ (3.8), to avoid "singularities" in further proofs and definitions. In particular, the dependence on Z of the isomorphisms ϕ_Z and γ_Z can not be continuous in the whole $(S^{n-1})^r$, but when we restrict ourselves to $\mathcal{S}$, ϕ_Z and γ_Z can be chosen continuous. For instance, let $Ort : \mathbb{R}^{n \times r} \to \mathbb{R}^{n \times (n-r)}$ be such that $Ort(Z)$ provides an orthonormal basis of $\mathcal{R}(Z)^\perp$ for any $Z \in \mathbb{R}^{n \times r}$ of rank r. Then the isomorphism between $\mathcal{R}(Z)^\perp$ and $\mathbb{R}^{n-r}$ can be viewed as the bijection:

$$\phi_Z : \mathbb{R}^{n-r} \to \mathcal{R}(Z)^\perp, \ \phi_Z(v) = Ort(Z)v \text{ for all } v \in \mathbb{R}^{n-r}.$$

We can construct a continuous isomorphism between R^r and $\mathcal{R}(Z)$ as the following bijection

$$\gamma_Z : \mathbb{R}^r \to \mathcal{R}(Z), \ \gamma_Z(u) = Z(Z^\intercal Z)^{-1}u \text{ for all } u \in \mathbb{R}^r.$$

We are particularly interested in the isomorphism between $\Pi_Z S^{n-1} \subset \mathcal{R}(Z)$ and

$$B_Z := \gamma_Z^{-1} \circ \Pi_Z S^{n-1} = \{Z^\intercal x : x \in S^{n-1}\}.$$

Notice that for any $Z \in \mathcal{S}$ we can send $x \in \Pi_Z^\perp S^{n-1}$ to the unit sphere in $\mathbb{R}^{n-r}$ by normalizing x. Then, since $\mathrm{Stab}_{O_n}(Z)$ is isomorphic to O_{n-r}, any action of $\mathrm{Stab}_{O_n}(Z)$ on $\Pi^\perp S^{n-1}$ can be associated with an action of O_{n-r} on S^{n-r-1}. Hence we can use the result for the orthogonal group acting on the unit sphere, described in Theorem 3.7. To formalize this procedure, we need the following lemma.

Lemma 3.12. *The maps*

$$\begin{aligned} T_1 : S^{n-r-1} \times \{(u, Z) : Z \in \mathcal{S},\ u \in B_Z\} &\to \{(x, Z) : Z \in \mathcal{S},\ x \in S^{n-1}\} \\ \begin{bmatrix} v \\ u \\ Z \end{bmatrix} &\mapsto \begin{bmatrix} \phi_Z(v)\sqrt{1-\|\gamma_Z(u)\|^2}+\gamma_Z(u) \\ Z \end{bmatrix} \end{aligned}$$

$$\begin{aligned} T_2 : \{(x, Z) : Z \in \mathcal{S},\ x \in S^{n-1} \setminus \mathcal{R}(Z)\} &\to S^{n-r-1} \times \{(u, Z) : Z \in \mathcal{S}, u \in B_Z\} \\ \begin{bmatrix} x \\ Z \end{bmatrix} &\mapsto \begin{bmatrix} \frac{\phi_Z^{-1}(\Pi_Z^\perp x)}{\|\phi_Z^{-1}(\Pi_Z^\perp x)\|} \\ \gamma_Z^{-1}(\Pi_Z x) \\ Z \end{bmatrix} \end{aligned}$$

are continuous, T_1 is surjective, T_2 is injective, and $T_1 \circ T_2 = \mathrm{id}_{\{(x,Z):Z\in\mathcal{S},\, x\in S^{n-1}\setminus\mathcal{R}(Z)\}}$.

Proof. Continuity follows from the continuity of γ_Z, ϕ_Z and their inverses. It is a straightforward calculation to check that T_1 is a surjection and that we have $T_1 \circ T_2 = \mathrm{id}_{\{(x,Z):Z\in\mathcal{S},\, x\in S^{n-1}\setminus\mathcal{R}(Z)\}}$. T_2 is injective since for any $x_1, x_2 \in \mathbb{R}^n$, if $x_1 \neq x_2$ then either $\Pi_Z x_1 \neq \Pi_Z x_2$ or $\Pi_Z^\perp x_1 \neq \Pi_Z^\perp x_2$. □

Proof of Theorem 3.8 . First, denote

$$\mathcal{B} := \{(u, Z) : Z \in \mathcal{S},\ u \in B_Z\} = \left\{(Z^\intercal x, Z) : Z \in \mathcal{S},\ x \in S^{n-1}\right\},$$

and consider the projection subbundle $f : \mathcal{B} \to \mathcal{S}$. Clearly, f is proper. Let K be a kernel on $\pi_{S^{n-1},(S^{n-1})^r}$. Throughout the proof we use a function L on the cylinder ${}_{S^{n-r-1}\times}f$ defined by

$$L_Z\left(\begin{bmatrix} u_1 \\ v_1 \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \end{bmatrix}\right) = K\left(T_1\left(\begin{bmatrix} u_1 \\ v_1 \\ Z \end{bmatrix}\right), T_1\left(\begin{bmatrix} u_2 \\ v_2 \\ Z \end{bmatrix}\right)\right), \tag{3.12}$$

for all $v_1, v_2 \in S^{n-r-1}$, $Z \in \mathcal{S}$, $u_1, u_2 \in B_Z$. By our assumptions K is a kernel on $\pi_{S^{n-1},(S^{n-1})^r}$. Since T_1 is surjective by Lemma 3.12, we have that L is a kernel on ${}_{S^{n-r-1}\times}f$ by Lemma 3.11.

For the "only if" direction of the theorem, let K be a kernel on $\pi_{S^{n-1},(S^{n-1})^r}$ that has expansion (3.9). Then K is invariant under the action of O_n as it is continuous and depends only on inner products of $x, y, z_1, \ldots, z_r$ on a dense subset of its domain. To show that K is p.d., notice that expansion (3.9) of K implies L has an expansion of type (3.6). Hence L is a p.d. kernel on ${}_{S^{n-r-1}\times}f$ by Theorem 3.7. Therefore K is p.d. when restricted to a kernel on $\pi_{S^{n-1},\mathcal{S}}$ by Lemma 3.11, and thus K is p.d. by continuity.

For the "if" direction of the theorem, let K be p.d. on the projection bundle $\pi_{S^{n-1},(S^{n-1})^r}$ and invariant under the action of O_n. That is, $K_{PZ}(Px, Py) = K_Z(x,y)$ for all $P \in O_n, x, y \in S^{n-1}$ and $Z \in (S^{n-1})^r$. Then for any $Z \in (S^{n-1})^r$, K_Z is invariant under the action of $\mathrm{Stab}_{O_n}(Z)$. Using Lemma 3.11, we obtain that L is a p.d. kernel on ${}_{S^{n-r-1}\times}f$. As $\mathrm{Stab}_{O_n}(Z)$ fixes $\Pi_Z S^{n-1}$ and acts transitively on $S^{n-1} \cap \mathcal{R}(Z)^\perp$, we have that L is invariant under the horizontal action of O_{n-r} on ${}_{S^{n-r-1}\times}f$. From Theorem 3.7 there are p.d. kernels d_i on f, $i = 0, 1, \ldots$ such that for all $v_1, v_2 \in S^{n-r-1}$, all $Z \in \mathcal{S}$ and $u_1, u_2 \in B_Z$

$$L_Z\left(\left[\begin{smallmatrix} u_1 \\ v_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} u_2 \\ v_2 \end{smallmatrix}\right]\right) = \sum_{i\geq 0} (d_i)_Z(u_1, u_2) P_i^{\frac{n-r}{2}-1}(v_1^\intercal v_2). \tag{3.13}$$

Now, $T_1 \circ T_2 = \mathrm{id}_{\{(x,Z):Z\in\mathcal{S},\, x\in S^{n-1}\setminus\mathcal{R}(Z)\}}$ from Lemma 3.12. Thus for any $Z \in \mathcal{S}$ and $x_1, x_2 \in S^{n-1} \setminus \mathcal{R}(Z)$, we have

$$\begin{aligned} K_Z(x_1, x_2) &= L_Z\left(T_2\left(\left[\begin{smallmatrix} x_1 \\ Z \end{smallmatrix}\right]\right), T_2\left(\left[\begin{smallmatrix} x_2 \\ Z \end{smallmatrix}\right]\right)\right) \\ &= \sum_{i\geq 0} (d_i)_Z\left(\gamma_Z^{-1}(\Pi_Z x_1), \gamma_Z^{-1}(\Pi_Z x_2)\right) P_i^{\frac{n-r}{2}-1}\left(\left(\frac{\phi_Z^{-1}(\Pi_Z^\perp x_1)}{\|\Pi_Z^\perp x_1\|}\right)^\intercal \frac{\phi_Z^{-1}(\Pi_Z^\perp x_2)}{\|\Pi_Z^\perp x_2\|}\right) \\ &= \sum_{i\geq 0} (d_i)_Z(Z^\intercal x_1, Z^\intercal x_2) P_i^{\frac{n-r}{2}-1}\left(\frac{\left(\Pi_Z^\perp x_1\right)^\intercal \Pi_Z^\perp x_2}{\|\Pi_Z^\perp x_1\|\|\Pi_Z^\perp x_2\|}\right), \end{aligned} \tag{3.14}$$

where we have used $\phi_Z^{-1}(a_1)^\intercal \phi_Z^{-1}(a_2) = a_1^\intercal a_2$.

To finish, we specify the form of the d_i's. Let $k \geq 0$. By Lemma 3.11 we have that d_k is a p.d. kernel on $f : \mathcal{B} \to \mathcal{S}$. Notice that O_n acts vertically on f since for every $P \in O_n$, $x \in S^{n-1}$ and $Z \in \mathcal{S}$, we have $(Z^\intercal x, Z)^P = (Z^\intercal x, PZ)$. Next we show that d_k is invariant under this vertical action. That is, that for all $Z \in \mathcal{S}$, $v_1, v_2 \in B_Z$ and $P \in O_n$ we have $(d_k)_Z(v_1, v_2) = (d_k)_{PZ}(v_1, v_2)$. By (3.7) from Theorem 3.7,

$$(d_k)_{PZ}(v_1, v_2) = \frac{1}{p_{(n-r)/2-1,k}} \int_{u_1,u_2\in S^{n-r-1}} L_{PZ}\left(\left[\begin{smallmatrix} u_1 \\ v_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} u_2 \\ v_2 \end{smallmatrix}\right]\right) P_k^{\frac{n-r}{2}-1}(v_1^\intercal v_2) d\omega_n(v_1) d\omega_n(v_2).$$

Therefore it is enough to show that L is invariant under the vertical action of O_n on ${}_{S^{n-r-1}\times}f$. By construction of ϕ_Z, γ_Z we have

$$\phi_{PZ}(v) = P\phi_Z(v),\ \gamma_{PZ}(u) = P\gamma_Z(u),\ \text{and } \|\gamma_{PZ}(u)\| = \|\gamma_Z(u)\|.$$

Therefore

$$\begin{aligned} L_{PZ}\left(\left[\begin{smallmatrix} u_1 \\ v_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} u_2 \\ v_2 \end{smallmatrix}\right]\right) &= K\left(T_1\left(\left[\begin{smallmatrix} u_1 \\ v_1 \\ PZ \end{smallmatrix}\right]\right), T_1\left(\left[\begin{smallmatrix} u_2 \\ v_2 \\ PZ \end{smallmatrix}\right]\right)\right) \\ &= K\left(\left[\begin{smallmatrix} P & 0 \\ 0 & P \end{smallmatrix}\right] T_1\left(\left[\begin{smallmatrix} u_1 \\ v_1 \\ Z \end{smallmatrix}\right]\right), \left[\begin{smallmatrix} P & 0 \\ 0 & P \end{smallmatrix}\right] T_1\left(\left[\begin{smallmatrix} u_2 \\ v_2 \\ Z \end{smallmatrix}\right]\right)\right) = L_Z\left(\left[\begin{smallmatrix} u_1 \\ v_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} u_2 \\ v_2 \end{smallmatrix}\right]\right), \end{aligned}$$

where the last but one equality holds by invariance of K under O_n. Now, we can reduce ourselves to the orbit bundle of f under the vertical action of O_n. We have that $\mathcal{O}_{O_n}(\mathcal{S}) = \operatorname{cor}\mathcal{E}^r$, $\mathcal{O}_{O_n}(\mathcal{B}) = \left\{\left[\begin{smallmatrix} 1 & y^\intercal \\ y & Y \end{smallmatrix}\right] \in \mathcal{E}^{r+1} : Y \succ 0\right\}$ and $\mathcal{O}_{O_n}(f)$ is a projection. From Proposition 3.6 there is a p.d. kernel c_i on $\mathcal{O}_{O_n}(f) : \left\{\left[\begin{smallmatrix} 1 & y^\intercal \\ y & Y \end{smallmatrix}\right] \in \mathcal{E}^{r+1} : Y \succ 0\right\} \to \{Y \in \mathcal{E}^r : Y \succ 0\}$ such that for all $Z \in \mathcal{S}$, $v_1, v_2 \in B_Z$ we have that $(c_k)_{Z^\intercal Z}(v_1, v_2) = (d_k)_Z(v_1, v_2)$. □

3.5 Further observations

One question for further research is what shape the coefficients c_i from expansion (3.9) could have. We show one result in this direction: a Gegenbauer polynomial of order $\left(\frac{n-r}{2} - 1\right)$ used in (3.9) on $\pi_{S^{n-1},(S^{n-1})^r}$ can be considered as a p.d. kernel on $\pi_{S^{n-1},(S^{n-1})^{r+1}}$, and therefore it can be expressed in a series of the form (3.9) with coefficients c_i of a particular but rather complex form. To simplify the notation, given $x, y \in S^{n-1}$ and $Z \in (S^{n-1})^r$, define

$$\langle x, y\rangle_Z := (\Pi_Z^\perp x)^\intercal \Pi_Z^\perp y = x^\intercal y - (Z^\intercal x)^\intercal (Z^\intercal Z)^{-1} Z^\intercal y.$$

Proposition 3.13. *Let $r \in \mathbb{N}$, $x, y, q \in S^{n-1}$ and $Z \in (S^{n-1})^r$, then*

$$P_k^{\frac{n-r}{2}-1}\left(\frac{\langle x,y\rangle_Z}{\sqrt{\langle x,x\rangle_Z}\sqrt{\langle y,y\rangle_Z}}\right) = \sum_{i=0}^{k} c_{k,i}^{n,r} \left(\frac{\langle x,x\rangle_{[Zq]}}{\langle x,x\rangle_Z}\right)^{i/2} \left(\frac{\langle y,y\rangle_{[Zq]}}{\langle y,y\rangle_Z}\right)^{i/2} P_i^{\frac{n-r}{2}-\frac{3}{2}}\left(\frac{\langle x,y\rangle_{[Zq]}}{\sqrt{\langle x,x\rangle_{[Zq]}}\sqrt{\langle y,y\rangle_{[Zq]}}}\right)$$
$$\cdot P_{k-i}^{\frac{n-r}{2}+i-1}\left(\frac{\langle x,q\rangle_Z}{\sqrt{\langle x,x\rangle_Z}\sqrt{\langle q,q\rangle_Z}}\right) P_{k-i}^{\frac{n-r}{2}+i-1}\left(\frac{\langle y,q\rangle_Z}{\sqrt{\langle y,y\rangle_Z}\sqrt{\langle q,q\rangle_Z}}\right)$$

where $c_{k,i}^{n,r}$ are positive constants.

Proof. We use the addition theorem for Gegenbauer polynomials. Let $\alpha > 0$ and $k \in \mathbb{N}$, then for angles γ, θ, τ

$$P_k^\alpha(\cos\theta\cos\tau + \sin\theta\sin\tau\cos\gamma)$$
$$= \sum_{i=0}^{k} c_{k,i}^\alpha (\sin\theta)^i (\sin\tau)^i P_i^{\alpha-\frac{1}{2}}(\cos\gamma) P_{k-i}^{\alpha+i}(\cos\theta) P_{k-i}^{\alpha+i}(\cos\tau),$$

where $c_{k,i}^\alpha$ are some positive constants that depend on α, k, i. More details about the formula can be found in [105]. Define

$$\cos\gamma = \frac{\langle x,y\rangle_{[Zq]}}{\sqrt{\langle x,x\rangle_{[Zq]}}\sqrt{\langle y,y\rangle_{[Zq]}}}$$
$$\cos\theta = \frac{\langle x,q\rangle_Z}{\sqrt{\langle x,x\rangle_Z}\sqrt{\langle q,q\rangle_Z}}, \quad \sin\theta = \sqrt{1-\cos^2\theta},$$
$$\cos\tau = \frac{\langle y,q\rangle_Z}{\sqrt{\langle y,y\rangle_Z}\sqrt{\langle q,q\rangle_Z}}, \quad \sin\tau = \sqrt{1-\cos^2\tau}.$$

Now, expanding the expressions for the sines above and using the inverse formula for block matrices, one can show that

$$\sin\theta = \sqrt{\frac{\langle x,x\rangle_{[Zq]}}{\langle x,x\rangle_Z}}$$

and

$$\cos\theta\cos\tau + \sin\theta\sin\tau\cos\gamma = \frac{\langle x,y\rangle_Z}{\sqrt{\langle x,x\rangle_Z}\sqrt{\langle y,y\rangle_Z}}.$$

Substituting the corresponding expressions and orders of Gegenbauer polynomials in the addition formula with $\alpha = \frac{n-r}{2} - 1$, the result follows. □

The proof of Proposition 3.13 is based on the addition theorem for Gegenbauer polynomials [105] and is related to the approach in Musin [151]. Namely, Musin [151] modifies the addition theorem for Gegenbauer polynomials to characterize p.d. kernels on S^{n-1} invariant under the action of $\mathrm{Stab}_{O_n}(Z)$ for a given $Z \in (S^{n-1})^r$.

3.6 Application to the kissing number problem

In this section we apply the results derived in this chapter to obtain bounds on the kissing number problem introduced in Chapter 2. In particular, we show how to implement problem (2.25). Consider the case $r = 0$. In this case the condition $F \in C\left((S^{n-1})^{r+2}\right)$ is 2-p.d. reduces to the condition F is p.d. To implement this condition, we use Schoenberg's theorem 3.1.

Schoenberg's theorem was used in [44] to obtain upper bounds on the spherical codes, and the kissing number in particular. We use it when $r = 0$ to substitute the condition $S \in C\left((S^{n-1})^{r+2}\right)^{O_n}$ in problem (2.25) with expansion (3.1). This results in an LP with infinitely many unknowns, the coefficients $c_i,\ i \in \mathbb{R}_+$. This LP is equivalent to the kissing number upper bound problem in [44] (see Section 3.7).

Now we move to the case $r \geq 1$ and use Theorem 3.8. We use (3.9) to substitute the condition $S \in C\left((S^{n-1})^{r+2}\right)^{O_n}$ in problem (2.25) with $r \geq 1$. Now we can rewrite problem (2.25).

$$\nu_r(\mathcal{G}_n) = \inf_{\phi,S} \quad \phi(1) + 1 \tag{3.15}$$

$$\begin{aligned} \text{s.t.}\quad & \phi(u) \leq -1, \text{ for all } u \in [-1, \tfrac{1}{2}], \\ & \phi(x_1^\intercal x_2) + \cdots + \phi(x_{r+1}^\intercal x_{r+2}) - \sigma\left(S(x_1, \ldots, x_{r+2})\right) \geq 0, \qquad (3.16) \\ & \text{for all } x_1, \ldots x_{r+2} \in S^{n-1}, \\ & \phi \in C([-1,1]),\ S \in C\left((S^{n-1})^{r+2}\right) \\ & S \text{ is of the form (3.9).} \end{aligned}$$

By restricting the last condition in this problem to a particular set of polynomials, we obtain a tractable conic problem. This procedure is described in the next subsection.

3.6.1 Implementation and numerical results

We use Stone-Weierstrass theorem and approximate continuous functions on $\mathcal{E}^{r+2}$ by polynomials of $\binom{d}{2}$ variables. Recall that each variable corresponds to an inner product between a pair of variables $x, y, z_1, \dots, z_r \in S^{n-1}$. Inspired by [12], we also restrict the functions $\{c_i\}_{i\in\mathbb{N}}$ to use in (3.9). For an r-variate vector of variables X and $d \in \mathbb{N}$, let $m^d(X)$ be the vector of all possible monomials in variables X of degrees up to d. Let $x, y \in S^{n-1}$ and $Z = [z_1, \dots, z_r] \in (\in S^{n-1})^r$. We denote

$$\chi := Z^\intercal x, \quad \upsilon := Z^\intercal y, \quad \Omega := Z^\intercal Z. \tag{3.17}$$

We define $c_i(\chi, \upsilon, \Omega)$ for every $i \in \mathbb{N}$ as follows:

$$c_i(\chi, \upsilon, \Omega) := m^{d_i}(\chi)^\intercal C_i(\Omega) m^{d_i}(\upsilon) \, |\Omega|^i \sqrt{(1-\chi^\intercal\Omega^{-1}\chi)\,(1-\upsilon^\intercal\Omega^{-1}\upsilon)}^{\,i}, \tag{3.18}$$

where $d \in \mathbb{N}$, $|\Omega|$ denotes the determinant of the matrix Ω, and for any $\Omega \in \mathcal{E}^r$ we have $C_i(\Omega) \succeq 0$.

Proposition 3.14. *Let $n \in \mathbb{N}_+$ and $r, N \in \mathbb{N}$. Let $x, y \in S^{n-1}$ and $Z = [z_1, \dots, z_r] \in \mathcal{S}$, where $\mathcal{S}$ is defined in* (3.8)*. Consider a function $S \in C\left((S^{n-1})^{r+2}\right)$ with representation* (3.9) *such that c_i is of the form* (3.18) *for $i \le N$, and $c_i = 0$ for all $i > N$. Then S is 2-p.d. and invariant under the action of O_n, and representation* (3.9) *is a polynomial in the inner products of $x, y, z_1, \dots, z_r$.*

Proof. Since $Z \in \mathcal{S}$, $\Omega \in \mathcal{E}^r$ is of full rank. By definition (3.18), c_i is a p.d. kernel a fixed $\Omega \in \mathcal{E}^r$ for all $i \le N$. Since the zero kernel is p.d., for all $i \in \mathbb{N}_+$, the functions c_i can be viewed as p.d. kernels on the projection subbundle f : $\left\{\begin{bmatrix} 1 & y^\intercal \\ y & Y \end{bmatrix} \in \mathcal{E}^{r+1} : Y \succ 0\right\} \to \{Y \in \mathcal{E}^r : Y \succ 0\}$. Therefore S is 2-p.d. and invariant under the action of O_n by Theorem 3.8. Finally, c_i is defined in such a way that each element in expansion (3.9) is a polynomial in the inner products of $x, y, z_1, \dots, z_r$. $\square$

For $r = 1$, we have $\Omega = 1$, therefore in case of $Q_1^{S^{n-1}}$, $C_i(\Omega)$ in (3.18) are symmetric matrices, and the condition $C_i(\Omega) \succeq 0$ can be addressed with any SDP solver. For $r = 2$, we define $z := z_1^\intercal z_2 \in [-1, 1]$. Then $C_i(\Omega) = C_i(z)$ are *polynomial matrices in variables* z; that is, matrices where each entry is a polynomial in $\mathbb{R}[z]$. Our implementation of the condition $C_i(z) \succeq 0$ for $r = 2$ is inspired by the following existing result.

Theorem 3.15 (Theorem 2 by Hol and Scherer [198])**.** *Let $M(z) : \mathbb{R}^n \to \mathbb{S}^k$ be a symmetric-valued polynomial matrix in variables z. Let $h_1, \dots, h_m \in \mathbb{R}[z]$ be such that their quadratic module is Archimedean, as in Definition 1.11. If $M(z) \succ 0$ for all $z \in S = \{z \in \mathbb{R}^n : h_1(z) \ge 0, ..., h_m(z) \ge 0\}$, then there exist $\varepsilon > 0$ and (not necessarily square) polynomial matrices $T_0(z), \dots, T_m(z)$ such that*

$$M(z) = T_0(z)^\intercal T_0(z) + \sum_{j=1}^{m} T_j(z)^\intercal T_j(z) h_j(z) + \varepsilon I_k.$$

Theorem 3.15 provides a certificate of positive definiteness of $M(z)$ on S: if $M(z)$ has the representation from the theorem, it is clearly positive definite for any $z \in S$. Based on this certificate, we require $C_i(z)$ to have the following form:

$$C_i(z) = T_i^1(z)^\intercal T_i^1(z) + (1-z^2)T_i^2(z)^\intercal T_i^2(z), \tag{3.19}$$

where $T_i(z),\ T_i^2(z)$ are polynomial matrices in variables z.

Representation (3.19) contains products of polynomial matrices. If we use this representation directly, we end up with a non-linear program which is not efficiently solvable. Therefore we reformulate representation (3.19) using PSD matrices. Our reformulation is based on another result by Hol and Scherer [198].

Theorem 3.16 (Lemma 1 by Hol and Scherer [198])**.** *Let $M(z) : \mathbb{R}^n \to \mathbb{S}^k$ be a symmetric-valued polynomial matrix in variables z. Then $M = T(z)^\intercal T(z)$ for some (not necessarily square) polynomial matrix $T(z)$ if and only if there exist $d > 0$ and a matrix $H \in \mathbb{S}^{k\binom{n+d}{d}}$ such that*

$$M(z) = (m^d(z) \otimes I_k)^\intercal\ H\ (m^d(z) \otimes I_k),\ \textit{and}\ H \succeq 0. \tag{3.20}$$

Now, (3.18) implies that $C_i(z)$ is a matrix of the size $\binom{2+d_i}{d_i}\times\binom{2+d_i}{d_i}$. Based on Theorem 3.16, we obtain:

$$C_i(z) := M_i^1(z)^\intercal H_i^1 M_i^1(z) + (1 - z^2)M_i^2(z)^\intercal H_i^2 M_i^2(z), \tag{3.21}$$

where $H_i^1 \succeq 0, H_i^2 \succeq 0$, and $M_i^j(z) = m^{d_i^j}(z) \otimes I_{\binom{2+d_i}{d_i}}$ for some $d_i^j > 0$ and all $j \in \{1, 2\}$.

Next, we provide the details on the implementation of problem (3.15) using expansion (3.9) and coefficients c_i (3.18) with C_i as in (3.21). For $Q_0^{S^{n-1}}, Q_1^{S^{n-1}}, Q_2^{S^{n-1}}$ we use Gegenbauer polynomials of degrees up to $N_0 = 24, N_1 = 12, N_2 = 4$, respectively. For $Q_1^{S^{n-1}}$ we set the degree d_i in (3.18) for each $i \in [0, 1, \ldots, N_1]$ to $2N_1 - 2i$. For $Q_2^{S^{n-1}}$ for each $i \in [0, 1, \ldots, N_2]$, we set $d_i^1 = N_2 - i$, $d_i^2 = N_2 - i - 1$ and $d_i = 2(N_2 - i)$.

We set some restrictions on the structure of the problem so that the total degree of the polynomial in $Q_1^{S^{n-1}}$ is $2N_1$, and the degree of each variable is at most N_1; the total degree of the polynomial in $Q_2^{S^{n-1}}$ is $4N_2$, and the maximal degree of each variable is $2N_2$. This structure allows us to generate fewer monomials while solving polynomial optimization problems. We also partially restrict the inequality constraint (3.16). For $Q_0^{S^{n-1}}$, this inequality can be replaced by equality without loss of generality due to the shape of the objective: the constraint will hold as equality in optimum. For $Q_1^{S^{n-1}}, Q_2^{S^{n-1}}$, we replace the non-negativity condition in (3.16) by simpler, though stronger, conditions. For $Q_2^{S^{n-1}}$, we require that the left-hand side in (3.16) is a sum of squares of polynomials (SOS). For $Q_1^{S^{n-1}}$, we say that the left-hand side in (3.16) equals $q_1 + (1 - (x^\intercal z_1)^2)q_2$, where q_1, q_2 are SOS.

For $Q_1^{S^{n-1}}$, the structural restrictions do not influence the rounded values of the bounds, but the problem becomes substantially smaller. However, for $Q_2^{S^{n-1}}$, we could solve the restricted problems up to N_2=7, but the restrictions cause feasibility issues. Therefore we present the bounds for N_2=4 only.

Further, we undertake several steps which make the problem smaller and more numerically stable. First, Gegenbauer polynomials are normalized so that $P_i^{\frac{n-r}{2}-1}(1) = 1$ for all $i \in [N]$, this can be done without loss of generality and provides more numerically stable optimization results. Next, we exploit invariance of problem (3.15) under all possible permutations of the corresponding variables $x, y, z_1, z_2 \in S^{n-1}$. This allows us to decrease the number of constraints in the problem; when equating polynomials to SOS in (3.16), we only generate the constraints corresponding to the orbits of monomials.

Table 3.1 shows the bounds we obtain. The bounds from problem (3.15) in the table are rounded up. We do not use the exact arithmetic to check the feasibility of the bounds, but we verify that all SDP constraints are satisfied, and equality constraints violations are of the order 10^{-4} or smaller. The bound for $Q_2^{S^{n-1}}$ cannot be computed for $n \leq 3$ since in this case $n < r + 2$, and Proposition 3.8 does not apply.

The bounds we could compute are not better than any of the best existing bounds, however, for $n \leq 16$ the results for $Q_1^{S^{n-1}}$ are close to the best bounds, while obtained in a much shorter time than in [136, 143]. The main reason is that our problems are by construction smaller than the ones from [136, 143], due to the structural restrictions mentioned earlier in this section. Also, we do not use the solver SDPA-GMP with arbitrary precision arithmetic in contrast to [136, 143]. The bounds obtained with SDPA-GMP can be feasible with high precision at the cost of longer running times. The running times of our problems would increase if we solved them with high precision.

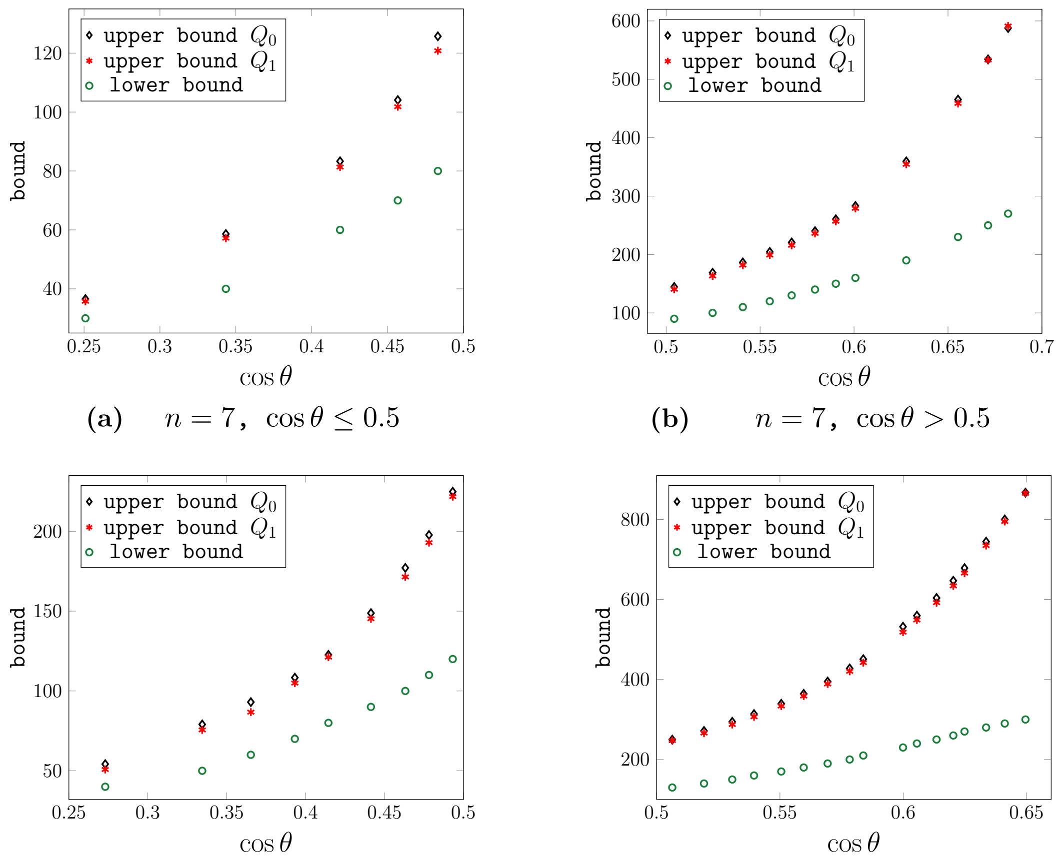
Table 3.1 – The upper bounds from problem (3.15) and the best existing upper and lower bounds on the kissing number. The bounds for $Q_1^{S^{n-1}}$ which conicide with the best bounds are marked in bold.

Dimension n	$Q_0^{S^{n-1}}$ 24 Gegenb. polynom.	$Q_1^{S^{n-1}}$ 12 Gegenb. polynom.	$Q_2^{S^{n-1}}$ 4 Gegenb. polynom.	Best upper bound	Best lower bound [154]
3	13.16	**12.54**	-	12[205]	12
4	25.56	**24.50**	25.77	24 [150]	24
5	46.34	45.16	47.74	44 [143]	40
6	82.64	**78.90**	85.93	78 [12]	72
7	140.17	136.30	154.90	134 [143]	126
8	240.00	**240.00**	297.21	240 [125, 164]	240
9	380.10	371.75	724.69	364 [136]	306
10	595.83	580.68	4,525.45	554 [136]	500
11	915.39	899.68	infeasible	870 [136]	582
12	1,416.10	1,384.68	infeasible	1,357 [136]	840
13	2,233.64	2,152.53	infeasible	2,069 [136]	1,154
14	3,492.22	3,307.45	infeasible	3,183 [136]	1,606
15	5,431.03	5,043.03	infeasible	4,866 [136]	2,564
16	8,313.79	7,863.00	infeasible	7,355 [136]	4,320
17	12,218.68	12,050.45	infeasible	11,072 [136]	5,346
18	17,877.07	18,008.00	infeasible	16,572 [136]	7,398
19	25,900.79	26,672.27	infeasible	24,812 [136]	10,668
20	37,974.01	39,554.23	infeasible	36,764 [136]	17,400
21	56,851.69	59,458.13	infeasible	54,584 [136]	27,720
22	86,537.49	88,326.01	infeasible	82,340 [136]	49,896
23	128,095.86	130,270.28	infeasible	124,416 [136]	93,150
24	196,560.00	**196,560.02**	infeasible	196,560 [125, 164]	196,560
25	278,364.38	282,690.33	infeasible	278,083 [177]	197,040
26	396,977.00	403,772.00	infeasible	396,447 [177]	198,480
30	1,653,914.18	1,749,936.18	infeasible		219,008
35	10,510,137.84	13,835,411.99	infeasible		370,892
40	infeasible	infeasible	infeasible		1,063,216
Approx. solution time	$\leq$ 1 sec.	10 sec.	25 sec.	[136]: 12 hours, [143]: $\geq$1 week	-

Since the sum of copositive kernels is copositive, we could set the kernel in problem (3.15) to be the sum of the functions from $Q_0^{S^{n-1}}, Q_1^{S^{n-1}}$ and $Q_2^{S^{n-1}}$. Table 3.1 suggests that this could provide potentially stronger bounds, but the resulting problems are numerically unstable and do not improve on the bounds from Table 3.1. Nevertheless, using this approach we can compare the optimization problems for our bounds and the existing SDP upper bound used in [12, 136, 143]. This is done in the next section.

To conclude this section, we recall that the kissing number problem is a particular case of the spherical codes problem. As it was mentioned before, in the spherical codes problem we are interested in the maximum number of points on the unit sphere in $\mathbb{R}^n$ for which the pairwise angular distance is not smaller than some value θ. The kissing number problem corresponds to $\theta = \frac{\pi}{3}$. Figure 3.1 shows how the bounds from $Q_0^{S^{n-1}}$ and $Q_1^{S^{n-1}}$ change when θ changes. For a more informative comparison, we add lower bounds computed using the algorithm by Roebers [193].

Figure 3.1 – Upper bounds using $Q_0^{S^{n-1}}$ and $Q_1^{S^{n-1}}$ and lower bounds on the size of the spherical codes for various angular distances θ.



(a) $n = 7,\ \cos\theta \le 0.5$

(b) $n = 7,\ \cos\theta > 0.5$

(c) $n = 8,\ \cos\theta \le 0.5$

(d) $n = 8,\ \cos\theta > 0.5$

We have examined $3 \leq n \leq 8$, and the results exhibit the same pattern for each n. We choose $n \in \{7, 8\}$ since for these dimensions the difference in performance is most visible. The choice of θ is motivated by the lower bound algorithm in [193]: for a given number of points on the unit sphere (which is equal to the lower bound in our case), the algorithm finds a feasible allocation of the points with the minimum distance θ. We run the algorithm for 30 min. using the lower bounds $[20, 40, 60, \dots, 400]$ to obtain θ for each of these lower bounds. Next, we solve problem (3.15) replacing $\frac{1}{2}$ in the first constraint with the corresponding $\cos\theta$. We present the bounds for which all SDP constraints are satisfied, and equality constraints violations are of the order 10^{-4} or smaller.

3.7 Connection to the existing upper bound approaches

In this section we provide some remarks on the connection between our relaxations and the existing LP bound by Delsarte, Goethals and Seidel [44] and SDP bound by Bachoc and Vallentin [12]. We start by showing that the LP bound equals our bound for $Q_0^{S^{n-1}}$. As explained in Section 3.6.1, inequality (3.16) can be replaced by equality. That is, in (3.16) we can consider p.d. kernels instead of the sum of a non-negative and p.d. kernel. Hence our formulation for $Q_0^{S^{n-1}}$ is equivalent to

$$\begin{aligned} \inf\ & \sum_{k=0}^{N_0} a_k + 1 \\ \text{s. t. } & \sum_{k=0}^{N_0} a_k P_i^{\frac{n}{2}-1}(u) \leq -1, \text{ for all } u \in [-1, \tfrac{1}{2}], \\ & a_k \geq 0, \text{ for all } k \in \{0, \dots, N_0\}. \end{aligned}$$

From the shape of the objective, using $P_0^{\frac{n}{2}-1}(u) = 1$, it is clear that in optimality $a_0 = 0$. Therefore, our bound from $Q_0^{S^{n-1}}$ coincides with the LP bound by Delsarte, Goethals and Seidel [44].

Now let us consider the SDP bound. This bound is formulated in Theorem 4.2 from [12]. Fix $d > 0$, then

$$\begin{aligned} \alpha(\mathcal{G}_n) \leq \inf\ & \sum_{k=1}^{d} a_k + b_{11} + \langle F_0, S_0^n(1,1,1)\rangle + 1 \\ \text{s. t. } & \sum_{k=1}^{d} a_k P_i^{\frac{n}{2}-1}(u) + 2b_{12} + b_{22} + 3\sum_{k=0}^{d} \langle F_k, S_k^n(u,u,1)\rangle \leq -1, \\ & \quad \text{for all } u \in [-1, \tfrac{1}{2}] \\ & a_k \geq 0, \text{ for all } k \in \{1, \dots, d\}, \end{aligned} \tag{3.22}$$

$$F_k \succeq 0, \text{ for all } k \in \{0,\dots,d\},$$

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{bmatrix} \succeq 0, \tag{3.23}$$

$$b_{22} + \sum_{k=0}^{d} \langle F_k, S_k^n(u,v,t)\rangle \le 0,$$

$$\text{for all } u,v,t \in [-1,\tfrac{1}{2}] \text{ such that } 1+2uvt-u^2-v^2-t^2 \ge 0, \tag{3.24}$$

where $S_k^n(u,v,t) \in \mathbb{S}^{d-k}$ are defined for all $k \in \{0,\dots,d\}$ by

$$(S_k^n)_{ij}(u,v,t) = \tfrac{1}{6} \sum_{\pi \in Sym(3)} (Y_k^n)_{ij}(\pi[u,v,t]), \text{ for all } i,j \in \{0,\dots,d-k\}$$

$$(Y_k^n)_{ij}(u,v,t) = u^i v^j Q_k^n(u,v,t), \text{ for all } i,j \in \{0,\dots,d-k\}$$

$$Q_k^n(u,v,t) = P_k^{\frac{n}{2}-1}\left(\tfrac{t-uv}{\sqrt{1-u^2}\sqrt{1-v^2}}\right).$$

Hence for a positive semi-definite matrix F_k we have

$$\langle F_k, S_k^n(u,v,t)\rangle = \sigma\left(m^{d-k}(u)^\intercal F_k m^{d-k}(v) Q_k^n(u,v,t)\right).$$

To obtain an upper bound on $\alpha(\mathcal{G}_n)$ comparable to problem (3.22), we use the fact that the sum of copositive kernels is copositive. This implies we can relax (2.14) for $\alpha(\mathcal{G}_n)$ using the sum of $Q_0^{S^{n-1}}$ and $Q_1^{S^{n-1}}$ of degree d as follows:

$$\alpha(\mathcal{G}_n) \le \inf \sum_{k=1}^{d} a_k + g(1) + 1 \tag{3.25}$$

$$\text{s. t. } \sum_{k=1}^{d} a_k P_k^{\frac{n}{2}-1}(u) + g(u) \le -1, \text{ for all } u \in [-1,\tfrac{1}{2}],$$

$$a_k \ge 0, \text{ for all } k \in \{1,\dots,d\},$$

$$C_k \succeq 0, \text{ for all } k \in \{0,\dots,d\},$$

$$\sum_{k=0}^{d} \sigma\left(m^{d-k}(u)^\intercal C_k m^{d-k}(v) Q_k^n(u,v,t)\right) - g(u) - g(v) - g(t) \le 0,$$

$$\text{for all } u,v,t \in [-1,1] \text{ such that } 1+2uvt-u^2-v^2-t^2 \ge 0, \tag{3.26}$$

where the last constraint follows from Proposition 3.14 with r=1 and $N = d$.

Problems (3.22) and (3.25) are formulated using similar sets of variables and constraints, however our comparison of these problems is inconclusive because of the following reasons. The variables b_{11}, b_{12}, b_{22} and the corresponding constraint (3.23) appear naturally from the dual of problem (3.22), but do not fit naturally into our framework. At the same time, the variables g do not appear in problem (3.22). However, the relation between the variables b_{11}, b_{12}, b_{22} and g is not straightforward since

g is a polynomial while b_{11}, b_{12}, b_{22} are constant terms. We expect the bound (3.25) to be weaker (3.22) since working with constant terms might be less restrictive than working with a continuous function. Moreover, the requirement we impose by (3.26) is stronger than (3.24) imposed in (3.22).

Problem (3.15) where the sum of $Q_0^{S^{n-1}}$, $Q_1^{S^{n-1}}$ and $Q_2^{S^{n-1}}$ is used could potentially provide stronger bounds than problem (3.22) by Bachoc and Vallentin [12], but those problems are even harder to compare. For the future, it would be interesting to see whether one can combine our approach with the currently best approaches for the kissing number problem by Bachoc and Vallentin [12] and Pfender [177].

Acknowledgements
We would like to thank Lorenz Roebers for providing his code to compute the lower bounds in Section 3.6.1

CHAPTER 4

Copositive certificates of non-negativity for polynomials on unbounded sets

4.1 Introduction

Certificates of non-negativity are fundamental tools in optimization, and they underlie powerful algorithmic techniques for various types of optimization problems. Commonly used certificates of non-negativity of polynomials on basic semialgebraic sets include the classical Pólya's Positivstellensatz [83], the more modern Schmüdgen's Positivstellensatz [199], and Putinar's Positivstellensatz [183]. Herein, we use the terms Positivstellensatz and certificate of non-negativity interchangeably (see Section 1.2.2 in Chapter 1 for more details).

To illustrate the concept of a certificate of non-negativity, let $p, h_1, ..., h_m$ be polynomials. Assume we would like to know whether p is non-negative on the set $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, ..., h_m(x) \geq 0\}$. If there exist a polynomial $F(x, u)$ non-negative for all $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m_+$ such that $p(x) = F(x, h_1(x), \ldots, h_m(x))$, then we are sure that p is non-negative on S. We call such F *a certificate of non-negativity* for p. For instance, one could have $F(x, u) = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x)u_i$, where $\sigma_0, \ldots \sigma_m$ are sums-of-squares (SOS) polynomials [19]. From Putinar's Positivstellensatz [183], it is known that the latter certificate exists for p on S if the quadratic module generated by $h_1 \ldots, h_m$ is Archimedean and $p(x) > 0$ for all $x \in S$.

In this chapter we study certificates of non-negativity based on *copositivity*. Polynomials that are non-negative on the non-negative orthant are called *copositive* polynomials [see, e.g. 22]. More specifically, one can show that p is non-negative on $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, ..., h_m(x) \geq 0\}$ by demonstrating that for some $k \geq 0$

$$(1 + e^\intercal y + e^\intercal z)^k p(y - z) = F(y, z, h_1(y - z), \ldots, h_m(y - z)), \quad \text{where } F(y, z, u) \text{ is copositive.} \tag{4.1}$$

Such F is called a *copositive certificate of non-negativity* of p on S. For any $x \in S$, taking $x^+ = \max(x, 0)$ and $x^- = -\min(x, 0)$, where the maximum and minimum are taken component-wise, we have that $x^+, x^- \geq 0$ and therefore,

$$\begin{aligned} p(x) &= p(x^+ - x^-) \\ &= F(x^+, x^-, h_1(x^+ - x^-), \dots, h_m(x^+ - x^-))(1 + e^\intercal x^+ + e^\intercal x^-)^{-k} \\ &= F(x^+, x^-, h_1(x), \dots, h_m(x))(1 + e^\intercal |x|)^{-k} \geq 0, \end{aligned}$$

as F is copositive. As before, we use e to denote the vector of all-ones of appropriate dimension. For $x \in \mathbb{R}^n$, $|x|$ stands for the component-wise absolute value of x (i.e., $|x|_i = |x_i|, i = 1, \dots, n$). In Theorems 4.8 and 4.11 we prove the existence of copositive certificates under mild assumptions which hold generically. In particular, no compactness or similar properties are assumed.

One essential property of the copositive certificates of non-negativity we propose is that the degree of F in (4.1) is known a priori. Namely, this degree is bounded by the maximum of the degree of p and twice the degree of the polynomials defining the set S. As a consequence, questions on the non-negativity of polynomials on generic basic semialgebraic sets reduce to finding a copositive polynomial satisfying (4.1) of small and, more importantly, known degree. This result is in line with recent results by Huq [89] on small copositive extended formulations for some combinatorial problems.

Optimization over the cone of copositive polynomials is hard [148]; however, this cone has been well studied. In particular, there exists a plenty of tractable approximations to it [see, e.g., 27, 92, 120, 134, 162, 231], as well as several certificates of copositivity [for instance, 47, 83]. The main benefit of our copositive certificates of non-negativity is their ability to translate results known exclusively for copositive polynomials to more general basic closed semialgebraic sets (see Section 4.1.2 for a more detailed explanation of our contributions).

4.1.1 Certificates of non-negativity and polynomial optimization

Classically, certificates of non-negativity based on SOS and non-negative coefficients (SOS-certificates), have been used to solve/approximate polynomial optimization (PO) problems [114, 172, 210]. PO encompass a wide variety of optimization problems including combinatorial and some non-convex optimization problems. Pólya's, Schmüdgen's, and Putinar's Positivstellensatzen are examples of SOS certificates, and their applications in PO are illustrated in recent works [e.g., 35, 84, 97, 114, 116, 168, 172, 173, among numerous others]. Searching for a given SOS-certificate of non-negativity of a fixed degree translates into solving a number of linear matrix inequalities (LMI). As the degree of the SOS-certificate is not known a priori, this method constructs a hierarchy of LMI approximations to the underlying problem. That is, optimization problems with a linear objective and LMI constraints [see 19]. LMI problems usually have the form of a linear program (LP), second-order cone program (SOCP) or semidefinite program (SDP), which can be solved to a given precision using interior-point methods [see 188]. The main drawbacks of using SOS certificates are the exponential growth of the LMI hierarchies in terms of the certificate's degree and the lack of SOS-certificates for many interesting cases. To guarantee the existence of SOS-certificates, usually some form of compactness is needed [see 194, for a detailed analysis].

To deal with the fast-growing size of SOS certificates, one could use subsets of SOS polynomials whose LMI reformulations do not result in full dimensional SDPs. For instance, in certain cases the structure of the problem allows arguing that sparse SOS certificates can be used as not all monomials have to be present in the certificates. This approach results in smaller convergent approximations to PO problems over some compact sets [examples are presented in 96, 115, 218, 221]. For non-structured problems, one could use scaled diagonally dominant sums-of-squares (SDSOS) instead of classical SOS. SDSOS are a type of SOS which result in LP or SOCP relaxations of PO problems. Such relaxations are computationally cheaper than SDPs and provide valid bounds [3] on PO problems. However these bounds are either not proven to converge or require the use of additional methods to ensure convergence [2].

Another way to deal with the flaws of SOS certificates would be to replace SOS in the expressions of certificates with different non-negative polynomials. Some existing examples include hyperbolic polynomials and non-negative circuit polynomials. The set of hyperbolic polynomials contains the set of SOS polynomials as a strict subset. Hence replacing SOS with hyperbolic polynomials provides hyperbolic programming relaxations of PO [197], which could potentially result in stronger bounds or faster convergence compared to classical SOS relaxations. Hyperbolic programs can be solved using interior-point methods, but efficient hyperbolic solvers are still under development, and the hyperbolic cone is not yet fully understood [189]. Non-negative circuit polynomials form neither a subset nor a superset of the cone of SOS polynomials. The relation between the two sets of polynomials depends on the degree and the number of variables [91]. Certificates based on non-negative circuit polynomials result in geometric programming relaxations of PO problems [59, 219] which converge under certain Archimedean conditions.

Given the key role that compactness plays for SOS certificates and their alternatives, a question that has attracted much research attention is which certificates exist on non-compact sets. In particular, Marshall [141], Powers [181] derive certificates of non-negativity for the case in which the underlying domain is a cylinder with a compact cross-section. Nguyen and Powers [158] derive certificates of non-negativity for the case in which the underlying domain is a strip or a half-strip. For more general settings, Demmel et al. [45], Marshall [140], Nie et al. [161], Vui and So'n [217], Wang [220] provide certificates of non-negativity, based on Putinar's and Schmüdgen's Positivstellensatzen, that do not require the underlying set to be compact. The latter certificates exploit gradient, Jacobian and KKT ideals. More recently, Jeyakumar et al. [95] have provided certificates of non-negativity for non-compact semialgebraic sets if a certain modification of the set is compact. Following the results in [95], Jeyakumar et al. [96] provide a certificate of non-negativity, based on Putinar's Positivstellensatz, for coercive polynomials over possibly unbounded semialgebraic sets. Also recently, Guo et al. [79] have derived conditions under which Schmüdgen's Positivstellensatz can be used to certify the non-negativity of a polynomial on a possibly unbounded convex set. Two other examples of research in this direction that are related to the results in this chapter, are the works of Putinar and Vasilescu [185] and Dickinson and Povh [48].

4.1.2 Contributions

Now we give more details about our contributions.

Existence of copositive certificates of non-negativity

A common assumption for the existence of SOS-certificates of non-negativity for a polynomial p is the positivity of p. As we are interested in certifying the non-negativity of a polynomial on a given set S that might be unbounded, we request p to be not only positive on S, but also "strongly positive" on S (see Definition 4.5). In Theorem 4.11 we show that, given polynomials $h_1, \dots, h_m$ of degree at most d such that $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ is non-empty, for all p strongly positive on S, we always have a copositive certificate F as given by (4.1) of degree $\max\{2d, \deg(p)\}$. In particular, $k = \max\{2d - \deg(p), 0\}$ in (4.1).

As we are interested in certifying the copositivity of F, and certificates of copositivity usually exist for the interior of the cone of copositive polynomials, we show that for the compact case we can construct copositive certificates that lie in this interior (see Theorem 4.11). We also provide several equivalent characterizations of the interior of the cone of copositive polynomials (see Corollary 4.19).

In Section 4.3.2 we show that the strong positivity condition is generic since it is implied by a particular generic algebraic condition on S considered in [78, 79, 159].

Structure-rich certificates of non-negativity

The copositive approach we propose allows constructing a certificate of non-negativity from any certificate of copositivity (and any certificate of non-negativity on the non-negative orthant or standard simplex, in particular). This provides a universal procedure to obtain new certificates with desired properties on generic basic semialgebraic sets. To illustrate this approach, in Section 4.4, we construct two new certificates of non-negativity on compact sets which do not require full-dimensional SOS polynomials. The special structure of these certificates provides computational advantages when compared to classical SOS-based certificates. Notice that, even though we focus on SOS, our methods could be used to obtain certificates of non-negativity based on circuit, hyperbolic polynomials and/or any general type of certificate of non-negativity on $\mathbb{R}^n$ (see Corollary 4.28).

Besides the new certificates, we also obtain an elementary proof of the seminal theorem by Handelman [82] and an alternative proof of Schmüdgen's Positivstellensatz [199] which shortcuts the proof by Schweighofer [206].

Applications to polynomial optimization
Our contribution to PO is twofold. On the one hand, our certificates allow us to apply to generic basic semialgebraic sets a variety of results which are valid only for optimization over the non-negative orthant [see, e.g., 27, 47, 120, 134]. In particular, we can use both inner and outer approximations to the cone of copositive polynomials to obtain LMI hierarchies of upper and lower bounds for generic PO problems (see, Section 4.5). This is in contrast with commonly used LMI hierarchies which only provide lower bounds for (minimization) PO problems [see, 6, 19]. On the other hand, under our assumptions a PO problem can be reformulated as an optimization problem over copositive polynomials of a fixed degree. This result connects copositive optimization and PO in general and advances the ongoing research on copositive reformulations of optimization problems. This line of research started with the work by Bomze et al. [25] showing that (potentially non-convex) standard quadratic optimization problems can be reformulated as copositive optimization problems. Further, Burer [28], Arima et al. [8], Bai et al. [14], Bomze and Jarre [21], Burer and Dong [29], Dickinson et al. [51], Eichfelder and Povh [61], Peña et al. [175], Xia and Zuluaga [226], among many others, considered copositive reformulations of more general PO problems.

The rest of the chapter is organized as follows. In Section 4.2, we introduce the notation used throughout the chapter and some necessary basic results. Section 4.3 contains the main Theorems 4.8 and 4.11 which connect non-negative and copositive polynomials under certain assumptions that hold generically. In Section 4.4 we derive two new certificates of non-negativity. In Section 4.5 we show how to reformulate PO problems and obtain tractable upper and lower bounds on these problems using Theorem 4.11. Section 4.6 shows alternative proofs of Handelman's [82], Schmüdgen's [199] Positivstellensatzen. We conclude in Section 4.7 with some closing remarks.

4.2 Preliminaries

For $p \in \mathbb{R}[x]$, with degree $\deg p = d$, let $C_{d,\alpha}$ denote the multinomial coefficient $C_{d,\alpha} := \frac{d!}{(d-e^\intercal\alpha)!\alpha_1!\cdots\alpha_d!}$ and $\mathbb{N}^n_d := \{\alpha \in \mathbb{N}^n : e^\intercal\alpha \leq d\}$. Then, given $p(x) \in \mathbb{R}[x]$ with $\deg p = d$, we can write $p(x) = \sum_{\alpha\in\mathbb{N}^n_d} C_{d,\alpha}p_\alpha x^\alpha$ for some $p_\alpha \in \mathbb{R}$, where $x^\alpha := x_1^{\alpha_1}\cdots x_n^{\alpha_n}$. We define $\|p\| = \max\{|p_\alpha| : \alpha \in \mathbb{N}^n, e^\intercal\alpha \leq d\}$.

Lemma 4.1. *Let $p \in \mathbb{R}[x]$. For any $x \in \mathbb{R}^n$ we have*

$$p(x) \leq \|p\|(1 + e^\intercal|x|)^{\deg p}.$$

Proof. Given $p \in \mathbb{R}[x]$ with $\deg p = d$, and $x \in \mathbb{R}^n$ we have

$$p(x) \leq \sum_{\alpha\in\mathbb{N}^n_d} C_{d,\alpha}|p_\alpha||x|^\alpha \leq \|p\| \sum_{\alpha\in\mathbb{N}^n_d} C_{d,\alpha}|x|^\alpha \leq \|p\|(1 + e^\intercal|x|)^d.$$

□

For any $S \subseteq \mathbb{R}^n$, we define

$$\mathcal{P}(S) = \{p \in \mathbb{R}[x] : p(x) \geq 0 \text{ for all } x \in S\},$$

as the set of polynomials non-negative on S. Similarly, we define

$$\mathcal{P}^+(S) = \{p \in \mathbb{R}[x] : p(x) > 0 \text{ for all } x \in S\},$$

as the set of polynomials positive on S. Furthermore, let $\mathcal{P}_d(S) := \mathcal{P}(S) \cap \mathbb{R}_d[x]$ (resp. $\mathcal{P}^+_d(S) := \mathcal{P}^+(S) \cap \mathbb{R}_d[x]$) denote the set of polynomials of degree at most d that are non-negative (resp. positive) on S. In this chapter we usually deal with $\operatorname{int}\mathcal{P}_d(S)$ the interior of $\mathcal{P}_d(S)$. Since $\mathbb{R}_d[x]$ is a finite-dimensional vector space and $\mathcal{P}_d(S)$ is convex, the interior and the algebraic interior of $\mathcal{P}_d(S)$ coincide [see, e.g., 87, Chapter 17]. This fact is formally stated in Lemma 4.2.

Lemma 4.2. *Let $S \subseteq \mathbb{R}^n$. Then*

$$\operatorname{int}\mathcal{P}_d(S) = \{p \in \mathcal{P}_d(S) : \textit{for all } q \in \mathbb{R}_d[x] \textit{ there exists } \varepsilon > 0 \textit{ such that } p - \varepsilon q \in \mathcal{P}_d(S)\}.$$

Central to our discussion are copositive polynomials [22] and sum-of-squares polynomials (SOS) [19]. A polynomial is copositive if it is non-negative on the non-negative orthant. Formally, a polynomial $p \in \mathbb{R}_d[x]$ is copositive if $p \in \mathcal{P}_d(\mathbb{R}^n_+)$.

We call a set of the form $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}$ where $h_1, \ldots, h_m \in \mathbb{R}[x]$ a *basic semialgebraic set*.

4.2.1 Strong positivity

One of the assumptions for the existence of classical SOS-certificates of non-negativity of p on S like the ones derived by Schmüdgen [199], Putinar [183], and Handelman [82], is the positivity of p on S. In these classical theorems the assumptions also imply compactness of the semialgebraic set S. Notice that

$$S \text{ compact} \Rightarrow \operatorname{int}\mathcal{P}_d(S) = \mathcal{P}_d^+(S). \tag{4.2}$$

When S is not compact, $\mathcal{P}_d^+(S) \supset \operatorname{int}\mathcal{P}_d(S)$. For example, the polynomial $p(x) := 1$ belongs to $\operatorname{int}\mathcal{P}_d(S)$ only when S is compact. We are interested in certifying the non-negativity of a polynomial p on a given basic semialgebraic set S that might be unbounded (i.e., not compact). In the existing results over non-compact sets, the positivity on S alone is not enough. Usually, assumptions on the behaviour of *p at infinity*; that is, the behaviour of p on the "directions" in which S becomes unbounded, are necessary [see, e.g., 184, 190]. Our certificates are not an exception to this rule, they exist for a subset of $\operatorname{int}\mathcal{P}_d(S)$ with a certain behavior at infinity which we describe next.

Given a polynomial $p \in \mathbb{R}_d[x]$, let $\tilde{p}(x)$ denote the homogeneous component of $p(x)$ of the highest total degree. That is, $\tilde{p}(x)$ is obtained by dropping from $p(x)$ all the terms whose total degree is less than $\deg p$. Notice that $\tilde{p}(x)$ determines the behavior of p at infinity. Namely, if $\tilde{p}(y) > 0$ for some $y \in \mathbb{R}^n$, then there is $t_0 \in \mathbb{R}$ such that $p(ty) > 0$ for all $t > t_0$, since the homogeneous component of the highest degree will eventually dominate the behavior of p. Similarly if $\tilde{p}(y) < 0$, p will become eventually negative in the y direction. However, if $\tilde{p}(y) = 0$, we do not know how $p(ty)$ behaves when t goes to infinity.

Definition 4.3. *Let $h_1, \dots, h_m, g_1, \dots, g_r \in \mathbb{R}[x]$ and let $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0,\ g_1(x) = 0, \dots, g_r(x) = 0\}$. We denote by $\tilde{S}$ the following set*

$$\tilde{S} = \left\{x \in \mathbb{R}^n : \tilde{h}_1(x) \geq 0, \dots, \tilde{h}_m(x) \geq 0,\ \tilde{g}_1(x) = 0, \dots, \tilde{g}_r(x) = 0\right\}. \tag{4.3}$$

Remark 4.4. *Note that from Definition 4.3 it follows that if $S' = S \cap \mathbb{R}^n_+$, then $\tilde{S}' = \tilde{S} \cap \mathbb{R}^n_+$, a fact that we will use throughout the chapter.*

Definition 4.5 (Strong positivity)**.** *We say that p is* strongly positive *on S when*

$$p \in \mathcal{P}^+(S) \text{ and } \tilde{p} \in \mathcal{P}^+(\tilde{S} \setminus \{0\}). \tag{4.4}$$

Strong positivity has been used in [184, Thm. 4.2] and [48, Property 3.5]. In particular, strong positivity on S is sufficient for the certificates of non-negativity in [48] to exist. Theorem 4.11 shows that for any semialgebraic set S, copositive certificates of non-negativity exists for polynomials that are strongly positive on S. Strongly positive polynomials belong to $\operatorname{int}\mathcal{P}_d(S)$, as formally stated in Proposition 4.6.

Proposition 4.6. *Let S be a basic semialgebraic set. Then,*

$$\left\{p \in \mathbb{R}_{=d}[x] : p \in \mathcal{P}_d^+(S),\ \tilde{p} \in \mathcal{P}_d^+(\tilde{S} \setminus \{0\})\right\} \subseteq \operatorname{int} \mathcal{P}_d(S)$$

Proof. The inclusion follows from Proposition 4.17 and Lemma 4.16b in Section 4.3.2. □

4.3 Copositive certificates of non-negativity

In this section, we prove our main results, namely, the existence of copositive certificates of non-negativity of the form (4.1) for all polynomials that are strongly positive on a basic semialgebraic set. We first consider the particular case in which the semialgebraic set of interest is defined by equality constraints only. Lemma 4.7 is the stepping stone to our copositive certificates of non-negativity.

Lemma 4.7. *Let $p, g_1, \ldots, g_m \in \mathbb{R}_d[x]$ be such that $g_1, \ldots, g_m \in \mathcal{P}(\mathbb{R}^n_+)$, and $S = \left\{x \in \mathbb{R}^n_+ : g_1(x) = 0, \ldots, g_m(x) = 0\right\}$ be non-empty. Let $p \in \mathbb{R}_{=d}[x]$ be such that $p \in \mathcal{P}^+(S)$ and $\tilde{p} \in \mathcal{P}^+(\tilde{S} \setminus \{0\})$. Then there are $F \in \operatorname{int} \mathcal{P}_d(\mathbb{R}^n_+)$ and $\alpha_j \in \mathbb{R}_{d-\deg g_j}[x]$ for $j = 1, \ldots, m$ such that*

$$p(x) = F(x) + \sum_{j=1}^{m} \alpha_j(x) g_j(x).$$

Since the proof of Lemma 4.7 is long and technical, we postpone it to Section 4.7.1. Lemma 4.7 is an important result on its own. However, we do not focus much attention on it since this lemma is inspired by a similar result by Peña et al. [176]. We do not present this result in the thesis as the manuscript [176] is unpublished, and the proofs in it are not full. Our proof of Lemma 4.7 is not related to the results in [176].

For ease of presentation, in what follows we often assume that $S \subseteq \mathbb{R}^n_+$. However, as shown in Section 4.3.1 this assumption can be made without loss of generality for compact sets and can be removed after doubling the number of variables for non-compact sets. Now, we prove the existence of copositive certificates under the extra assumption $S \subseteq \mathbb{R}^n_+$.

Theorem 4.8. *Let $p, h_1, \ldots, h_m \in \mathbb{R}[x]$, and $S = \left\{x \in \mathbb{R}^n_+ : h_1(x) \geq 0, \ldots, h_m(x) \geq 0\right\}$ be non-empty. Let $d_{max} = \max\{\deg h_1, \ldots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. Assume that $p \in \mathcal{P}^+_{2d_{\max}}(S)$ and $\tilde{p} \in \mathcal{P}^+_{2d_{\max}}(\tilde{S} \setminus \{0\})$. Then there exists $F \in \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m}_+)$ such that*

$$(1 + e^\intercal x)^{2d_{\max} - \deg p} p(x) = F(x, h_1(x), \ldots, h_m(x)).$$

Proof. Let $d_j = \deg h_j$, $j \in \{1,\dots,m\}$. Define $g_j : \mathbb{R}^{n+m} \to \mathbb{R}$ as $g_j(x,u) := \left((1+e^\intercal x)^{d_{max}-d_j} h_j(x) - u_j^{d_{max}}\right)^2$ for $j = 1,\dots,m$. Let

$$U = \left\{(x,u) \in \mathbb{R}_+^{n+m} : g_1(x,u) = 0, \dots, g_m(x,u) = 0\right\},$$

and let $q(x) := (1+e^\intercal x)^{2d_{\max}-\deg p} p(x)$. We apply Lemma 4.7 to U and q. To do this, we first check that the assumptions of the proposition hold. First, note that S non-empy implies U non-empty. Also, for any $(x,u) \in U$ we have $x \in S$ and thus $q(x) > 0$; that is, $q \in \mathcal{P}^+_{2d_{\max}}(U)$. Moreover, let $(z,v) \in \tilde{U}$. We have that $\tilde{g}_j(z,v) = (\tilde{h}_j(z)(e^\intercal z)^{d_{max}-d_j} - v_j^{d_{max}})^2$, $j = 1,\dots,m$. Hence, if $z = 0$, then $v = 0$. If $z \neq 0$, then $\tilde{h}_j(z)(e^\intercal z)^{d_{max}-d_j} = v_j^{d_{max}} \geq 0$ for $j = 1,\dots,m$. Therefore $z \in \tilde{S}$, which implies $\tilde{q}(z) = (e^\intercal z)^{2d_{\max}-\deg p}\tilde{p}(z) > 0$, since $\tilde{p} \in \mathcal{P}^+_{2d_{\max}}(\tilde{S} \setminus \{0\})$. Hence $\tilde{q} \in \mathcal{P}^+_{2d_{\max}}(\tilde{U} \setminus \{0\})$.

Lemma 4.7 implies that there is $G \in \operatorname{int} \mathcal{P}_{2d_{\max}}(\mathbb{R}_+^{n+m})$ and $\alpha_j \in \mathbb{R}$ such that

$$\begin{aligned} q(x) &= G(x,u) + \sum_{j=1}^m \alpha_j g_j(x,u) \\ &= G(x,u) + \underbrace{\sum_{j=1}^m \alpha_j \left(h_j(x)(1+e^\intercal x)^{d_{max}-d_j} - u_j^{d_{max}}\right)^2}_{*}. \end{aligned}$$

The pre-multipliers $\alpha_1,\dots,\alpha_m$ are real numbers since the degrees of $g_1,\dots,g_m$ are equal to $2d_{\max}$, as well as the degree of G. Hence, from the statement of Lemma 4.7, $\alpha_1,\dots,\alpha_m$ are polynomials of degree zero, i.e., constants. Since the right-hand side of the representation depends on x and u while the left-hand side depends on x only, u has to cancel out on the right-hand side. Since $\alpha_j \in \mathbb{R}$ and $g_j(x,u)$ depends on u_j only, the monomials with $u_1,\dots,u_m$ in the expression marked by * do not cancel out with each other. Thus all these monomials have to cancel out with monomials of G. Moreover, G cannot contain any other monomials with $u_1,\dots,u_j$. Therefore in all monomials in G the degrees of u_j are d_{max} or $2d_{max}$, for all $j \in \{1,\dots,m\}$.

Now, taking $u_j = ((1+e^\intercal x)^{d_{\max}-d_j} h_j(x))^{1/d_{\max}}$ for all $j \in \{1,\dots,m\}$, we obtain

$$(1+e^\intercal x)^{2d_{\max}-\deg p} p(x) = F(x, h_1(x)(1+e^\intercal x)^{d_{\max}-d_1}, \dots, h_m(x)(1+e^\intercal x)^{d_{\max}-d_m}),$$

where $F(x,u_1,\dots,u_m) := G(x, u_1^{1/d_{\max}}, \dots, u_m^{1/d_{\max}})$ is a polynomial. To finish, notice that $G \in \mathcal{P}_{2d_{\max}}(\mathbb{R}_+^{n+m})$ implies $F \in \mathcal{P}_{2d_{\max}}(\mathbb{R}_+^{n+m})$. $\square$

Next, we show a stronger version of Theorem 4.8 for compact sets. Namely, for compact sets the pre-multiplier $(1+e^\intercal x)^{2d_{\max}-\deg p}$ can be omitted, and the copositive certificate F belongs to the interior of the cone of copositive polynomials.

Theorem 4.9. *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$, and let $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ be non-empty. Define $d_{\max} = \max\{\deg h_1, \dots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. Let $M > 0$ be such that $S \subseteq \{x \in \mathbb{R}^n_+ : e^\intercal x \leq M\}$. If $p \in \mathcal{P}^+_{2d_{\max}}(S)$, then there exists $F \in \operatorname{int} \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m+1}_+)$ such that*

$$p(x) = F\left(x, h_1(x), \dots, h_m(x), M - e^\intercal x + \sum_{j=1}^m \Big((1+M)^{d_j} \|h_j\| - h_j(x)\Big)\right).$$

Proof. Since S is bounded, S is compact. Since $p \in \mathcal{P}^+_{2d_{\max}}(S) = \operatorname{int} \mathcal{P}_{2d_{\max}}(S)$ (recall (4.2)) and S is compact, there exists $\varepsilon > 0$ such that $q(x) = p(x) - \varepsilon(1 + e^\intercal x)^{2d_{\max}} \in \mathcal{P}^+_{2d_{\max}}(S)$. Let $d_j = \deg h_j$, $j \in \{1, \dots, m\}$. Define $g_j : \mathbb{R}^{n+m} \to \mathbb{R}$ as $g_j(x, u) := (h_j(x) - u_j)^2$ for $j = 1, \dots, m$. Also, let $\hat{M} = \sum_{j=1}^m (1+M)^{d_j} \|h_j\|$ and

$$U := \left\{(x, u, v) \in \mathbb{R}^{n+m+1}_+ : g_j(x, u) = 0, (\hat{M} + M - e^\intercal x - e^\intercal u - v)^2 = 0\right\}.$$

We apply Lemma 4.7 to U and q. To do this, we first check that the assumptions of the proposition hold. Let $x \in S$. For $j = 1, \dots, m$ let $u_j = h_j(x) \geq 0$, from Lemma 4.1. Let $v = \hat{M} + M - e^\intercal x - e^\intercal u \geq 0$, from the assumption on M. Thus U is non empty as $(x, u, v) \in U$. For any $(x, u, v) \in U$ we have $x \in S$ and thus $q(x) > 0$; that is, $q \in \mathcal{P}^+_{2d_{\max}}(U)$. Moreover, $(x, u, v) \in \tilde{U}$ implies $(x, u, v) \in \mathbb{R}^{n+m+1}_+$ and $-e^\intercal u - e^\intercal x = v$. Therefore $\tilde{U} = \{0\}$. Hence $\tilde{q} \in \mathcal{P}^+_{2d_{\max}}(\tilde{U} \setminus \{0\})$. Thus, Lemma 4.7 implies that there is $G \in \operatorname{int} \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m+1}_+)$, $\alpha_j \in \mathbb{R}_{2(d_{\max} - d_j)}[x, u, v]$, for all $j \in \{1, \dots, m\}$, and $\beta \in \mathbb{R}_{2d_{\max}-2}[x, u, v]$ such that

$$q(x) = G(x, u, v) + \sum_{j=1}^m \alpha_j(x, u, v) g_j(x, u) + \beta(x, u, v)(\hat{M} + M - e^\intercal x - e^\intercal u - v)^2.$$

Now, for any given x, take $u_j = h_j(x)$ for $j \in \{1, \dots, m\}$, and $v = \hat{M} + M - e^\intercal x - e^\intercal u$ to obtain

$$\begin{aligned} p(x) &= G\Big(x, h_1(x), \dots, h_m(x), \hat{M} + M - e^\intercal x - \sum_{j=1}^m h_j(x)\Big) + \varepsilon(1 + e^\intercal x)^{2d_{max}} \\ &= F\left(x, h_1(x), \dots, h_m(x), M - e^\intercal x + \sum_{j=1}^m \Big((1+M)^{d_j} \|h_j\| - h_j(x)\Big)\right), \end{aligned}$$

where $F(x, u, v) = G(x, u, v) + \varepsilon(1 + e^\intercal x)^{2d_{max}}$. By Lemma 4.2, $G \in \operatorname{int} \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m+1}_+)$, and $(1 + e^\intercal x)^{2d_{max}} \in \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m+1}_+)$ imply that $F \in \operatorname{int} \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m+1}_+)$.

□

Notice that (as mentioned in the proof above) under the assumptions of Theorem 4.9 we have that $M - e^\intercal x + \sum_{j=1}^m \Big((1+M)^{d_j} \|h_j\| - h_j(x)\Big) \geq 0$ for all $x \in S$, by Lemma 4.1. Therefore the representation of p we obtain in this theorem is clearly non-negative on S and defines a copositive certificate of non-negativity of p on S. Since F in Theorem 4.9 lies in the interior of the cone of copositive polynomials, we can use the existing certificates of copositivity to obtain new certificates of non-negativity on compact sets (see Section 4.4).

We would like to emphasize the differences between Theorem 4.9 and Schmüdgen's Positivstellensatz (Theorem 4.40). Let $S = \left\{x \in \mathbb{R}^n_+ : h_1(x) \geq 0, \dots, h_m(x) \geq 0\right\}$ be compact and let $p \in \mathcal{P}^+(S)$. Schmüdgen's Positivstellensatz shows that $p(x) = F(x, h_1(x), \dots, h_m(x))$ for some $F \in \mathbb{R}[x, u]$ such that $F(x, u) = \sum_{\alpha \in \{0,1\}^m} \sigma_\alpha(x) u^\alpha$, where σ_α is an SOS polynomial for all $\alpha \in \{0,1\}^m$. Such F is clearly copositive, however, the degree bounds for σ_α can be high and are nontrivial to compute [207]. On the contrary, Theorem 4.9 guarantees a representation $p(x) = F\Big(x, h_1(x), \dots, h_m(x), M - e^\intercal x + \sum_{j=1}^m \Big((1+M)^{d_j} \|h_j\| - h_j(x)\Big)\Big)$ of degree $2d_{\max}$, where $d_{\max} = \max\{\deg h_1, \dots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. Notice that the situation is similar when comparing Theorem 4.9 with Putinar's Positivstellensatz (presented in (4.9)): the degree bounds for the latter certificate are exponential in the degree of p and the number of variables [160].

4.3.1 Removing the condition $S \subseteq \mathbb{R}^n_+$

In Theorem 4.8 we require that the basic semialgebraic set S is a subset of the non-negative orthant. In general, the condition can be dropped after doubling the number of variables; that is, by using the well known substitution $x_i = y_i - z_i$, with $y_i, z_i \in \mathbb{R}^+$ for each $i \in \{1, \dots, n\}$. Next, in Lemma 4.10 we show how to do this, while maintaining the validity of the other assumptions of Theorem 4.8.

Lemma 4.10. *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$, and let $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ be non-empty. Assume $p \in \mathcal{P}^+(S)$ and $\tilde{p} \in \mathcal{P}^+(\tilde{S} \setminus \{0\})$. Define*

$$T := \{(y, z) \in \mathbb{R}^{2n}_+ : h_1(y - z) \geq 0, \dots, h_m(y - z) \geq 0\} = \{(y, z) \in \mathbb{R}^{2n}_+ : y - z \in S\},$$

then T is non empty, $p(y - z) \in \mathcal{P}^+(T)$ and $\tilde{p}(y - z) \in \mathcal{P}^+(\tilde{T} \setminus \{0\})$.

Proof. The statement follows since $x \in S$ implies $(\max\{0, x\}, -\min\{0, x\}) \in T$, and $\tilde{T} = \{(y, z) \in \mathbb{R}^{2n}_+ : y - z \in \tilde{S}\}$. □

Theorem 4.11. *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$, and let $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ be non-empty. Denote $d_{\max} = \max\{\deg h_1, \dots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. Assume that $p \in \mathcal{P}^+_{2d_{\max}}(S)$ and $\tilde{p} \in \mathcal{P}^+_{2d_{\max}}(\tilde{S} \setminus \{0\})$. Then there is $F \in \mathcal{P}_{2d_{\max}}(\mathbb{R}^{2n+m}_+)$ such that*

$$(1 + e^\intercal y + e^\intercal z)^{2d_{\max} - \deg p} p(y - z) = F(y, z, h_1(y - z), \dots, h_m(y - z)).$$

Proof. Define $T := \{(y, z) \in \mathbb{R}^{2n}_+ : h_1(y - z) \geq 0, \dots, h_m(y - z) \geq 0\} = \{(y, z) \in \mathbb{R}^{2n}_+ : y - z \in S\}$. By Lemma 4.10, the conditions of Theorem 4.8 are satisfied for the polynomial $p(y-z) \in \mathbb{R}[y,z]$ and the set $T \subseteq \mathbb{R}^{2n}$. Thus the result follows after applying Theorem 4.8 to $p(y-z) \in \mathbb{R}[y,z]$ and $T \subseteq \mathbb{R}^{2n}_+$. □

Doubling the number of variables to use results valid for set contained in the non-negative orthant is a standard technique in polynomial optimization. Among others, there is a recent work by Ahmadi and Hall [2] which exploits this technique. Ahmadi and Hall [2] construct new certificates of non-negativity using certificates of global positivity. Our results are connected to the results in [2] in the sense that we also construct new certificates based on the existing ones (the certificates of *copositivity*). Moreover, we use similar tools, such as doubling the number of variables or Pólya's theorem, in our proofs. However, in contrast to [2], our copositive certificates have a known pre-multiplier in front of p in all our Theorems 4.8, 4.9, and 4.11. Moreover, the certificates we propose exist for generic basic semialgebraic sets while the certificates from [2] are proven to exist on compact sets. It would be interesting to see whether the latter certificates exist on more general sets too.

For compact semialgebraic sets $S \subseteq \mathbb{R}^n$ that do not belong to the non-negative orthant, doubling the number of the variables is not needed since we can translate the set to the non-negative orthant. Similar to Lemma 4.10, the conditions of Theorem 4.8 will be maintained after applying the translation. We use this fact in Section 4.5 to reformulate PO problems over compact sets (see the proof of Corollary 4.36).

4.3.2 Genericity of strong positivity

We say that a property holds generically on a given set if it holds on a dense subset of this set. In this section we show that the strong positivity condition holds generically. First we introduce some additional definitions. We define the homogenization of $p \in \mathbb{R}[x]$ [see, e.g., 190] as the polynomial

$$p^h(x_0, x) = p\left(\tfrac{x_1}{x_0}, \ldots, \tfrac{x_n}{x_0}\right) x_0^{\deg(p)}. \tag{4.5}$$

Notice that by construction,

$$p(x) = p^h(1, x) \text{ and } \tilde{p}(x) = p^h(0, x). \tag{4.6}$$

Definition 4.12. *Let $h_1, \ldots, h_m \in \mathbb{R}_d[x]$ and let $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}$. We denote by S^h the following set*

$$S^h = \left\{(x_0, x) \in \mathbb{R}^{n+1} : x_0 \geq 0,\ h_1^h(x_0, x) \geq 0, \ldots, h_m^h(x_0, x) \geq 0\right\}. \tag{4.7}$$

Definition 4.13 (Guo et al. [78, 79] and Nie [159])**.** *The semialgebraic set S is called* closed at infinity *if*

$$\mathrm{cl}(\mathrm{cone}(\{1\} \times S)) = S^h \tag{4.8}$$

Closedness at infinity is one of the sufficient conditions for hierarchies of relaxations to PO problems proposed in [78, 79, 159] to converge to the optimal value [see, e.g., 159, Thm 2.5, condition (d)]. In [78, 159], this condition is shown to hold generically.

Proposition 4.14 (Nie [159, Sec. 3], Guo et al. [78, Sec. 2.2])**.** *Generically, the defining polynomials of a basic semialgebraic set S are such that S is closed at infinity. That is, generically* $\operatorname{cl}(\operatorname{cone}(\{1\} \times S)) = S^h$.

To connect closedness at infinity with strong positivity, we introduce the *horizon cone* of $S \subseteq \mathbb{R}^n$ [191],

$$S^\infty := \{x : (0, x) \in \operatorname{cl}(\operatorname{cone}(\{1\} \times S))\}.$$

The notation stems from the work by Peña et al. [175] who use an alternative definition of S^∞ to obtain completely positive reformulations of equality constrained PO problems. The following properties of the horizon cone are important throughout the chapter.

Proposition 4.15 ([see, 175, 192])**.** *For any $S, V \subseteq \mathbb{R}^n$ the following holds.*

(a) If $S \neq \emptyset$ is bounded, then $S^\infty = \{0\}$.

(b) $(S \cup V)^\infty = S^\infty \cup V^\infty$.

(c) $(S \cap V)^\infty \subseteq S^\infty \cap V^\infty$, but the reverse inclusion does not necessarily hold.

(d) Let $A \in R^{n\times m}$, $b \in \mathbb{R}^m$ and $S = \{x \in \mathbb{R}^n : Ax \geq b\}$. Then $S^\infty = \{x \in \mathbb{R}^n : Ax \geq 0\}$.

Lemma 4.16. *Let $p \in \mathbb{R}[x]$ and $S \subseteq \mathbb{R}^n$. Then*

(a) If p is bounded on S from below, then $\tilde{p} \in \mathcal{P}_{\deg p}(S^\infty)$.

(b) If S is a basic semialgebraic set, then $S^\infty \subseteq \tilde{S}$.

Proof. Statement (b) follows from (a). Now, we prove (a). Let p be bounded on S from below with the lower bound λ_{lb}. Then $p(x) - \lambda_{lb} \in \mathcal{P}_{\deg p}(S)$. Let $y \in S^\infty$. Then there exists $\lambda^k > 0$ and $x^k \in S$ for $k = 1, \dots$ such that $\lim_{k\to\infty} \lambda^k = 0$ and $\lim_{k\to\infty} \lambda^k x^k = y$. For $0 \leq \ell < \deg p$, let $f^\ell(x)$ be the homogeneous component of $p(x) - \lambda_{lb}$ of degree ℓ. We have that

$$\begin{aligned}\lim_{k\to\infty} (\lambda^k)^{\deg p}(p(x^k) - \lambda_{lb}) &= \lim_{k\to\infty} (\lambda^k)^{\deg p} \left(\tilde{p}(x^k) + \sum_{\ell<\deg p} f^\ell(x^k)\right) \\ &= \lim_{k\to\infty} \left(\tilde{p}(\lambda^k x^k) + \sum_{\ell<\deg p} (\lambda^k)^{\deg p-\ell} f^\ell(\lambda^k x^k)\right) = \tilde{p}(y).\end{aligned}$$

As $(\lambda^k)^d(\lambda^k)^d(p(x^k) - \lambda_{lb}) \geq 0$ for all k we obtain $\tilde{p}(y) \geq 0$. □

Using the horizon cone and Lemma 4.16, we can characterize $\operatorname{int} P_d(S)$ for unbounded S.

Proposition 4.17. *Let $d > 0$, and let $S \subseteq \mathbb{R}^n$ be unbounded. Then* $\operatorname{int} P_d(S) = \left\{p \in \mathbb{R}_{=d}[x] : p \in \mathcal{P}_d^+(S),\ \tilde{p} \in \mathcal{P}_d^+(S^\infty \setminus \{0\})\right\}$

Proof. Let $p \in \operatorname{int} P_d(S)$, then it follows that $p \in \mathbb{R}_{=d}[x]$ and $p \in \mathcal{P}_d^+(S)$. To show that $\tilde{p} \in \mathcal{P}_d^+(S^\infty \setminus \{0\})$, let $y \in S^\infty$, $y \neq 0$. Without loss of generality, $y_1 > 0$. Then, for some $\varepsilon > 0$ the polynomial $q(x) = p(x) - \varepsilon x_1^d \in P_d(S)$. From Lemma 4.16a, $\tilde{q} \in \mathcal{P}_d(S^\infty)$, therefore $\tilde{p}(y) \geq \varepsilon y_1^d > 0$. Thus, $\operatorname{int} P_d(S) \subseteq \{p \in \mathbb{R}_{=d}[x] : p \in \mathcal{P}_d^+(S),\ \tilde{p} \in \mathcal{P}_d^+(S^\infty \setminus \{0\})\}$. To show that $\operatorname{int} P_d(S) \supseteq \{p \in \mathbb{R}_{=d}[x] : p \in \mathcal{P}_d^+(S), \tilde{p} \in \mathcal{P}_d^+(S^\infty \setminus \{0\})\}$, let $p \in \mathbb{R}_{=d}[x]$ such that $p \in \mathcal{P}_d^+(S)$ and $\tilde{p} \in \mathcal{P}_d^+(S^\infty \setminus \{0\})$. For the sake of contradiction, assume $p \notin \operatorname{int} P_d(S)$. Then there exists $q \in \mathbb{R}_d[x]$ such that for $k = 1, 2 \ldots$ there exists $x^k \in S$ such that

$$p(x^k) - \tfrac{1}{k} q(x^k) < 0$$

The sequence x^k, $k = 1, \ldots$ is unbounded. Define $\lambda^k := \frac{1}{\|x^k\|_2}$, $k = 1, \ldots$ so that $\lim_{k\to\infty} \lambda^k = 0$. The sequence $\lambda^k x^k$, $k = 1, \ldots$ is bounded and thus has a convergent subsequence with a limit $y \in S' := \{y \in S^\infty : \|y\| = 1\}$. We have then, for all ε,

$$0 > \lim_{k\to\infty} (\lambda^k)^d (p(x^k) - \varepsilon q(x^k)) = \begin{cases} \tilde{p}(y), & \text{if } \deg q < d \\ \tilde{p}(y) - \varepsilon \tilde{q}(y), & \text{if } \deg q = d. \end{cases}$$

But $\tilde{p} \in \mathcal{P}_d^+(S')$ and S' is compact. Thus for some $\varepsilon > 0$ small enough we obtain a contradiction. □

Lemma 4.16 b and Proposition 4.17 together imply that every polynomial of degree d that is strongly positive on S is in $\operatorname{int} \mathcal{P}_d(S)$ (as stated in Proposition 4.6).

Corollary 4.18. *Let $d > 0$. Generically, the defining polynomials of a basic semi-algebraic set S and a polynomial $p \in \mathcal{P}_d(S)$ are such that p is strongly positive on S.*

Proof. From Proposition 4.14, generically $\operatorname{cl}(\operatorname{cone}(\{1\} \times S)) = S^h$. Hence $\tilde{S} = \{x : (0, x) \in S^h\} = S^\infty$. If S is compact, then $\tilde{S} = \{0\}$. Otherwise, using Proposition 4.17,

$$\operatorname{int} P_d(S) = \left\{p \in \mathbb{R}_{=d}[x] : p \in \mathcal{P}_d^+(S),\ \tilde{p} \in \mathcal{P}_d^+(\tilde{S} \setminus \{0\})\right\}.$$

Since $p \in \mathcal{P}_d(S)$, generically $p \in \operatorname{int} P_d(S)$, and thus p is strongly positive. □

Since we are especially interested in copositive polynomials, we next look at the interior of the cone of copositive polynomials of degree at most d. Proposition 4.17 implies the following characterizations of the interior of $\mathbb{R}^n_+$.

Corollary 4.19. *For any $p \in \mathbb{R}_d[x]$ the following statements are equivalent*

(a) $p \in \operatorname{int} \mathcal{P}_d(\mathbb{R}^n_+)$.

(b) $\deg p = d$, $p \in \mathcal{P}^+_d(\mathbb{R}^n_+)$ *and* $\tilde{p} \in \mathcal{P}^+_d(\mathbb{R}^n_+ \setminus \{0\})$.

(c) $\deg p = d$ *and* $p^h \in \mathcal{P}^+_d(\mathbb{R}^{n+1}_+ \setminus \{0\})$.

Proof. Statement (b) follows from Proposition 4.17. Statement (c) follows from statement (b) and (4.6). □

4.3.3 Examples

By Proposition 4.17, if S is closed at infinity, then generically a non-negative polynomial on S is strongly positive on S. Next, we present some examples of sets that are closed or not closed at infinity and show several sufficient conditions for closedness at infinty. One could expect that this condition is always satisfied for compact sets or for sets generated by one constraint. However, Example 4.20 shows that both statements are false.

Example 4.20 (Violation of closedness at infinity for a compact set generated by one constraint)**.** Let $h(x_1, x_2) = -x_1^4 - x_2^2 + 1$. And let $S = \{(x_1, x_2) \in \mathbb{R}^2 : h(x_1, x_2) \geq 0\}$. Notice that $(0, 0) \in S$, so S is non-empty. Also, $S \subset [-1, 1]^2$ and thus compact since it is a bounded basic closed semialgebraic set. We have $(0, 0, 1) \in S^h$, but we claim that $(x_0, x_1, x_2) \in \operatorname{cl}(\operatorname{cone}(\{1\} \times S))$ and $x_0 = 0$ implies $x_2 = 0$. This is because for any $(x_0, x_1, x_2) \in \operatorname{cone}(\{1\} \times S)$ we have $x_1^4 + x_2^2 x_0^2 \leq x_0^4$ which implies $|x_2| \leq x_0$.

Example 4.21 (Violation of closedness at infinity for an unbounded set)**.** Let $h_1(x) = x_1, h_2(x) = x_2, h_3(x) = (x_1 x_2 + 1)(x_1 - x_2)$ and let

$$S = \left\{ x \in \mathbb{R}^2 : h_1(x) \geq 0, h_2(x) \geq 0, h_3(x) \geq 0 \right\}.$$

For any $t \geq 0$ we have that $(0, 0, t) \in S^h$. On the other hand, $x = (x_0, x_1, x_2) \in \operatorname{cone}(\{1\} \times S) \setminus \{0\}$ we have $x_1 \geq 0$, $x_2 \geq 0$ and $(x_1 x_2 + x_0^2)(x_1 - x_2) \geq 0$, that is $x_1 \leq x_2$. Thus $(0, 0, t) \notin \operatorname{cl} \operatorname{cone}(\{1\} \times S)$ for $t > 0$.

Now, we turn our attention to sufficient conditions for closedness at infinity. As this property holds generically (see Proposition 4.14), it is not a surprise that there are several families of semialgebraic sets for which closedness at infinity is straightforward to verify.

Proposition 4.22. *Let* $h_1, \ldots, h_m \in \mathbb{R}[x]$ *and* $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}$. *If any of the following conditions holds, then* $\operatorname{cl}(\operatorname{cone}(\{1\} \times S)) = S^h$.

(a) $h_m(x) = N - \|x\|^2$ *for some* $N > 0$.

(b) $h_1, \dots, h_m$ *are homogeneous.*

(c) $h_i(x) = q_1^i(x) \cdots q_{k_i}^i(x)$ *for some* $k_i > 0$ *and* $q_1^i, \dots, q_{k_i}^i \in \mathbb{R}_1[x]$. *Notice that in this case* S *is a union of polyhedra.*

(d) $n \geq 2$ *and* $S = \{x \in \mathbb{R}^n : \left(x_n - \sum_{i=1}^{n-1} x_i^2 - b\right) q(x) \geq 0,\ x_n \geq 0\}$, *where* $b \in \mathbb{R}$ *and* $q \in \mathbb{R}[x]$ *is such that* $\tilde{q} \in \mathcal{P}^+(\mathbb{R}^n \setminus \{0\})$.

Proof. By construction, $\tilde{S} = \{x : (0, x) \in S^h\}$. Hence $\tilde{S} = S^\infty$ if and only if $\mathrm{cl}(\mathrm{cone}(\{1\} \times S)) = S^h$. Moreover, from Lemma 4.16 (b) it follows that $\tilde{S} \supseteq S^\infty$. Therefore throughout this proof we only show that $\tilde{S} \subseteq S^\infty$.

Proof of statement (a). Without loss of generality, let $h_1(x) = N - \|x\|^2$. Then $\tilde{h}_1(x) = -\|x\|^2$, therefore $\tilde{S} = \{0\} = S^\infty$. The latter equality follows from Proposition 4.15 (a).

Proof of statement (b). Let $x \in \tilde{S}$. Then for every every $j \in [m]$ and $k > 0$, $h_j(kx) = \tilde{h}_j(kx) = k^{\deg h_j} \tilde{h}_j(x) \geq 0$. Since $x = \lim_{k \to \infty} \frac{kx}{k}$, we see that $x \in S^\infty$.

Proof of statement (c). The statement follows from Proposition 4.15 (b) and (d). If h_1 is a product of finitely many polynomials of degree one, then S can be written as a union of sets generated by linear equalities and inequalities. For example, if $h_1(x) = x_1(x_2 + 1)$, then

$$S = \mathbb{R}^n \cap \{\{x_1 \geq 0, x_2 + 1 \geq 0\} \cup \{-x_1 \geq 0, -x_2 - 1 \geq 0\} \cup \{x_1 = 0\} \cup \{x_2 + 1 = 0\}\}.$$

Proof of statement (d). On the one hand,

$$\tilde{S} = \left\{x \in \mathbb{R}^n_+ : \left(-\sum_{i=1}^{n-1} x_i^2\right)\tilde{q} \geq 0\right\} = \{(0, \dots, 0, x_n)^\intercal : x_n \geq 0\}.$$

On the other hand, for any $x_n \geq 0$ and $k > 0$, let $x^k = \left(\frac{k}{\sqrt{n-1}}, \dots, \frac{k}{\sqrt{n-1}}, k^2 + b\right)$, and $\lambda_k = \frac{x_n}{k^2}$. Then $x^k \in S$ for all k, $\lambda_k \downarrow 0$ and $\lim_{k \to \infty} \lambda_k x^k = (0, \dots, 0, x_n)^\intercal$. Therefore

$$S^\infty \supseteq \{(0, \dots, 0, x_n)^\intercal : x_n \geq 0\} = \tilde{S}.$$

□

An important question in algebraic geometry and in optimization is when the non-negativity of a polynomial on a set can or cannot be certified using the *quadratic module* [see 19]. Putinar [183] answers this question affirmatively when the quadratic module is Archimedean. Putinar's Positivstellensatz [183] underlines LMI approximations of PO problems with compact or "compactifiable" feasible sets S [see, e.g., 95, 114, 115, 118] since one could add the norm-constraint $N - \|x\|^2 \geq 0$ to the description of such S. In our next example, we show that copositive certificates of non-negativity could exist on the sets where the certificates based on the quadratic module do not exist.

We say that polynomials $\{h_1(x), \dots, h_m(x)\}$ satisfy the *strong moment property* (SMP) if $\mathrm{cl}(\mathcal{QM}(h_1(x), \dots, h_m(x))) = \mathcal{P}(S)$, where $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$, (see Proposition 3.4.1. in [139]). SMP implies that $p \in \operatorname{int} \mathcal{P}(S)$ can be written in the form

$$p(x) = \sigma_0(x) + \sum_{j=1}^{m} \sigma_j(x) h_j(x), \text{ where } \sigma_j \text{ is SOS for all } j = 0, \dots, m. \tag{4.9}$$

Definition 4.23 (Tentacles)**.** *Given a compact set $K \subseteq \mathbb{R}^n$ with nonempty interior, a* tentacle *of K in direction z is the set*

$$T_{K,z} := \{(\lambda^{z_1} x_1, \dots, \lambda^{z_n} x_n) : \lambda \geq 1, x = (x_1, \dots, x_n) \in K\}.$$

Netzer [157] shows that if S contains tentacles of a certain type, this set does not satisfy the SMP.

Example 4.24 (A set that violates the SMP but is closed at infinity)**.** Let $n \geq 2$ and consider the set

$$S = \left\{ x \in \mathbb{R}^n : \left(x_n - \sum_{i=1}^{n-1} x_i^2 \right) \left(2 \sum_{i=1}^{n-1} x_i^2 - x_n \right) \geq 0,\ x_n \geq 0 \right\}.$$

Figure 4.1 shows this set for $n = 2$.

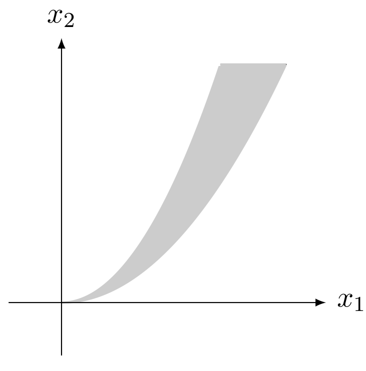


Figure 4.1 – Illustration of the set $S = \{x \in \mathbb{R}^2 : (x_2 - x_1^2)(2x_1^2 - x_2) \geq 0,\ x_2 \geq 0\}$ (in gray).

This example is similar to one presented in [157, Section 6]. Let $z = (1, \dots, 1, 2)^\intercal$ and

$$K = \left\{ x \in \mathbb{R}^n : |x_n - n + \tfrac{1}{2}| \leq \tfrac{1}{10n} \text{ and } |x_i - 1| \leq \tfrac{1}{10n}\, i = 1, \dots, n{-}1 \right\}.$$

We claim that the tentacle $T_{K,z} \subseteq S$. From [157, Thm. 5.4] we obtain that $\{(x_n - \sum_{i=1}^{n-1} x_i^2)(2\sum_{i=1}^{n-1} x_i^2 - x_n)\}$ does not satisfy the SMP. Thus, for some $d > 0$ there is $p \in \operatorname{int} \mathcal{P}_d(S)$ for which no certificate of non-negativity of the form (4.9) exists. On the other hand, from Proposition 4.22d the closedness at infinity condition holds. Hence Theorem 4.11 implies that for all $p \in \operatorname{int} P_d(S)$ a copositive certificate of non-negativity exists.

To prove the claim, notice that if $x \in S$, then for every $\lambda > 0$, $(\lambda x_1, \ldots, \lambda x_{n-1}, \lambda^2 x_n) \in S$. Thus, it is enough to show that $K \subset S$. Since $n \geq 2$, for $x \in K$ we have

$$x_n - \sum_{i=1}^{n-1} x_i^2 \geq n - \tfrac{1}{2} - \tfrac{1}{10n} - (n-1)\left(1 + \tfrac{1}{10n}\right)^2 = \tfrac{3}{10} + \tfrac{1+9n}{100n^2} > 0,$$

$$2\sum_{i=1}^{n-1} x_i^2 - x_n \geq 2(n-1)\left(1 - \tfrac{1}{10n}\right)^2 - n + \tfrac{1}{2} - \tfrac{1}{10n} = n - \tfrac{19}{10} + \tfrac{16n-1}{50n^2} > 0.$$

4.4 LP-based and sparse certificates of non-negativity on compact sets

Using Theorems 4.8, 4.9 and 4.11, one can construct from any certificate of copositivity a corresponding certificate of non-negativity for any given semialgebraic set S and any strongly positive polynomial on S. In this section we use two certificates of copositivity to illustrate this approach and obtain new certificates of non-negativity.

Our first example (see Corollary 4.26) is based on the celebrated Pólya's certificate of copositivity. Applying this certificate in optimization leads to LP approximations of PO problems. More importantly, the certificate can be strengthened so that instead of non-negative constant polynomials (resulting in LP approximations) one can use any set of non-negative polynomials with non-zero constant terms, such as SOS, scaled diagonally dominant SOS (SDSOS) [3], hyperbolic polynomials [197], non-negative circuit polynomials [91], etc. (see Corollary 4.28). As a result, one obtains convergent LMI hierarchies of approximations to PO problems which could provide stronger bounds than the mentioned LP hierarchies.

Our second example is a sparse certificate of non-negativity for generic semialgebraic sets. More precisely, we present an SOS-based certificate where all but two SOS polynomials are univariate which results in the use of lower dimensional SDP constraints in LMI approximations of PO problems (see Corollary 4.31). To obtain this result, we propose a new sparse certificate of copositivity in Theorem 4.30.

In both examples we use certificates of copositivity which are guaranteed to exists only for polynomials in $\operatorname{int} \mathcal{P}_d(\mathbb{R}^n_+)$. For this reason we limit ourselves to compact sets, in order to take advantage of Theorem 4.9 which guarantees the existence of copositive certificate lying in the interior of the cone of copositive polynomials. Also, for ease of presentation we work with $S \subset \mathbb{R}^n_+$ since any compact set can be translated to the non-negative orthant. Corollary 4.36 in Section 4.5 shows how to implement the results in this section for general compact sets.

The certificates we obtain in this section use rational polynomial expressions to certify the non-negativity of a polynomial on a set $S = \{x \in \mathbb{R}^n_+ : h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}$. That is, the certificates are (basically) of the form $G(h_1(x), \ldots, h_m(x))p(x) = F(h_1(x), \ldots, h_m(x))$, where F, G are copositive polynomials. The existence of such rational certificates is in general guaranteed by the Krivine-Stengle Positivstellensatz [107, 213]. However, the problem of finding such certificates is not tractable in general because the denominator G is unknown [see, e.g., 95, for more details]. The rational certificates of non-negativity introduced in this section have fixed denominators. Hence, these certificates provide efficiently solvable lower bound approximations to PO problems. We present examples of such approximations in Section 4.5.

4.4.1 LP certificates

Our first illustration of constructing new certificates of non-negativity using certificates of copositivity is based on Pólya's certificate of copositivity [see, e.g., 83] (see Theorem 1.8).

Corollary 4.25. *Let $d > 0$ and $F \in \mathbb{R}_d[x]$ be such that $F \in \operatorname{int} \mathcal{P}_d(\mathbb{R}^n_+)$. Then for some $r > 0$ all the coefficients of $(1 + e^\intercal x)^r F(x)$ are non-negative.*

Proof. The result follows from Corollary 4.19 c by applying Theorem 1.8 to F^h. □

Combining Theorem 4.9 and Corollary 4.25, we obtain the new certificate of non-negativity stated in Corollary 4.26 below.

As before, for $n > 0$ and $d \geq 0$ we define $\mathbb{N}^n_d = \{\alpha \in \mathbb{N}^n : e^\intercal \alpha \leq d\}$. Given $h_1, \ldots, h_m \in \mathbb{R}[x]$ and $\alpha \in \mathbb{N}^m_d$, we use $h^\alpha := \prod_{j=1}^m h_j^{\alpha_j}$. In particular $x^\alpha = \prod_{j=1}^m x_j^{\alpha_j}$. Also, we use the notation h to arrange the polynomials $h_1, \ldots, h_m$ in an array; that is, $h := [h_1, \ldots, h_m]^\intercal$.

Corollary 4.26. *Let $p, h_1, \ldots, h_m \in \mathbb{R}[x]$ and $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}$ be non-empty. Also, let $M > 0$ be such that $S \subseteq \{x \in \mathbb{R}^n_+ : e^\intercal x \leq M\}$, and let any $a \in \mathbb{R}^n_+$ and $b \in \mathbb{R}^m_+$ be given. Denote $d_{max} = \max\{\deg h_1, \ldots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. If $p \in \mathcal{P}^+_{2d_{\max}}(S)$ then there exists $r \geq 0$ and $c_{\alpha,\beta,\eta} \geq 0$ for $(\alpha, \beta, \gamma) \in \mathbb{N}^{n+m+1}_{d_{\max}(2d_{\max}+r)}$ such that*

$$(1 + a^\intercal x + b^\intercal h(x))^r \, p(x) = \sum_{(\alpha,\beta,\gamma) \in \mathbb{N}^{n+m+1}_{d_{\max}(2d_{\max}+r)}} c_{\alpha,\beta,\gamma} x^\alpha h(x)^\beta (M - e^\intercal x)^\gamma. \quad (4.10)$$

Proof. Let $d_j = \deg h_j$. Denote $g_j(x) = (1 + M)^{d_j} \|h_j\| - h_j(x)$. By Theorem 4.9, there is $F \in \operatorname{int} \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m+1}_+)$ such that

$$p(x) = F(x, h(x), M - e^\intercal x + e^\intercal g(x)). \tag{4.11}$$

Let $\hat{M} = 1 + M + \sum_{j=1}^m (1+M)^{d_j} \|h_j\|$. $\hat{a} = \hat{M}a + e$, $\hat{b}_j = \hat{M}b + e$. By construction $\hat{a} > 0$, $\hat{b} > 0$. Denote $\hat{x}_i = x_i \hat{a}_i$ for $i \in \{1, \ldots, n\}$ and $\hat{u}_j = u_j \hat{b}_j$ for $j \in \{1, \ldots, m\}$. Then

$$F(\hat{x}, \hat{u}, v) \in \operatorname{int} \mathcal{P}_{2d_{\max}}(\mathbb{R}_+^{n+m+1}).$$

Applying Corollary 4.25, we obtain that there is $r \geq 0$ and $k_{\alpha,\beta,\gamma} \geq 0$ for $(\alpha, \beta, \gamma) \in \mathbb{N}_{2d_{\max}+r}^{n+m+1}$ such that

$$(1 + e^\intercal \hat{x} + e^\intercal \hat{u} + v)^r F(\hat{x}, \hat{u}, v) = \sum_{(\alpha,\beta,\gamma) \in \mathbb{N}_{2d_{\max}+r}^{n+m+1}} k_{\alpha,\beta,\gamma} \hat{x}^\alpha \hat{u}^\beta v^\gamma.$$

Thus,

$$(1 + \hat{a}^\intercal x + \hat{b}^\intercal u + v)^r F(x, u, v) = \sum_{(\alpha,\beta,\gamma) \in \mathbb{N}_{2d_{\max}+r}^{n+m+1}} \hat{k}_{\alpha,\beta,\gamma} x^\alpha u^\beta v^\gamma,$$

where $\hat{k}_{\alpha,\beta,\gamma} \geq 0$ for all $(\alpha, \beta, \gamma) \in \mathbb{N}_{2d_{\max}+r}^{n+m+1}$. Using (4.11) we obtain that

$$\begin{aligned} &\Big(1 + \hat{a}^\intercal x + \hat{b}^\intercal h(x) + M - e^\intercal x + e^\intercal g(x)\Big)^r p(x) = \\ &\qquad \sum_{(\alpha,\beta,\gamma) \in \mathbb{N}_{2d_{\max}+r}^{n+m+1}} \hat{k}_{\alpha,\beta,\gamma} x^\alpha h(x)^\beta \Big(M - e^\intercal x + e^\intercal g(x)\Big)^\gamma. \end{aligned} \tag{4.12}$$

To finish the proof, notice that

$$\begin{aligned} 1 + \hat{a}^\intercal x + \hat{b}^\intercal h(x) + M - e^\intercal x + e^\intercal g(x) =& 1 + \Big(\hat{M}a + e\Big)^\intercal x + \sum_{j=1}^m (\hat{M}b_j + 1) h_j(x) \\ &+ \sum_{j=1}^m \Big((1+M)^{d_j} \|h_j\| - h_j(x)\Big) + M - e^\intercal x \\ =& \hat{M}(1 + a^\intercal x + b^\intercal h(x)), \end{aligned}$$

which, up to a positive constant multiplier, is equivalent to the left-hand side factor of $p(x)$ in (4.10). Also, for each $j = 1, \ldots, m$ we have

$$\begin{aligned} g_j(x) &= (1+M)^{d_j} \|h_j\| - h_j(x) \\ &= \|h_j\|((1+M)^{d_j} - (1 + e^\intercal x)^{d_j}) + \sum_{\alpha \in \mathbb{N}_{d_j}^n} C_{d_j,\alpha} (\|h_j\| - (h_j)_\alpha) x^\alpha \\ &= \|h_j\|(M - e^\intercal x) \sum_{k=0}^{d_j - 1} (M+1)^{d_j - k - 1} (1 + e^\intercal x)^k + \sum_{\alpha \in \mathbb{N}_{d_j}^n} C_{\deg h_j,\alpha} (\|h_j\| - (h_j)_\alpha) x^\alpha. \end{aligned}$$

After replacing the expression for $g_j(x), j = 1, \ldots, m$ above into $e^\intercal g(x)$ in the right hand side of (4.12), the right-hand side of (4.12) is equivalent to the right-hand side of (4.10). □

Remark 4.27. *The choice of the vectors a, b in Corollary 4.26 is free, and we can set $a = 0,\ b = 0$ to eliminate the pre-multiplier in front of $p(x)$ in* (4.10)*.*

For every $r \in \mathbb{N}$, the only unknowns in the certificate from Corollary 4.26 are the non-negative constants $c_{\alpha,\beta,\gamma} \geq 0$ for all $(\alpha, \beta, \gamma) \in \mathbb{N}^{n+m+1}_{d_{\max}(2d_{\max}+r)}$. The representation (4.10) is linear in these constants. As we show in Section 4.5, we can use the hierarchy (4.10) for every $r \in N$, to obtain LP lower bounds for PO problems over compact semialgebraic sets. Setting $a = e,\ b = 0$ in (4.10) results in the certificate from Theorem 1.14 by Dickinson and Povh [48].

Recently, it has been a topic of great interest to replace SOS-based certificates by certificates based on other types of non-negative polynomials. The idea is to provide alternative certificates that can lead to LMI relaxation bounds that are computationally cheaper to compute, but still provide quality bounds for the PO problem. This is typically done by replacing the full dimensional SOS polynomials on non-negative certificates based on Putinar's Positivstellensatz by SDSOS [3] (which result in second-order cone programming (SOCP) relaxations), hyperbolic polynomials [197] (which results hyperbolic programming relaxations) and non-negative circuit polynomials [59, 219] (which result in geometric programming relaxations). Since these alternative sets of polynomials are not necessarily supersets of SOS polynomials, the resulting LMI hierarchies of bounds on PO problems can require additional assumptions to converge [see, for instance, 59], are not proven to converge [e.g., 3] or require the use of additional methods to ensure convergence [see e.g., 2]. In contrast, all earlier mentioned subsets of non-negative polynomials can be used to strengthen the LP-based certificates from Corollary 4.26 to obtain convergent LMI hierarchies of bounds on PO problems with compact feasible sets.

Corollary 4.28. *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$ and $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ be non-empty. Also, let M be such that $S \subseteq \{x \in \mathbb{R}^n_+ : e^\intercal x \leq M\}$, and let any $a \in \mathbb{R}^n_+$ and $b \in \mathbb{R}^m_+$ be given. Denote $d_{\max} = \max\{\deg h_1, \dots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. Let $d \geq 0$ and let $\mathcal{K} \subset \mathcal{P}_d(\mathbb{R}^n)$ be such that $\mathbb{R}_+ \subseteq \mathcal{K}$. If $p \in \mathcal{P}^+_{2d_{\max}}(S)$ then there exists $r \geq 0$ and $c_{\alpha,\beta,\gamma} \in \mathcal{K}$ for $(\alpha, \beta, \gamma) \in \mathbb{N}^{n+m+1}_{d_{\max}(2d_{\max}+r)}$ such that*

$$\left(1 + a^\intercal x + b^\intercal h(x)\right)^r p(x) = \sum_{(\alpha,\beta,\gamma)\in\mathbb{N}^{n+m+1}_{d_{\max}(2d_{\max}+r)}} c_{\alpha,\beta,\gamma}(x) x^\alpha h(x)^\beta (M - e^\intercal x)^\gamma. \quad (4.13)$$

The computational benefits of the certificates arising from Corollary 4.28 have been explored by Kuang et al. [108], who consider setting K to be: $\mathbb{R}_+$, as well as the cone of quadratic diagonally dominant SOS (DSOS), SDSOS, and SOS polynomials. The authors conclude that PO-hierarchies based on certificate (4.13) can be more computationally efficient compared to the broadly used Lasserre's hierarchies [114, 121], as well as the SDSOS hierarchies in [3].

A final noteworthy characteristic of the LP certificates proposed in this section is that the unknown r could be small, even without considering the additional strengthenings, beyond the use of non-negative constant polynomials, in Corollary 4.28. That is, the polynomials involved in the certificate (4.10) can have low degrees. This implies that the LP one has to solve to find a certificate is not too large, as illustrated by Example 4.29. This situation is in contrast to the existing limited research on LP certificates of non-negativity of polynomials [see, e.g., 35, 103, 119, 195, for noteworthy examples].

Example 4.29 (Low degree convergence)**.** We show that the polyhedral hierarchy (4.10) could convergence for small r by considering an instance of the *Celis-Dennis-Tapia* (CDT) problem [see 31]. This classical problem is concerned with the non-negativity of a quadratic polynomial on the intersection of two ellipses. Recent advances on this problem have been made thanks to the use of polynomial optimization techniques [see 24]. Specifically, for $n \geq 3$ consider the polynomial $q \in \mathbb{R}[x]$:

$$q(x) := -2x_1 + 8x_1 \sum_{i=1}^{n} x_i.$$

Note that q is not a copositive polynomial; that is, $q \notin \mathcal{P}(\mathbb{R}^n_+)$. In particular,

$$q(x_1, 0, \ldots, 0) < 0 \text{ for } 0 < x_1 < 1/4.$$

However, we can use Corollary 4.26 to certify that $q \in \mathcal{P}(\mathcal{B}_e \cap \mathcal{B}_{e/2})$, where

$$\mathcal{B}_c = \left\{ x \in \mathbb{R}^n_+ : b_c(x) := 1 - \|x - c\|^2 \geq 0 \right\},$$

is the unitary ball centered at $c \in \mathbb{R}^n$. In particular notice that

$$\begin{aligned}\left(1 + e^\intercal x + b_e(x) + b_{e/2}(x)\right) q(x) &= 8x_1(e^\intercal x)\left(b_e(x) + b_{e/2}(x)\right)\\ &\quad + x_1\left(\left(\tfrac{5}{2}n - 6\right) + 8(e^\intercal x)^2 + 4\sum_{i=2}^{n} x_i^2\right),\end{aligned}$$

for $n \geq 3$. After expanding the right hand side, the expression above has the form (4.10) with $r = 1$. In particular, this certifies that q is non-negative on the $\mathcal{B}_e \cap \mathcal{B}_{e/2}$.

4.4.2 Sparse certificates

As another illustration of the power of our approach, we construct sparse SOS-certificates of non-negativity of polynomials over compact semialgebraic sets. For that purpose, we first construct sparse certificates of copositivity. Then, using Theorem 4.9, we translate the result to any compact semialgebraic set.

Theorem 4.30. *Let $F \in \operatorname{int} \mathcal{P}_d(\mathbb{R}^n_+)$. Then there exist $r \geq 0$, n-variate SOS polynomials σ_0 and σ_1, and bivariate homogeneous SOS polynomials $\hat{\sigma}_0, \dots, \hat{\sigma}_n$ such that*

$$(1+e^\intercal x)^r F(x) = \sigma_0(x)+\sigma_1(x) \sum_{0\leq i\leq j\leq n} x_i x_j+(1+e^\intercal x)\sum_{i=0}^{n} \hat{\sigma}_i\left(x_i, 1+e^\intercal x\right) x_i, \quad (4.14)$$

where $x_0 := 1$.

Combining Theorem 4.30 and Theorem 4.9 we obtain a sparse certificate of non-negativity on compact sets.

Corollary 4.31. *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$, and $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ be non-empty. Let $M > 0$ be such that $S \subseteq \{x \in \mathbb{R}^n_+ : e^\intercal x \leq M\}$. Denote $X = \Big(x_1, \dots, x_n, h_1(x), \dots, h_m(x), M - e^\intercal x + \sum_{j=1}^m \Big((1+M)^{d_j}\|h_j\| - h_j(x)\Big)\Big)$. If $p \in \mathcal{P}^+(S)$, then there exist n-variate SOS polynomials σ_0 and σ_1, and univariate SOS polynomials $\hat{\sigma}_1, \dots, \hat{\sigma}_{n+m+1}$ such that*

$$\begin{aligned} p(x) = \sigma_0(x) + \sigma_1(x)\left(\Big(1+M+\sum_{j=1}^m (1+M)^{d_j}\|h_j\|\Big)^2 - 1 - \sum_{i=1}^{n+m+1} X_i^2\right) \quad (4.15) \\ + \sum_{i=1}^{n+m+1} \hat{\sigma}_i\left(X_i\right) X_i. \end{aligned}$$

Proof. Denote $X_0 = 1$. From Theorem 4.30, with $d = \deg p$, and Theorem 4.9 we obtain that there are $r \geq 0$ and n-variate SOS polynomials σ_0 and σ_1, and homogeneous bivariate SOS polynomials $\hat{\sigma}_0, \dots, \hat{\sigma}_{n+m+1}$ such that

$$\begin{aligned} (1+e^\intercal X)^r p(x) = \sigma_0(x) + \sigma_1(x) \sum_{0\leq i\leq j\leq n+m+1} X_i X_j \\ + (1+e^\intercal X) \sum_{i=0}^{n+m+1} \hat{\sigma}_i\left(X_i, 1+e^\intercal X\right) X_i. \end{aligned}$$

Using $X_0 + e^\intercal X = 1 + e^\intercal X = 1 + M + \sum_{j=1}^m (1+M)^{d_j}\|h_j\|$ and

$$\sum_{0\leq i\leq j\leq n+m+1} 2X_i X_j = (X_0 + e^\intercal X)^2 - \sum_{i=0}^{n+m+1} X_i^2,$$

we obtain (4.15), up to a positive constant multiplier. □

The certificates constructed in Theorem 4.30 and Corollary 4.31 are sparse in the sense that the SOS polynomial multipliers $\hat{\sigma}_i$, $i = 1, \dots, n+m+1$ are all sparse. Indeed while σ_0 and σ_1 are full SOS, each $\hat{\sigma}_i$ is univariate. A univariate SOS of degree d can be represented using a $(d+1) \times (d+1)$ SDP matrix which is much smaller than the one needed to represent a multivariate SOS of the same degree.

The rest of this section shows the proof of Theorem 4.30.

Lemma 4.32. *Let $S \subset \mathbb{R}^n$ be non-empty and compact, and let $p \in \mathbb{R}[x]$. Then $p \in \mathcal{P}^+(S \cap \mathbb{R}^n_+)$ if and only if*

$$p(x) = q(x) + \sum_{i=1}^{n} x_i \sigma_i(x_i), \tag{4.16}$$

where $\sigma_1, ..., \sigma_n$ are univariate SOS polynomials and $q \in \mathcal{P}^+(S)$.

Proof. If $S \subset \mathbb{R}^n_+$, then the result is straightforward, thus further in the proof we assume that $S \not\subseteq \mathbb{R}^n_+$. Without loss of generality, there exists $k \leq n$ such that $\{x \in S, x_i < 0\} \neq \emptyset$ for all $i \in \{1, \dots, k\}$, and $\{x \in S, x_i < 0\} = \emptyset$ for all $i \in \{k+1, \dots, n\}$. Since $p \in \mathcal{P}^+(S \cap \mathbb{R}^n_+)$, there exists $\varepsilon > 0$ such that $x \in S$ and $x > -\varepsilon$ implies $p(x) \geq 0$. Also, let $M > 0$ be such that $x \in S$ implies $x < M$. Let $p^0_{\min} = \min\{p(x) : x \in S \cap \mathbb{R}^n_+\}$, and let $p^i_{\min} = \min\{p(x) : x \in S, x_i \leq -\varepsilon\}$ for $i \in \{1, \dots, k\}$. Consider the function $f_i(x) = a_i x_i e^{-b_i x_i}$ for some $a_i > 0$, $b_i > 0$. For any $x \in \mathbb{R}^n$ and $i \in \{1, \dots, n\}$, we have that $f_i(x)$ is positive for $x_i > 0$ and negative for $x_i < 0$. For $i \in \{1, \dots, k\}$, we can tailor a_i and b_i so that $\max\{f_i(x) : x \in S : x_i \leq -\varepsilon\} \leq -\varepsilon a_i < \frac{p^i_{\min}}{n}$ and $\max\{f_i(x) : x \in S \cap \mathbb{R}^n_+\} \leq a_i M e^{-b_i M} < \frac{p^0_{\min}}{n}$. For $i \in \{k+1, \dots, n\}$, we let $f_i(x) = 0$. Defining $f(x) = \sum_{i=1}^n f_i(x)$ we obtain $p(x) > f(x)$ for all $x \in S$.

Let $i \in \{1, \dots, k\}$. We show that $f_i(x)$ can be approximated as closely as desired by $x_i \sigma_i(x_i)$, where σ_i is a univariate SOS, which implies $p(x) = q(x) + \sum_{i=1}^n x_i \sigma_i(x_i)$ where $q(x) \geq 0$ for all $x \in S$.

For any $l \geq 0$ consider the Taylor approximation of e^t with $2l$ terms:

$$T_l(t) = \sum_{j=0}^{2l} \tfrac{t^j}{j!} = 1 + t + \tfrac{t^2}{2!} + \tfrac{t^3}{3!} + ... + \tfrac{t^{2k}}{(2l)!}.$$

Since the Taylor series converges uniformly on bounded intervals, by growing l, one can approximate $f_i(x)$ to any desired accuracy by $a_i x_i T_l(-b_i x_i)$. Hence it is enough to show that $T_l(t)$ is an SOS, or equivalently, given that T_l is a univariate polynomial, that $T_l(t) \geq 0$ for all t [see, e.g., 190]. We prove the non-negativity of T_l by contradiction. Assume T_l is not non-negative. Then it must have a zero as $T_l(0) = 1$. Let t^* be the largest zero of T_l. Then $t^* < 0$ and $T_l'(t^*) > 0$. But for any t, $T_l'(t) = \sum_{j=0}^{2l-1} \frac{t^j}{j!} = T_l(t) - \frac{t^{2l}}{(2l)!}$. Thus $0 < T_l'(t^*) = -\frac{(t*)^{2l}}{(2l)!} < 0$, which is a contradiction. □

Remark 4.33. *From the proof of Lemma 4.32 it follows that one could use the same SOS polynomial $\sigma_i = \sigma$ for $i = 1, \dots, n$ in* (4.16)*.*

Now, we use the representation from Lemma 4.32 and some of the known certificates of non-negativity on compact sets [174, 199] to prove Theorem 4.30.

Proof of Theorem 4.30. Let Δ^n be the standard simplex in $\mathbb{R}^n$: $\Delta^n = \left\{x \in \mathbb{R}^n_+ : e^\intercal x = 1\right\}$. Denote the unit ball in $\mathbb{R}^n$ by $\mathcal{B}^n = \{x \in \mathbb{R}^n : \|x\|^2 \leq 1\}$. We can write

$$\Delta^{n+1} = \{(x_0, x) \in \mathbb{R}^{n+1}_+ : (x_0 + e^\intercal x - 1)^2 = 0\} \cap \mathcal{B}^{n+1}.$$

Since $F \in \operatorname{int} \mathcal{P}_d(\mathbb{R}^n_+)$, we have by Corollary 4.19c that $F^h \in \mathcal{P}^+_d(\mathbb{R}^{n+1}_+ \setminus \{0\}) \in \mathcal{P}^+(\Delta^{n+1})$. Also, from [174, Corollary 2],

$$(x_0 + e^\intercal x)^2 F^h(x_0, x) = q(x_0, x) + h(x_0, x)(x_0 + e^\intercal x - 1)^2, \tag{4.17}$$

where $h \in \mathbb{R}_d[x_0, x]$ and $q \in \mathcal{P}^+_{d+2}\left(\mathbb{R}^{n+1}_+ \cap \mathcal{B}^{n+1}\right)$.

Hence by Lemma 4.32,

$$q(x_0, x) = g(x_0, x) + \sum_{i=0}^{n} x_i \hat{\sigma}_i(x_i), \tag{4.18}$$

where $g(x_0, x) \in \mathcal{P}^+(\mathcal{B}^{n+1})$ and $\hat{\sigma}_0, ..., \hat{\sigma}_n$ are univariate SOS polynomials. By Schmüdgen's Positivstellensatz [199] (Theorem 4.40), we obtain

$$g(x_0, x) = \sigma_0(x_0, x) + \sigma_1(x_0, x)\left(1 - \sum_{i=0}^{n} x_i^2\right) \tag{4.19}$$

where σ_0, σ_1 are SOS polynomials. Now, we use the substitution $(x_0, x) \to \left(\frac{1}{1+e^\intercal x}, \frac{x}{1+e^\intercal x}\right)$, and (4.17)-(4.19) to obtain:

$$\begin{aligned} F^h\left(\tfrac{1}{1+e^\intercal x}, \tfrac{x}{1+e^\intercal x}\right) &= \sigma_0'\left(\tfrac{1}{1+e^\intercal x}, \tfrac{x}{1+e^\intercal x}\right) \\ &\quad + \sigma_1'\left(\tfrac{1}{1+e^\intercal x}, \tfrac{x}{1+e^\intercal x}\right) \sum_{n \geq j > i \geq 0} \tfrac{x_i x_j}{(1+e^\intercal x)^2} + \sum_{i=0}^{n} \hat{\sigma}_i'\left(\tfrac{x_i}{1+e^\intercal x}\right) \tfrac{x_i}{1+e^\intercal x}. \end{aligned} \tag{4.20}$$

Note that: (i) from (4.5), it follows that $F^h(\frac{1}{1+e^\intercal x}, \frac{x}{1+e^\intercal x}) = (1+e^\intercal x)^{-d} F(x)$; (ii) for any even large enough $M \in \mathbb{N}$, we have that if $\sigma(x_0, x)$ is a SOS polynomial, then $(1+e^\intercal x)^M \sigma\left(\frac{1}{1+e^\intercal x}, \frac{x}{1+e^\intercal x}\right)$ is a SOS polynomial in n variables; and (iii) if $\sigma(x_i)$ is a SOS polynomial for any $i = 1, \ldots, n$, then $(1+e^\intercal x)^M \sigma\left(\frac{x_i}{1+e^\intercal x}\right)$ is a *bivariate homogeneous* SOS polynomial in x_i and $(1+e^\intercal x)$. Hence the theorem follows by multiplying (4.20) by $(1+e^\intercal x)^M$ for an even large enough $M \in \mathbb{N}$. □

4.5 Copositive certificates of non-negativity in polynomial optimization

In the spirit of the seminal work of Lasserre [114] and a large body of literature in PO, now we present a convex reformulation of PO problems using Theorem 4.11. More precisely, we reformulate a PO problem as an equivalent linear optimization problem over the cone of copositive polynomials of a known fixed degree. An advantage of using the copositive reformulation is that it allows constructing both inner and outer LMI hierarchies of approximations. Thus, it is possible to obtain arbitrarily close upper and lower bounds to the underlying PO problem. Consider the following standard PO problem:

$$\lambda^* = \inf_x \{p(x) : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}. \tag{4.21}$$

Theorem 4.34. *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$ and $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ be non-empty. Denote $d_{\max} = \max\{\deg h_1, \dots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. Then $\lambda^*_{\text{cop}} \leq \lambda^*$, where*

$$\begin{aligned} \lambda^*_{\text{cop}} = \sup_{\lambda, F} \ & \lambda \\ \text{s. t. } & (1 + e^\intercal y + e^\intercal z)^{2d_{\max} - \deg p}(p(y-z) - \lambda) \\ & \qquad = F(y, z, h_1(y-z), \dots, h_m(y-z)), \\ & F \in \mathcal{P}_{2d_{\max}}(\mathbb{R}^{2n+m}_+), \end{aligned} \tag{4.22}$$

and if $S \subseteq R^n_+$, then $z = 0$ and $F \in \mathcal{P}_{2d_{\max}}(\mathbb{R}^{n+m}_+)$. If $\tilde{p} \in \mathcal{P}^+_{2d_{\max}}(\tilde{S} \setminus \{0\})$, then $\lambda^*_{\text{cop}} = \lambda^*$.

Proof. If (λ, F) is a feasible solution to (4.22), then F is a copositive certificate of non-negativity for $p(x) - \lambda$ on S, that is $\lambda \leq \lambda^*$. Thus $\lambda^*_{\text{cop}} \leq \lambda^*$. Assume $\tilde{p} \in \mathcal{P}^+(\tilde{S} \setminus \{0\})$. If p is unbounded from below on S, then (4.21) is infeasible and its optimal value is $\lambda^* = -\infty$. It then follows from $\lambda^*_{\text{cop}} \leq \lambda^*$ that $\lambda^*_{\text{cop}} = -\infty$ also. Assume therefore that p is bounded from below. Consider any $\lambda < \lambda^*$. Then we have $q := p - \lambda \in \mathcal{P}^+_{2d_{\max}}(S)$, and $\tilde{q} = \tilde{p} \in \mathcal{P}^+_{2d_{\max}}(\tilde{S} \setminus \{0\})$. Hence the result follows by applying Theorem 4.11 to q and S. Finally, if $S \subseteq R^n_+$, then we can set $z = 0$ by Theorem 4.8. □

Corollary 4.35. *Generically, the defining polynomials of a basic semialgebraic set S and a polynomial p are such that if we minimize p on S, then $\lambda^*_{\text{cop}} = \lambda^*$.*

Proof. By genericity of closedness at infinity, we have $\text{cl}(\text{cone}(\{1\} \times S)) = S^h$ from Proposition 4.14. Hence $\tilde{S} = \{x : (0, x) \in S^h\} = S^\infty$. If p is unbounded on S from below, then $\lambda^* = -\infty = \lambda_{\text{cop}}$. If p is bounded on S from below, then $\tilde{p} \in \mathcal{P}(S^\infty)$ by Lemma 4.16 (a). Hence $\tilde{p} \in \mathcal{P}(\tilde{S})$. Therefore generically $p \in \mathbb{R}[x]$ is either unbounded from below on S or has $\tilde{p} \in \mathcal{P}^+(\tilde{S} \setminus \{0\})$. We conclude by applying Theorem 4.34 to the problem of minimizing p on S. □

To numerically use problem (4.22), one can replace the condition $F \in \mathcal{P}_{d_{\max}}(\mathbb{R}^{2n+m}_+)$ by any certificate of copositivity. Possible choices are Pólya's Positivstellensatz [83, Sec. 2.2], the certificate of copositivity we propose in Theorem 4.30, or the certificate of copositivity by Dickinson and Povh [46, Thm. 2.4]. One could also construct convergent inner hierarchies for the cone of copositive tensors based on the method by Bundfuss and Dür [27].

On compact sets, we obtain stronger results. Namely, certificates (4.13) and (4.15) provide convergent lower bounds for (4.21). As an example, we present the usage of certificate (4.13) below.

Corollary 4.36. *Let* $p, h_1, \dots, h_m \in \mathbb{R}[x]$*, and* $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ *be non-empty. Define* $d_{\max} = \max\{\deg h_1, \dots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$*. Let* $M > 0$ *be such that* $S \subseteq \{x \in \mathbb{R}^n : |x| \leq Me\}$*. Let* $d \geq 0$ *and let* $\mathcal{K} \subset \mathcal{P}_d(\mathbb{R}^n)$ *be such that* $\mathbb{R}_+ \subseteq \mathcal{K}$*. For any* $r \in \mathbb{N}$*, define*

$$\begin{aligned}\lambda^r = \sup_{\lambda, (c_{\alpha,\beta,\gamma})} \quad & \lambda \\ \text{s.t.} \quad & p(y - Me) - \lambda = \\ & \sum_{(\alpha,\beta,\gamma) \in \mathbb{N}^{n+m+1}_{d_{\max}(2d_{\max}+r)}} c_{\alpha,\beta,\gamma}(y) y^\alpha h(y - Me)^\beta (2nM - e^\intercal y)^\gamma, \\ & c_{\alpha,\beta,\gamma} \in \mathcal{K} \text{ for } (\alpha, \beta, \gamma) \in \mathbb{N}^{n+m+1}_{d_{\max}(2d_{\max}+r)}.\end{aligned} \tag{4.23}$$

Then $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda^*$*, and* $\lim_{r \to \infty} \lambda^r = \lambda^*$*.*

Proof. Let $y := x + Me$. Then $y \in \mathbb{R}^n_+$ and

$$e^\intercal y \leq e^\intercal x + nM \leq e^\intercal |x| + nM \leq 2nM.$$

Define $T := \{y \in \mathbb{R}^n_+ : h_1(y - Me) \geq 0, \dots, h_m(y - Me) \geq 0\} = \{y \in \mathbb{R}^n_+ : y - Me \in S\}$. Since S is compact and non-empty, T is compact and non-empty. Moreover, $p(y - Me) \in \mathcal{P}^+_{2d_{\max}}(T)$. Hence the conditions of Corollary 4.28 are satisfied for $p(y - Me) \in \mathbb{R}[y]$ on $T \subseteq \mathbb{R}^n_+$. First,

$$\begin{aligned}\lambda^r &\leq \inf\{p(y - Me) : y \in T\} = \inf\{p(y - Me) : y - Me \in S\} \\ &= \inf\{p(x) : x \in S\} = \lambda^*.\end{aligned}$$

Now, notice that if $(\lambda, (c_{\alpha,\beta,\gamma}))$ is feasible for problem (4.23) with r, then it is also feasible for problem (4.23) with $r + 1$. Hence λ^r is non-decreasing in r. To prove the convergence, it is left to show that for any $k > 0$ there is r such that $\lambda^r \geq \lambda^* - \frac{1}{k}$. Compactness of T implies that $p(y - Me)$ is bounded on T from below. Consider any $\lambda^* - \frac{1}{k} < \lambda < \lambda^*$. Then we have $q := p(y - Me) - \lambda \in \mathcal{P}^+_{2d_{\max}}(T)$. Hence, by Corollary 4.28 with $a = 0$, $b = 0$, there is r and $(c_{\alpha,\beta,\gamma})$ such that $(\lambda, (c_{\alpha,\beta,\gamma}))$ is feasible for problem (4.23) with r. Therefore $\lambda^r \geq \lambda \geq \lambda^* - \frac{1}{k}$. □

To obtain more information on λ^* for general sets, we can additionally construct upper bounds on λ^* by applying to problem (4.22) outer – instead of inner – approximations to the cone of copositive polynomials. We do not present this approach here since the resulting bound is an upper bound on λ^* only when (4.22) is a reformulation of (4.21), for instance, when $\tilde{p} \in \mathcal{P}^+(\tilde{S} \setminus \{0\})$. Instead, in Proposition 4.37 we construct an upper bound on λ^* that is always valid.

Proposition 4.37. *Let $p, h_1, \dots, h_m \in \mathbb{R}[x]$ and $S = \{x \in \mathbb{R}^n_+ : h_1(x) \geq 0, \dots, h_m(x) \geq 0\}$ be non-empty. Define $d_{\max} = \max\{\deg h_1, \dots, \deg h_m, \lceil \frac{\deg p}{2} \rceil\}$. Let λ^* be the objective value of (4.21) and let $\lambda_{lb} \leq \lambda^*$ and $\varepsilon > 0$ be given. Define*

$$\begin{aligned}
\lambda_\varepsilon = \sup_{\lambda, F} \ & \lambda \qquad (4.24) \\
\text{s. t.} \ & (1 + e^\intercal y + e^\intercal z)^{2d_{\max} - \deg p}(p(y-z) - \lambda + \varepsilon(1 + e^\intercal y + e^\intercal z)^{\deg p}) \\
& \qquad = F(y, z, h_1(y-z), \dots, h_m(y-z), p(y-z) - \lambda_{lb}), \\
& F \in \mathcal{P}_{2d_{\max}}(\mathbb{R}^{2n+m+1}_+).
\end{aligned}$$

Then $\lambda_\varepsilon \geq \lambda^*$ and $\lim_{\varepsilon \to 0^+} \lambda_\varepsilon = \lambda^*$.

Proof. Let $S_{lb} := \{x \in S : p(x) \geq \lambda_{lb}\}$ and $T_{lb} := \{(y, z) \in \mathbb{R}^{2n}_+ : y - z \in S_{lb}\}$. We have $\tilde{S}_{lb} = \{x \in \tilde{S} : \tilde{p}(x) \geq 0\}$ and $\tilde{T}_{lb} := \{(y, z) \in \mathbb{R}^{2n}_+ : y - z \in \tilde{S}_{lb}\}$. Let $q(y, z) = (1 + e^\intercal y + e^\intercal z)^{\deg p}$. Notice that S_{lb} is non empty, which implies that T_{lb} is non empty. Also, $\tilde{p}(x) \in \mathcal{P}(\tilde{S}_{lb})$, which implies that $\tilde{p}(y - z) \in \mathcal{P}(\tilde{T}_{lb})$. Since $\tilde{q}(y, z) \in \mathcal{P}^+(\mathbb{R}^{2n}_+ \setminus \{0\})$, it follows that $\tilde{p}(y-z) + \varepsilon \tilde{q}(y, z) \in \mathcal{P}^+(\tilde{T}_{lb} \setminus \{0\})$. Then, using Theorem 4.34 we obtain that

$$\begin{aligned}
\lambda_\varepsilon &= \inf\{p(y-z) + \varepsilon q(y, z) : (y, z) \in T_{lb}\} \geq \inf\{p(y-z) : (y, z) \in T_{lb}\} \qquad (4.25) \\
&= \inf\{p(x) : x \in S_{lb}\} = \inf\{p(x) : x \in S\} = \lambda^*
\end{aligned}$$

To show the convergence, notice that λ_ε is non-increasing in ε, and thus it is enough to show that for any $k > 0$ there is ε such that $\lambda_\varepsilon < \lambda^* + \frac{1}{k}$. Fix k, let $x^k \in S$ be such that $p(x^k) < \lambda^* + \frac{1}{2k}$. Define $y^k = \max(x, 0)$, $z^k = -\min(x, 0)$, and let $\varepsilon^k := \frac{1}{2kq(y^k, z^k)}$, we have $\lambda_{\varepsilon^k} \leq p(y^k - z^k) + \frac{q(y^k, z^k)}{2kq(y^k, z^k)} = p(x^k) + \frac{1}{2k} < \lambda^* + \frac{1}{k}$. ☐

To numerically use the upper bound λ_ε, we can use any outer approximation to the set of copositive polynomials. Some examples are the simplicial partitions approach by Bundfuss and Dür [27], the simplex discretization approach by Yildirim [231] and the moment matrices approach by Lasserre [117, 120].

Proposition 4.37 illustrates how the copositive certificates of non-negativity proposed in Theorems 4.8 and 4.11 can be used, in contrast to the use of classical certificates of non-negativity [see, e.g., 114], to obtain not only lower but also upper bounds for PO problems. This allows obtaining realistic estimates of how far the convergent lower bounds from Corollary 4.36 are from the optimal value λ^* of problem (4.21). Besides improving estimates for λ^*, the proposed construction of bounds extends the range of applications of results specific for copositive polynomials (such as Pólya's theorem or the results from [27, 46, 117, 231]) to general basic semialgebraic sets.

4.6 Relationship to Handelman's and Schmüdgen's Positivstellensatzen

In this section we obtain Handelman's Positivstellensatz [82] (see Theorem 1.13 in Chapter 1) and Schmüdgen's [199] Positivstellensatz (see Theorem 1.10 in Chapter 1) using Corollary 4.26. Although we mentioned both Positivstellensatzen in Chapter 1, we recite them later in this section for convenience. Besides the cited classical proofs by Handelman and Schmüdgen, there are a few other proofs [18, 182, 206] for the first theorem and [18, 204, 206] for the second one. The alternative proofs exploit tools from various fields of mathematics, but mainly abstract algebra. Our proofs are different in the sense that we use Corollary 4.26 and standard optimization tools, with minimum use of algebraic tools. Our approach to Schmüdgen's theorem partially follows the approach of Schweighofer [206]. Both our proof and the proof in [206] exploit a result by Berr and Wörmann [18] (Proposition 4.41). In both our approach and the proof of Schweighofer [206], the polynomial $p(x)$ is associated with some copositive polynomial $F(x, h_1(x), \dots, h_m(x))$. This polynomial is homogenized, and Pólya's theorem (Theorem 1.8) is applied to it. However, the ways in which the existence of $F(x, h_1(x), \dots, h_m(x))$ is established are different: our reasoning goes through Corollary 4.26, while Schweighofer [206] uses tools from algebraic geometry.

We apply our approach first to prove Handelman's Positivstellensatz [82], which we present below for convenience of the reader.

Theorem 4.38 (Handelman's Positivstellensatz [82])**.** *Let $A \in \mathbb{R}^{m\times n}$, and let $S = \{x : Ax \le b\}$ be a non-empty polytope. Let $p \in \mathcal{P}^+(S)$. Then*

$$p(x) = \sum_{\alpha\in\mathbb{N}^m} c_\alpha (b - Ax)^\alpha,$$

for some $c_\alpha \ge 0$ for all $\alpha \in \mathbb{N}^m$.

For our alternative proof, we use the following version of Farkas' lemma.

Proposition 4.39 (Ziegler [232, Proposition 1.9])**.** *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$ be such that $S = \{x : Ax \le b\}$ is non-empty. If $c_0 + c^\intercal x \in \mathcal{P}_1(S)$, then there exist $u, u_0 \ge 0$ such that $c_0 + c^\intercal x = u^\intercal(b - Ax) + u_0$.*

Proof of Theorem 4.38. Let $\hat{x}_i = \min_{x\in S} x_i$. We use the translation $x \to y - \hat{x}$. Define $S' = \{y \in R^n : A(y - \hat{x}) \le b\} \subset \mathbb{R}^n_+$ so that $S' \subset \mathbb{R}^n_+$. Clearly, S' is non empty. Also, as S is compact, S' is compact and there is $M > 0$ such that $S' \subseteq \{y \in \mathbb{R}^n_+ : e^\intercal y \le M\}$. From Corollary 4.26, after letting $a = 0$ and $b = 0$, it follows that as $p(y - \hat{x}) \in \mathcal{P}^+(S')$, there exists $d \ge 0$ and $c_{\alpha,\beta,\gamma} \ge 0$ for $(\alpha, \beta, \gamma) \in \mathbb{N}^{n+m+1}$ such that

$$p(y - \hat{x}) = \sum_{(\alpha,\beta,\gamma)\in\mathbb{N}^{n+m+1}} c_{\alpha,\beta,\gamma} y^\alpha (A(y - \hat{x}) - b)^\beta (M - e^\intercal y)^\gamma.$$

Now, substitute back $y \to x + \hat{x}$. We have $x_i + \hat{x}_i \in \mathcal{P}_1(S)$ for all $i = 1, \ldots, n$, and $M - e^\intercal(x + \hat{x}) \in \mathcal{P}_1(S)$. The result then follows by using Proposition 4.39 to replace $x_i + \hat{x}_i$ for all $i \in \{1, \ldots, n\}$, and $M - e^\intercal(x + \hat{x})$ in the representation above, respectively by expressions of the form ${u^i}^\intercal(b - Ax) + u_0^i$ for some $u^i, u_0^i \geq 0$, $i = 1, \ldots, n+1$. □

Now we prove Schmüdgen's Positivstellensatz, which we again present below for convenience.

Theorem 4.40 (Schmüdgen's Positivstellensatz [199])**.** *Let $p, h_1, \ldots, h_m \in \mathbb{R}[x]$ be such that $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, ..., h_m(x) \geq 0\}$ is non-empty and compact, and let $p \in \mathcal{P}^+(S)$. Then there is $r \geq 0$ such that*

$$p = \sum_{\alpha \in \{0,1\}^m} \sigma_\alpha h^\alpha, \tag{4.26}$$

for some SOS polynomials σ_α of degree $r - \deg(h^\alpha)$ for all $\alpha \in \{0,1\}^m$.

The approach to prove Theorem 4.40 is the same used to prove Theorem 4.38. First we use a weaker result that allows us to add redundant constraints to the semialgebraic set S that can then be written in terms of the original constraints defining S. For that we use a result by Berr and Wörmann [18].

Proposition 4.41 (Berr and Wörmann [18], Schweighofer [206])**.** *Let $h_1, \ldots, h_m \in \mathbb{R}[x]$ be such that $S = \{x \in \mathbb{R}^n : h_1(x) \geq 0, ..., h_m(x) \geq 0\}$ is non-empty and compact. Then for every polynomial $p \in \mathbb{R}[x]$ there exists $t \in \mathbb{R}_+$ such that $t + p$ and $t - p$ have a representation of the form* (4.26)*.*

The proposition is weaker than Schmüdgen's Positivstellensatz: it holds for every $p \in \mathbb{R}[x]$ and does not require positivity of p on S. Intuitively, since S is bounded, one can make the minimum of $t \pm p$ as large as desired by growing t. Theorem 4.40 shows that for $p > 0$ on S, there is representation of the form (4.26) for $t + p$, where $t = 0$.

Proof of Theorem 4.40. For $i = 1, \ldots n$, apply Proposition 4.41 on S to obtain $\hat{x} \in \mathbb{R}^n_+$ such that for $i = 1, \ldots, n$,

$$x_i + \hat{x}_i = \sum_{\alpha \in \{0,1\}^m} \sigma_\alpha^i(x) h(x)^\alpha. \tag{4.27}$$

and $M > 0$ such that

$$M - e^\intercal(x + \hat{x}) = \sum_{\alpha \in \{0,1\}^m} \hat{\sigma}_\alpha(x) \hat{h}(x)^\alpha. \tag{4.28}$$

Next, apply the translation $x \to y - \hat{x}$ obtaining $S' = \{y \in \mathbb{R}^n_+ : y - \hat{x} \in S\}$. Then we have that $S' \subseteq \mathbb{R}^n_+$ and non empty. From Corollary 4.26, after letting $a = 0$ and $b = 0$, it follows that as $p(y - \hat{x}) \in \mathcal{P}^+(S')$, there exists $d \geq 0$ and $c_{\alpha,\beta,\gamma} \geq 0$ for $(\alpha, \beta, \gamma) \in \mathbb{N}^{n+m+1}$ such that

$$p(y - \hat{x}) = \sum_{(\alpha,\beta,\gamma)\in\mathbb{N}^{n+m+1}} c_{\alpha,\beta,\gamma} y^\alpha h(y - \hat{x})^\beta (M - e^\intercal y)^\gamma, \tag{4.29}$$

The result then follows after replacing y by $x + \hat{x}$, substituting representations (4.27) and (4.28) into (4.29), expanding, and using the fact that the product of SOS polynomials is a SOS polynomial. □

4.7 Conclusion and questions for further research

In this chapter we propose copositive certificates of non-negativity of polynomials over semialgebraic sets. We show that under some mild assumptions, such a copositive certificate of small and known degree exists on a given basic closed semialgebraic set, not necessarily compact (see Theorems 4.8 and 4.11). Moreover, these assumptions hold generically. Certifying copositivity is an NP-hard problem. However, one can use existing outer and inner approximations to the set of copositive polynomials. These approximations, in combination with the copositive certificates we propose, deliver new results about the non-negativity of polynomials over generic semialgebraic sets. In particular, we obtain LMI hierarchies of upper and lower bounds on polynomial optimization problems and derive new structured certificates of non-negativity on compact sets. A question for future research is to evaluate the performance of the bounds and certificates we propose in optimization. Another question is to compare the performance of our approach with the performance of classical approximations to PO problems, such as Lasserre's hierarchy [114].

4.7.1 Proof of Lemma 4.7

Preliminaries about the notation in this section are provided in Section 1.1.1 of Chapter 1. Let $\mathcal{C}^n_d$ denote the cone of completely positive tensors:

$$\mathcal{C}^n_d := \text{cone}\left\{x^{\otimes d} : x \in \mathbb{R}^n_+\right\}. \tag{4.30}$$

Notice that $\mathcal{C}^n_d \in \mathbb{S}^n_d$, and if $T \in \mathcal{C}^n_d$, then for any $u \in [n]^d$ we have $T(u_1, \ldots, u_d) = x_{u_1} \cdots x_{u_d}$.

Proposition 4.42 ([175])**.** *For any $d, n > 0$, $\mathcal{C}^n_d$ is a proper cone (closed, pointed, convex, with non-empty interior).*

Now, let $\mathcal{COP}_d^n$ be the dual cone of $\mathcal{C}_d^n$, i.e.,

$$\mathcal{COP}_d^n := \{T \in \mathbb{S}_d^n : \langle T, C \rangle \geq 0 \text{ for all } C \in \mathcal{C}_d^n\}, \tag{4.31}$$

where the inner product of two tensors is defined in (1.2). Since $\mathcal{C}_d^n$ is proper, so is $\mathcal{COP}_d^n$. We call $\mathcal{COP}_d^n$ the cone of copositive tensors.

To connect copositive polynomials and completely positive tensors, we introduce the operator $C_d(p) : \mathbb{R}_d[x] \to \mathbb{S}_d^{n+1}$. For $p \in \mathbb{R}_{=d}[x]$ define its homogenization p^h as in (4.5). We can write $p^h(x_0, x) = \sum_{\alpha \in \mathbb{N}^{n+1}, e^\intercal \alpha = d} x_0^{\alpha_1} \dots x_n^{\alpha_{n+1}} p_\alpha^h$. For any $u \in [n+1]^d$ we define

$$C_d(p)(u_1, \dots, u_d) := \frac{\alpha_1! \dots \alpha_{n+1}!}{d!} p_\alpha^h, \tag{4.32}$$

where for every $k \in [n+1]$, α_k is the number of times $k-1$ appears in $u \in [n+1]^d$. With this notation, for any $p \in \mathbb{R}_d[x]$ and $a \in \mathbb{R}^n$ we have

$$p(a) = \langle C_d(p), \left[\begin{smallmatrix} 1 \\ a \end{smallmatrix}\right]^{\otimes d} \rangle \tag{4.33}$$

The latter is equivalent to writing $p(x) = p^h(1, x)$.

Let $q, h_1, ..., h_m \in \mathbb{R}_{=d}[x]$ and consider the following pair of problems

$$\begin{aligned} \inf\ & q(x) \\ \text{s. t. } & h_i(x) = 0 && \text{for } i = 1, ..., m \\ & x \geq 0 \end{aligned} \tag{4.34}$$

$$\begin{aligned} \inf\ & \langle C_d(q), Y \rangle \\ \text{s. t. } & \langle C_d(h_i), Y \rangle = 0 && \text{for } i = 1, ..., m \\ & \langle C_d(1), Y \rangle = 1 \\ & Y \in C_d^{n+1} \end{aligned} \tag{4.35}$$

Proposition 4.43 (Theorem 5 from [175])**.** *Let $q, h_1, ..., h_m \in \mathbb{R}_{=d}[x]$. Consider a non-empty set $S = \left\{x \in \mathbb{R}_+^n : h_1(x) = 0, \dots, h_m(x) = 0\right\}$. Assume that the following conditions hold for all $j \in [m]$:*

(a) $h_j(x) \geq 0$ for all $x \in \mathbb{R}_+^n$,

(b) $\tilde{q}(x) \in \mathcal{P}_d(\tilde{S})$.

Then the optimal values problems (4.34) and (4.35) are the same and one of the problems attains its optimal value if and only if the other one does.

We use Proposition 4.43 together with the standard conic dual problem of problem (4.35):

$$\begin{aligned} &\sup \lambda && (4.36)\\ &\text{s. t. } C_d(q) - \lambda C_d(1) - \sum_{i=1}^{m} C_d(h_i) y_i \in \mathcal{COP}_d^{n+1} \end{aligned}$$

It is important for our proof that strong duality holds between problems (4.35) and (4.36).

Theorem 4.44. *For $d > 0$ let $q, h_1, \dots, h_m \in \mathbb{R}_{=d}[x]$. Consider a non-empty set $S = \left\{x \in \mathbb{R}^n_+ : h_1(x) = 0, \dots, h_m(x) = 0\right\}$. If $q \in \mathcal{P}_d^+(S)$, $\tilde{q} \in \mathcal{P}_d^+(\tilde{S} \setminus \{0\})$, and the assumptions of Proposition 4.43 hold, then strong duality holds between problems* (4.35) *and* (4.36)*, and problem* (4.36) *has a feasible solution with $\lambda \geq 0$.*

Proof. $S \subseteq \mathbb{R}^n_+$ and is non-empty, hence problem (4.34) is feasible. Therefore problem (4.35) is also feasible from (4.33). Since $q \in \mathcal{P}^+(S)$, the optimal value of problem (4.34) is finite and positive. The conditions of Proposition 4.43 are satisfied. Hence the optimal value of problem (4.35) is equal to the optimal value of problem (4.34), is finite and positive. Therefore, if strong duality holds between problems (4.35) and (4.36), problem (4.36) has a feasible solution with $\lambda \geq 0$.
Now, we prove the statement about strong duality. Consider the following cone:

$$\mathcal{K} = \left\{(\langle C_d(h_1), Y\rangle, ..., \langle C_d(h_m), Y\rangle, \langle C_d(1), Y\rangle, \langle C_d(q), Y\rangle) : Y \in C_d^{n+1}\right\} \subset \mathbb{R}^{m+2}.$$

the fact that $\mathcal{K}$ is a cone follows from the definition of the tensor inner product and from $Y \in C_d^{n+1}$. We claim that $\mathcal{K}$ is closed. In this case, since problem (4.35) is feasible, by Theorem ("zero duality gap") in Chapter IV.7.2 in Barvinok [15], strong duality holds between problems (4.35) and (4.36). To finish the proof of the theorem, we show that $\mathcal{K}$ is closed.
Let $(K^k)_{k\in\mathbb{N}} \subset \mathcal{K}$ and $K = \lim_{k\to\infty} K^k$. Show that $K \in \mathcal{K}$. For each $k \in K^k$ by definition there is $Y^k \in C_d^{n+1}$. First let the sequence $(Y^k)_{k\in\mathbb{N}}$ have a convergent subsequence. Since C_d^{n+1} is closed, the limit Y of this subsequence is in C_d^{n+1}. Therefore, assuming that we work on a convergent subsequence of $(Y^k)_{k\in\mathbb{N}}$ and the corresponding subsequence of $(K^k)_{k\in\mathbb{N}}$,

$$\begin{aligned} K = \lim_{k\to\infty} K^k &= \lim_{k\to\infty} \left(\langle C_d(h_1), Y^k\rangle, ..., \langle C_d(h_m), Y^k\rangle, \langle C_d(1), Y^k\rangle, \langle C_d(q), Y^k\rangle\right)\\ &= \left(\langle C_d(h_1), \lim_{k\to\infty} Y^k\rangle, ..., \langle C_d(h_m), \lim_{k\to\infty} Y^k\rangle, \langle C_d(1), Y^k\rangle, \langle C_d(q), \lim_{k\to\infty} Y^k\rangle\right)\\ &= (\langle C_d(h_1), Y\rangle, ..., \langle C_d(h_m), Y\rangle, \langle C_d(1), Y\rangle, \langle C_d(q), Y\rangle) \text{ for some } Y \in C_d^{n+1}. \end{aligned}$$

Thus $K \in \mathcal{K}$. Now let the sequence $(Y^k)_{k\in\mathbb{N}}$ have no convergence subsequence. We show that this contradicts to the assumption that the sequence $(K^k)_{k\in\mathbb{N}}$ converges.

By definition, for every $k \in \mathbb{N}$,

$$Y^k = \sum_{i=1}^{p} \underbrace{\begin{bmatrix} x_0^k \\ x^k \end{bmatrix}_i \otimes \begin{bmatrix} x_0^k \\ x^k \end{bmatrix}_i \otimes \cdots \otimes \begin{bmatrix} x_0^k \\ x^k \end{bmatrix}_i}_{d},$$

where $\begin{bmatrix} x_0^k \\ x^k \end{bmatrix}_i \in \mathbb{R}_+^{n+1}$ for all $i \in [p]$. Since $(Y^k)_{k\in\mathbb{N}}$ is contained in a closed set and has no convergent subsequence, it is unbounded. Therefore the sequence $\left(\begin{bmatrix} x_0^k \\ x^k \end{bmatrix}_i\right)_{k\in\mathbb{N}}$ is unbounded for some $i \in [p]$. Denote this sequence by $\left(\begin{bmatrix} y_0^k \\ y^k \end{bmatrix}\right)_{k\in\mathbb{N}}$.
Let K_j denote the j^{th} element of K. By the assumptions of the theorem, $h_j \in \mathcal{P}_d(\mathbb{R}_+^n)$ for all $j \in [m]$. Therefore, since $\tilde{\mathbb{R}}_+^n = \mathbb{R}_+^n$ and using Lemma 4.16, we have $\tilde{h}_j \in \mathcal{P}_d(\mathbb{R}_+^n)$. Hence $h_j^h \in \mathcal{P}_d(\mathbb{R}_+^n)$ by (4.6) and

$$K_j = \lim_{k\to\infty} K_j^k = \lim_{k\to\infty} \langle C_d(h_j), Y^k\rangle = \lim_{k\to\infty} \sum_{i=1}^{p} h_j^h(x_{0,i}^k, x_i^k) \geq \lim_{k\to\infty} h_j^h(y_0^k, y^k),$$

$$K_{m+1} = \lim_{k\to\infty} K_{m+1}^k = \lim_{k\to\infty} \langle C_d(1), Y^k\rangle = \lim_{k\to\infty} \sum_{i=1}^{p} (x_{0,i}^k)^d \geq \lim_{k\to\infty} (y_0^k)^d,$$

Let $\Delta^{n+1} = \{x \in \mathbb{R}^{n+1} : e^\intercal x = 1\}$ be the standard simplex in $\mathbb{R}^{n+1}$. Since $\begin{bmatrix} y_0^k \\ y^k \end{bmatrix} \in \mathbb{R}_+^{n+1}$ is unbounded, for every $k \in \mathbb{N}$ large enough we can write

$$\begin{bmatrix} y_0^k \\ y^k \end{bmatrix} = (y_0^k + e^\intercal y^k) \begin{bmatrix} \bar{y}_0^k \\ \bar{y}^k \end{bmatrix}, \text{ where } \begin{bmatrix} \bar{y}_0^k \\ \bar{y}^k \end{bmatrix} = \begin{bmatrix} y_0^k/(y_0^k + e^\intercal y^k) \\ y^k/(y_0^k + e^\intercal y^k) \end{bmatrix} \in \Delta^{n+1}.$$

Therefore we have for all $j \in [m]$

$$K_j \geq \lim_{k\to\infty} h_j^h(y_0^k, y^k) = \lim_{k\to\infty} h_j^h(\bar{y}_0^k, \bar{y}^k)(y_0^k + e^\intercal y^k)^d,$$
$$K_{m+1} \geq \lim_{k\to\infty} (y_0^k)^d = \lim_{k\to\infty} (\bar{y}_0^k)^d (y_0^k + e^\intercal y^k)^d.$$

Since $\begin{bmatrix} y_0^k \\ y^k \end{bmatrix} \in \mathbb{R}_+^{n+1}$, the sequence $(y_0^k + e^\intercal y^k)_{k\in\mathbb{N}}$ diverges. The inequality above for K_{m+1} implies then that $\lim_{k\to\infty} \bar{y}_0^k = 0$. Hence the inequality above for K_j implies that for all $j \in [m]$ $\lim_{k\to\infty} h_j^h(\bar{y}_0^k, \bar{y}^k) = 0$. Since $\left(\begin{bmatrix} \bar{y}_0^k \\ \bar{y}^k \end{bmatrix}\right)_{k\in\mathbb{N}}$ is a sequence in Δ^{n+1}, it has a convergent subsequence with the limit $\begin{bmatrix} 0 \\ \bar{y} \end{bmatrix} \in \Delta^{n+1}$. Further we assume that we work on this convergent subsequence. As h_j^h is continuous and using (4.6),

$$0 = \lim_{k\to\infty} h_j^h(\bar{y}_0^k, \bar{y}^k) = h_j^h(0, \bar{y}) = \tilde{h}_j(\bar{y}).$$

Hence $\bar{y} \in \tilde{S}$. If $\tilde{S} = \{0\}$, we immediately obtain a contradiction since $\bar{y} \in \Delta^{n+1}$. If $\tilde{S} \neq \{0\}$, we use $\tilde{q}(\bar{y}) > 0$. Then

$$\lim_{k\to\infty} q^h(\bar{y}_0^k, \bar{y}^k) = q^h(0, \bar{y}) = \tilde{q}(\bar{y}) > 0.$$

Since q has degree d,

$$K_{m+2} = \lim_{k\to\infty} \langle C_d(q), Y^k\rangle \geq \lim_{k\to\infty} q^h(y_0^k, y^k) = \lim_{k\to\infty} q^h(\bar{y}_0^k, \bar{y}^k)(y_0^k + e^\intercal y^k)^d = \infty.$$

This contradicts to the fact that K_{m+2} is finite. Hence the sequence $(Y^k)_{k\in\mathbb{N}}$ is bounded and the cone $\mathcal{K}$ is closed. □

Proof of Lemma 4.7. We have $\deg p = d$, $p \in \mathcal{P}^+(S)$ and $\tilde{p} \in \mathcal{P}^+(\tilde{S} \setminus \{0\})$. Hence there exists $\varepsilon > 0$ such that $q(x) = p(x) - \varepsilon(1 + e^\intercal x)^d \in \mathcal{P}^+(S)$ and $\tilde{q} \in \mathcal{P}^+(\tilde{S} \setminus \{0\})$. Since $S \subseteq \mathbb{R}^n_+$, we can write $S = \left\{x \in R^n_+ : g_1(x) = 0, \dots, g_m(x) = 0\right\}$, where $g_j(x) = (1+e^\intercal x)^{d-\deg h_j} h_j(x)$ for all $j \in [m]$. Then the conditions of Theorem 4.44 are satisfied for q and S. Therefore problem (4.36) is feasible with some $\lambda \geq 0$ and from (4.33) we obtain

$$p(x) - \varepsilon(1 + e^\intercal x)^d = \lambda + G(x) + \sum_{j=1}^{m} y_j(x) g_j(x),$$

where $\deg G = d,\ G \in \mathcal{P}(\mathbb{R}^n_+)$ and $y_j \in \mathbb{R}_{d-\deg g_j}[x]$ for all $j \in \{1, \dots, m\}$.
Finally, denote $F(x) := \varepsilon(1+e^\intercal x)^d + \lambda + G(x)$. Then $F \in \operatorname{int} \mathcal{P}(\mathbb{R}^n_+)$ by Corollary 4.19 (b). □

CHAPTER 5

New SDP upper bounds on the maximum k-colorable subgraph problem

5.1 Introduction

For a given graph, the maximum k-colorable subgraph (MkCS) problem is to find the largest induced subgraph that can be colored in k colors such that no two adjacent vertices have the same color. The MkCS is also known as the maximum k-partite induced subgraph problem since the k-coloring corresponds to a k-partition of the subgraph. A natural extension of this problem is to assign weights to the vertices of the graph and ask for the k-colorable subgraph of the maximum weight.

We remark that in the literature the name "maximum k-colorable subgraph problem" is sometimes used for the maximum k-cut problem [67, 169]. In the latter problem one searches for the partition of the graph into k subsets such that the number of edges crossing the parts is maximized. If we color vertices in the resulting subsets in different colors, the crossing edges are properly colored, that is, the endpoints of these edges have different colors. Hence the goal of the maximum k-cut problem is to find a maximum properly k-colorable subgraph. This definition shows that the MkCS is different from the maximum k-cut problem. We do not consider the latter problem in this chapter, and refer interested readers to [67, 72, 169, 186, 212, 215] for more information on the maximum k-cut problem and semidefinite programming (SDP) relaxations for it.

The MkCS falls into the class of NP-complete problems considered by Lewis and Yannakakis [126]. Moreover, even approximating this problem is hard [133]. For $k = 1$ the MkCS reduces to the famous maximum stable set problem, which has been shown to be NP-complete by Karp [104]. Another well-known problem from the list of Karp [104] related to the MkCS is the chromatic number problem: to determine whether the vertices of a given graph can be colored in k colors. If we can solve the MkCS for a given number of colors, we can clearly solve the maximum stable set and the chromatic number problems. On the other hand, the MkCS on a given graph can be formulated as an instance of the stable set problem on the Cartesian product of that graph and the complete graph on k vertices. However, not all efficient algorithms for the maximum stable set and the chromatic number problems result in efficient algorithms for the MkCS. For instance, on perfect graphs, the chromatic number and the stability number can be computed in polynomial time while the MkCS is NP-complete on some perfect graphs [229]. In this sense the MkCS is harder than the maximum stable set and the chromatic number problems. Nevertheless, the MkCS is polynomial-time solvable on special classes of graphs. Some examples are graphs where every odd cycle has two non-crossing chords for any k [1], clique-separable graphs for $k = 2$ [1], chordal graphs for fixed k [229], interval graphs for any k [229], circular-arc graphs and tolerance graphs for $k = 2$ [152].

In contrast to the stable set and chromatic number problems, the MkCS has been rarely considered in the literature. Januschowski and Pfetsch [93] notice that such a lack of attention might be precisely related to the connection of the MkCS to the earlier mentioned prominent problems. However, MkCS has a number of applications, such as channel assignment in spectrum sharing networks (Wi-Fi or cellular) [16, 81, 85, 214], VLSI problems [66, 137] and human genetic research [66, 127]. Therefore we believe that the k-colorable subgraph deserves a thorough study on its own.

Besides the case $k = 1$, i.e., the stable set problem, the second case that has attracted some attention in the literature is $k = 2$. The corresponding problem is called the maximum bipartite subgraph problem. It has been studied, among others, by Fouilhoux and Mahjoub [65], Lee et al. [123], and Hüffner [88]. The case with $k > 2$ is the least studied. Some significant sources of information on this case are Narasimhan [152] and Narasimhan and Manber [153]. The latter paper introduces an upper bound on the optimal value of the MkCS called "the generalized ϑ-number". This name is after the famous ϑ-number by Lovász [129], which is an upper bound on the size of the largest stable set of a graph. For a given graph, the generalized ϑ-number is the minimum sum of the k largest eigenvalues over the family of matrices that have ones on the diagonal and in the entries corresponding to the non-edges of the graph. For $k = 1$ this corresponds to one of the definitions of the ϑ-number by Lovász [129].

Narasimhan and Manber [153] proposed their bound at the end of the 1980s. At that time, providing computational results on such a problem was not possible. Alizadeh [4] formulated the generalized ϑ-number problem by [153] using SDP and included it within the promising applications of the interior point methods for SDP. Also, Mohar and Poljak [144] presented the bound by Narasimhan and Manber [153] among important applications of eigenvalues of graphs in combinatorial optimization. However, to our knowledge, the quality of the generalized ϑ-number by Narasimhan and Manber [153] is still not evaluated.

Other related works look at integer programming (IP) formulations of the MkCS and solve or approximate them using linear programming (LP) [30, 93, 94]. Januschowski and Pfetsch [93], and Campêlo and Corrêa [30] provide computational results for some $k > 2$ and large graphs up to 1085 vertices. Finally, Hertz et al. [86] analyze the performance of various existing online algorithms on the MkCS.

Outline and main results

In this chapter we propose several SDP upper bounds on the optimal value of the MkCS. We evaluate the quality of these bounds numerically and find that our bounds outperform the existing approaches on all tested graphs. We work with the basic version of the MkCS, but our results extend immediately to the weighted version of the problem.

We begin our analysis of the MkCS with the bound by Narasimhan and Manber [153]. We write it as an SDP following the approach of Alizadeh [4] and show how to obtain the same bound starting with an IP formulation of the MkCS. Our next question of interest is how to tighten the generalized ϑ-number. Recall that for $k = 1$ the generalized ϑ-number equals the ϑ-number by Lovász [129]. A natural way to tighten the latter bound is to add non-negativity constraints to one of its SDP formulations. This results in the so-called Schrijver ϑ'-number [202]. Applying the same procedure to the generalized ϑ-number, we obtain the generalized ϑ'-number. For any k, both the generalized ϑ- and ϑ'-numbers require solving an SDP with one matrix of order n, for a graph with n vertices.

We proceed by considering alternative SDP upper bounds for the MkCS. For this purpose we use vector lifting or matrix lifting SDP relaxations of one of the classical IP formulations of the MkCS. The size of the resulting relaxations depends on n and k, and the larger k is, the larger the SDP matrices involved in the problems are. We reduce the sizes of the SDPs by exploiting the invariance of the MkCS under permutations of the colors; that is, it does not matter which label is assigned to each color in our SDP relaxations. All constraints in the problems are satisfied with any labeling, and the objective does not change if the labeling changes. This property is inherited by our SDP relaxations from the IP formulations of the MkCS. By exploiting the invariance, our strongest SDP relaxation reduces to a problem with two SDP constraints of the order $n+1$ and n, respectively, for a graph with n vertices. This matrix size does not depend on the value of k or the type of graph. Notice that the size of the SDP formulation of the generalized ϑ-number by Alizadeh [4] does not depend on the number of colors k and cannot be reduced using the invariance under the color permutations.

We evaluate the quality of all SDP bounds on graphs used in the earlier research [30, 93] with up to 200 vertices. Januschowski and Pfetsch [93] solve the MkCS for some graphs to optimality, and we compare our bounds to these optimal values. For some of the tested graphs, we obtain the optimal values too. Campêlo and Corrêa [30] propose an IP formulation of the MkCS by representatives and implement a Lagrangian decomposition of that formulation. To our knowledge, [30] contains the only available numerical experiments with upper bounds for the MkCS, which motivates the choice of [30] as a benchmark. Our computational results show that even the weakest among our SDP bounds is at least as good as both the generalized ϑ-number by Narasimhan and Manber [153] and its strengthening ϑ'-number. At the same time, the generalized ϑ-number is at least as good as the bounds from [30] for six out of eight compared graphs.

We introduce the MkCS formulation in Section 5.2. Next, in Section 5.3 we look at the generalized ϑ-number by Narasimhan and Manber [153] and strengthen this bound. In Section 5.5 we propose several new SDP bounds and compare them to each other. In section 5.6 we present the symmetry reduction procedure for problems from Section 5.5. Section 5.7 contains numerical experiments. We compare the performance of all SDP bounds to the performance of other bounds existing in the literature. We summarize the results and list some questions for future research in Section 5.8.

5.2 Problem formulation

In this section we formally introduce the maximum k-colorable subgraph problem.

Let $G = (V, E)$ be a simple undirected graph with the vertex set V and the edge set E. Let $|V| = n$, and let k be a given integer such that $1 \leq k \leq n - 1$. We say that G is k-colorable if one can assign to each vertex in G one of the k colors such that adjacent vertices in G do not have the same color. A graph $G' = (V', E')$ is called an induced subgraph of a given graph $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$ is such that E' contains all edges in E connecting the vertices in V'.

The maximum k-colorable subgraph problem is to find an induced k-colorable subgraph with maximum cardinality in the given graph. Any k-colorable subgraph is also k-partite, so MkCS is also called the maximum k-partite subgraph problem. We denote by $\alpha(G)$ and $\alpha_k(G)$ for $k \geq 2$ the stability number (the maximum cardinality of a stable set) and the maximum number of vertices in a k-partite subgraph of G, respectively.

The MkCS can be formulated as an IP. Let $X \in \{0,1\}^{n \times k}$ be the matrix with one in the entry (i, r) if vertex $i \in [n]$ is colored with color $r \in [k]$ and zero otherwise. We use $[n]$ to denote the set $\{1, \ldots, n\}$. The MkCS can be formulated as the following

IP:

$$\begin{aligned} \alpha_k(G) = \max_{X \in \{0,1\}^{n\times k}} \quad & \sum_{i\in[n], r\in[k]} X_{ir} && && (5.1)\\ \text{s. t.} \quad & X_{ir} + X_{jr} \leq 1, && \text{for all } \{ij\} \in E, r \in [k] \\ & \sum_{r\in[k]} X_{ir} \leq 1, && \text{for all } i \in [n]. \end{aligned}$$

This formulation is also used by Januschowski and Pfetsch [93, 94] to whose results we refer in Section 5.7. In Section 5.5 we provide another integer programming formulation for the MkCS problem that contains nonlinear constraints and is more suitable for deriving semidefinite programming relaxations.

5.3 The generalized ϑ-number

In this section we review the so-called generalized ϑ-number by Narasimhan and Manber [153], which is an eigenvalue upper bound for the MkCS, and its SDP formulation by Alizadeh [4]. Finally, we strengthen the resulting SDP bound and obtain the generalized ϑ'-number.

5.3.1 Eigenvalue and SDP formulations of the generalized ϑ-number

For a matrix $A \in \mathbb{R}^{n\times n}$, let $\lambda_i(A)$ be the i^{th} largest eigenvalue of A. For $A, B \in \mathbb{S}^n$, the trace inner product of A and B is denoted by $\langle A, B\rangle := \sum_{i,j\in[n]} A_{ij}B_{ij}$. Throughout the chapter we use the notation I_m (resp. J_m) for the identity matrix (resp. the matrix of all ones) of dimension $m\times m$. If the dimension of the matrices is clear from the context, we omit the subsctipt. Narasimhan and Manber [153] introduce the following upper bound for $\alpha_k(G)$:

$$\alpha_k(G) \leq \vartheta_k(G) = \min_{A\in\mathbb{S}^n} \left\{ \sum_{r=1}^{k} \lambda_r(A) : \; A_{ij} = 1 \text{ for } \{ij\} \notin E \text{ or } i = j \right\}. \qquad (5.2)$$

Notice that the minimum in problem (5.2) is attained since the maximal eigenvalue of feasible matrices can be bounded from above without loss of generality. This implies that the absolute values of entries of feasible matrices can be bounded, see, e.g., Lemma 4 in Narasimhan and Manber [153] for the proof. Therefore one can restrict the optimization to a compact subset of the feasible set. To show that we can bound the maximal eigenvalue, notice that problem (5.2) is feasible and the trace of all feasible solutions is n. Therefore $\vartheta_k(G)$ is finite. Let $(A_m)_{m>0}$ be a sequence of solutions to (5.2) such that $\lim_{m\to\infty} \sum_{i=1}^{k} \lambda_i(A_m) = \vartheta_k(G)$. As $\vartheta_k(G)$ is finite and for any m the trace of A_m is fixed, $\lambda_1(A_m)$ is bounded from above. Otherwise $\lambda_{k+1}(A_m), \ldots, \lambda_n(A_m)$ are unbounded from below, which leads to a contradiction.

To show that $\vartheta_k(G)$ is an upper bound on $\alpha_k(G)$, we follow the reasoning of Mohar and Poljak [144] who use Fan's theorem.

Theorem 5.1 (Fan [64]). *Let A be a symmetric matrix with eigenvalues $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$. Then*

$$\sum_{i=1}^{k} \lambda_i(A) = \max_X \left\{ \langle A, XX^\mathsf{T} \rangle \ \text{ s.t. } \ X^\mathsf{T}X = I_k \right\}. \tag{5.3}$$

Theorem 5.2 ([144]). *Let G be a graph, then $\alpha_k(G) \leq \vartheta_k(G)$.*

Proof. We sketch the proof by [144] for convenience. Let $X^* \in \{0,1\}^{n\times k}$ be the optimal solution to problem (5.1). That is, for every $r \in [k]$ the r^{th} column of X^* is the incidence vector of the stable set colored in color r. Let $\hat{X}$ be the matrix whose columns are the columns of X^* normalized to one. By construction $\hat{X}^\mathsf{T}\hat{X} = I_k$. Therefore, by Theorem 5.1 for any matrix A feasible to problem (5.2),

$$\alpha_k(G) = \langle A, \hat{X}\hat{X}^\mathsf{T} \rangle \leq \max_X \left\{ \langle A, XX^\mathsf{T} \rangle \ \text{ s.t. } \ X^\mathsf{T}X = I_k \right\} = \sum_{i=1}^{k} \lambda_i(A).$$

Hence $\alpha_k(G) \leq \vartheta_k(G)$. □

For $k = 1$, $\vartheta_k(G)$ is the eigenvalue formulation of the ϑ-number by Lovász [129]. In the original paper by Narasimhan and Manber [153], $\vartheta_k(G)$ is introduced for the clique number of G, which implies that $\vartheta_k(G)$ is defined for the complement of G. We work with the stability number $\alpha_k(G)$ and therefore define $\vartheta_k(G)$ for G. Our notation coincides with notation in Mohar and Poljak [144].

It is known that problem (5.3) can be formulated as an SDP. Next, we present several SDP formulations of (5.3). For this purpose we again use Fan's Theorem 5.1. For a given symmetric matrix A, consider the following pair of primal and dual SDPs:

$$p^* = \max_{Z\in\mathbb{S}^n} \left\{ \langle A, Z \rangle \ \text{ s.t. } \langle I, Z \rangle = k, \ Z \succeq 0, \ I - Z \succeq 0 \right\} \tag{5.4}$$

$$d^* = \min_{Y\in\mathbb{S}^n,\mu} \left\{ \langle I, Y \rangle + \mu k \ \text{ s.t. } -A + \mu I + Y \succeq 0, \ Y \succeq 0 \right\}. \tag{5.5}$$

The theorem below shows that the optimal values of both problems (5.4) and (5.5) are equal to the sum of k largest eigenvalues of a given matrix A.

Theorem 5.3 ([166]). *Let A be a symmetric matrix with eigenvalues $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$. Then $\sum_{i=1}^{k} \lambda_i(A) = p^*$, where p^* is defined in* (5.4).

The proof of Theorem 5.3 uses the properties of projection matrices and can be found in the work of Overton and Womersley [166].

Corollary 5.4. *Define p^* and d^* as in* (5.4) *and* (5.5)*, respectively. Then $p^* = d^*$, and both problems attain the optimal value.*

Proof. Define Λ as the diagonal matrix such that $\Lambda_{ii} = \lambda_i(A)$. Let $U\Lambda U^\intercal$ be the spectral decomposition of A and define the following diagonal matrix:

$$M_{ij} = \begin{cases} \lambda_i(A) - \lambda_k(A), & i = j \leq k \\ 0, & \text{otherwise.} \end{cases}$$

We show that $(Y, \mu) := \big(UMU^\intercal, \lambda_k(A)\big)$ is feasible for problem (5.5), with the objective value equal to p^*. Then $p^* = d^*$ follows from the fact that the problems are dual, using weak duality. First, $Y \succeq 0$ as Y has the same eigenvalues as M and $M \succeq 0$. Next, $-A + \mu I + Y = -U(\Lambda - M)U^\intercal + \lambda_k(A)I$. The matrix $-U(\Lambda - M)U^\intercal$ has the same eigenvectors as A, and its minimum eigenvalue is $-\lambda_k(A)$. Therefore $-A + \mu I + Y$ has the same eigenvectors as A and its minimum eigenvalue is zero. By Theorem 5.3 we obtain

$$d^* \leq \langle I, Y\rangle + \mu k = \sum_{i=1}^{k} (\lambda_i(A) - \lambda_k(A)) + k\lambda_k(A) = \sum_{i=1}^{k} \lambda_i(A) = p^*.$$

Hence strong duality holds and (Y, μ) is the optimal solution to problem (5.5). At the same time, let U_k be the matrix that consists of the first k columns of U. Since the i^{th} column is the i^{th} unitary eigenvector of A, $Z = U_k U_k^\intercal$ is the optimal solution to problem (5.4). Therefore the optimal value is attained in both problems. □

The first SDP reformulation of (5.2) follows from Theorem 5.3.

Corollary 5.5.

$$\begin{aligned} \vartheta_k(G) = \min_{A, Y \in \mathbb{S}^n, \mu} \quad & \langle I, Y\rangle + \mu k \\ \text{s.t.} \quad & -A + \mu I + Y \succeq 0, \\ & Y \succeq 0,\ A_{ij} = 1 \textit{ for } \{ij\} \notin E \textit{ or } i = j. \end{aligned} \tag{5.6}$$

Note that Y in problem (5.5), and thus in problem (5.6), is the dual variable corresponding to constraint $I - Z \succeq 0$ in (5.4). For $k = 1$ constraint $I - Z \succeq 0$ in problem (5.4) becomes redundant since Z is positive-semidefinite and its eigenvalues sum to one. Therefore when $k = 1$, Y in problem (5.6) can be set to zero, which leads to one of the standard formulations of the ϑ-number by Lovász [129]:

$$\begin{aligned} \vartheta(G) = \min_{A \in \mathbb{S}^n, \mu} \quad & \mu \\ \text{s.t.} \quad & -A + \mu I \succeq 0,\ A_{ij} = 1 \text{ for } \{ij\} \notin E \text{ or } i = j. \end{aligned} \tag{5.7}$$

From here on we use the following notation: for all $i \in [n]$, $\mathcal{E}^{ii} = e_i e_i^\intercal$, and for $i, j \in [n]$ such that $i \neq j$, $\mathcal{E}^{ij} = e_i e_j^\intercal + e_j e_i^\intercal$, where e_i is the unit vector with one in the i^{th} entry.

Corollary 5.6. *The optimal values of the following problems are equal to* $\vartheta_k(G)$

$$\begin{aligned} \min_{Y \in \mathbb{S}^n, \{x_{ij}\}_{\{ij\} \in E}, \mu} \quad & \langle I, Y \rangle + \mu k \\ \text{s. t.} \quad & \sum_{\{ij\} \in E} \mathcal{E}^{ij} x_{ij} - J + \mu I + Y \succeq 0, \; Y \succeq 0. \end{aligned} \tag{5.8}$$

$$\begin{aligned} \max_{Z \in \mathbb{S}^n} \quad & \langle J, Z \rangle \\ \text{s. t.} \quad & Z_{ij} = 0 \; for \; \{ij\} \in E, \\ & \langle I, Z \rangle = k, \\ & Z \succeq 0, \; I - Z \succeq 0, \end{aligned} \tag{5.9}$$

and problem (5.9) *is strictly feasible.*

Proof. Reformulation (5.8) follows immediately from the definition of A in problem (5.6) and the fact that $A - J = \sum_{\{ij\} \in E} \mathcal{E}^{ij} x_{ij}$ for some $x_{ij} \in \mathbb{R}$ for all $\{ij\} \in E$. Problem (5.9) is the dual of problem (5.8). Moreover, $Z = \frac{k}{n} I$ is strictly feasible for problem (5.9). This implies that strong duality holds for problems (5.8) and (5.9). □

Alternatively, SDP relaxation (5.9) can be obtained directly from problem (5.1). As before, let $X \in \{0,1\}^{n \times k}$ be the optimal solution to problem (5.1), and let $\hat{X}$ be the matrix whose columns are the columns of X normalized to one. Then the matrix $Z = \hat{X}\hat{X}^\intercal$ is feasible for problem (5.9) with the objective value $\alpha_k(G)$. Clearly, $Z \succeq 0$. Moreover, the objective value is $\alpha_k(G)$ since $\left(\hat{X}\hat{X}^\intercal\right)_{ij} = \frac{1}{c_r}$ if vertices i and j are colored with color r, and c_r is the total number of vertices colored in color r; $\left(\hat{X}\hat{X}^\intercal\right)_{ij} = 0$ otherwise. The first constraint is satisfied by construction of $\hat{X}$. The second and the last constraints in (5.9) are satisfied since the columns of $\hat{X}$ are normalized to one, $Z = \hat{X}\hat{X}^\intercal$ has k eigenvalues equal to one, and $n - k$ eigenvalues equal to zero.

For $k = 1$ constraint $I - Z \succeq 0$ in problem (5.9) becomes redundant since Z is positive-semidefinite and its eigenvalues sum to one. In this case problem (5.9) reduces to another formulation of the ϑ-number by Lovász [129], i.e.,

$$\begin{aligned} \vartheta(G) = \max_{Z \in \mathbb{S}^n} \; & \langle J, Z \rangle \\ \text{s. t. } & Z_{ij} = 0 \text{ for } \{ij\} \in E, \\ & \langle I, Z \rangle = 1, \\ & Z \succeq 0. \end{aligned} \tag{5.10}$$

Also, the dual of problem (5.10) is a classical formulation of the ϑ-number by Lovász [129], which corresponds to problem (5.8) with $Y = 0$.

5.3.2 Strengthening the generalized ϑ-number

Here we introduce a natural strengthening of the SDP relaxation (5.9). Note that all entries of the optimal solution to the integer programming problem (5.1) are non-negative. Therefore, we can add the non-negativity constraints to the matrix variable in (5.9) to strengthen the relaxation. This leads us to the following SDP relaxation:

$$\begin{aligned} \vartheta'_k(G) = \max_{Z \in \mathbb{S}^n} \ & \langle J, Z \rangle \\ \text{s. t.} \ & Z_{ij} = 0 \text{ for } \{ij\} \in E \\ & \langle I, Z \rangle = k \\ & Z \succeq 0, \ I - Z \succeq 0, \\ & Z \geq 0. \end{aligned} \tag{5.11}$$

Note that for $k = 1$, ϑ'_k equals the ϑ' upper bound on $\alpha(G)$ by Schrijver [202].

5.4 Copositive reformulation

We can obtain a copositive reformulation of problem (5.1) using the approach by Peña et al. [175]. One could also use the approach by Burer [28]. However, to satisfy the conditions in [28], one has to deal with the linear IP formulation of the MkCS (5.1) and add a number of slack variables and constraints to turn it into an equality constrained problem over a bounded feasibility set. This makes the final copositive formulation larger and motivates us to choose an alternative approach. The approach by Peña et al. [175] that we use applies to equality constrained polynomial optimization problems. Therefore we first write the MkCS in this form. We begin with replacing linear inequality "edge constraints" in problem (5.1) by non-linear equality constraints:

$$\alpha_k(G) = \max_{X \in \{0,1\}^{n \times k}} \sum_{i \in [n], r \in [k]} X_{ir} \tag{5.12}$$

$$\text{s. t.} \quad X_{ir} X_{jr} = 0, \quad \text{for all } \{ij\} \in E, r \in [k] \tag{5.13}$$

$$\sum_{r \in [k]} X_{ir} \leq 1, \quad \text{for all } i \in [n]. \tag{5.14}$$

Now, we add a binary slack variable to each inequality constraint in (5.12) to obtain an equality constrained problem.

$$\begin{aligned}\alpha_k(G) = \max_{X\in\{0,1\}^{n\times(k+1)}} \quad & \sum_{i\in[n],r\in[k]} X_{ir} \\ \text{s.t.} \quad & X_{i,r}X_{j,r} = 0, \quad \text{for all } \{ij\}\in E, r\in[k] \\ & \sum_{r\in[k+1]} X_{ir} = 1, \quad \text{for all } i\in[n].\end{aligned} \tag{5.15}$$

Next, we rewrite the binary constraints to obtain a polynomial optimization formulation of the MkCS. Then we square the second constraint and the objective to satisfy the conditions in (5.16).

$$\begin{aligned}\alpha_k(G) = \max_{X\in\mathbb{R}^{n\times(k+1)}} \quad & \sum_{i\in[n],r\in[k]} X_{ir}^2 \\ \text{s.t.} \quad & \Big(\sum_{r\in[k+1]} X_{ir} - 1\Big)^2 = 0, \quad \text{for all } i\in[n] \\ & X_{i,r}X_{j,r} = 0, \quad \text{for all } \{ij\}\in E,\ r\in[k] \\ & X_{ir}X_{il} = 0, \quad \text{for all } i\in[n],\ r,l\in[k],\ l\neq r \\ & X_{ir}(1-X_{ir}) = 0, \quad \text{for all } i\in[n],\ r\in[k] \\ & X_{ir}\geq 0, \quad \text{for all } i\in[n],\ r\in[k].\end{aligned} \tag{5.16}$$

The third constraint is redundant, we add it to obtain a more intuitive copositive reformulation. All polynomials and the order of the constraints in problem (5.16) are such that this problem satisfies the conditions of Theorem 4 by Peña et al. [175]. This allows us to write (5.16) as a copositive program. We refer the interested reader to [175] for the details of the reformulation since the procedure is straightforward, but requires much new notation that we do not use later in the chapter.

To write the copositive program, we need some new notation as well, and this notation is frequently used in the sequel. Assume that $X\in\{0,1\}^{n\times(k+1)}$ is the matrix with one in the entry (i,r) if vertex i is colored with color r and zero otherwise, where color $k{+}1$ represents the uncolored vertices. That is, X is feasible for problem (5.15), and the first k columns of X form a feasible solution to any of the formulations of the MkCS we use, such as problem (5.16). Let $\operatorname{vec}(\cdot)$ be the operator that produces a vector from a matrix by stacking its columns onto each other. Consider the following matrix:

$$\begin{bmatrix}1 & y^T\\ y & Y\end{bmatrix} = \begin{bmatrix}1\\ \operatorname{vec}(X)\end{bmatrix}\begin{bmatrix}1 & \operatorname{vec}(X)^\intercal\end{bmatrix} = \begin{bmatrix}1 & \operatorname{vec}(X)^\intercal\\ \operatorname{vec}(X) & \operatorname{vec}(X)\operatorname{vec}(X)^\intercal\end{bmatrix}. \tag{5.17}$$

This matrix has a certain block structure. Y consists of $(k+1)^2$ blocks of the size $n{\times}n$. We denote by Y^{rl} the $n{\times}n$ block of Y located in position $(r,l)\in[k{+}1]{\times}[k{+}1]$.

From here on we denote by e the vector of all ones, use the subscripts i, j to indicate vertices and use superscripts r, l to indicate colors.

Using the notation above and Theorem 4 by Peña et al. [175] for (5.16), we obtain the following copositive program for $\alpha_k(G)$:

$$\begin{aligned}
\alpha_k(G) = \max_{Y \in \mathbb{S}^{n(k+1)}, y \in \mathbb{R}^{n(k+1)}} \quad & \sum_{r\in[k]} \sum_{i\in[n]} Y_{ii}^{rr} \\
\text{s. t.} \quad & \sum_{r=1}^{k} Y_{ii}^{rr} + 2 \sum_{1\le r\neq l\le k} Y_{ii}^{rl} - 2\sum_{r=1}^{k} y_{(r-1)n+i} + 1 = 0, \\
& \qquad \text{for all } i \in [n] \\
& Y_{ij}^{rr} = 0, \quad \text{for all } \{ij\} \in E,\ r \in [k] \\
& Y_{ii}^{rl} = 0, \quad \text{for all } i \in [n],\ r, l \in [k{+}1],\ r \neq l \\
& Y_{ii}^{rr} = y_{(r-1)n+i}, \quad \text{for all } i \in [n],\ r \in [k{+}1] \\
& \begin{bmatrix} 1 & y^T \\ y & Y \end{bmatrix} \in \mathcal{CP}^{n(k+1)}.
\end{aligned}$$

$$\begin{aligned}
= \max_{Y \in \mathbb{S}^{n(k+1)}} \quad & \sum_{r\in[k]} \sum_{i\in[n]} Y_{ii}^{rr} \\
\text{s. t.} \quad & \sum_{r\in[k+1]} \operatorname{diag} Y^{rr} = e \\
& Y_{ij}^{rr} = 0, \quad \text{for all } \{ij\} \in E,\ r \in [k] \\
& Y_{ii}^{rl} = 0, \quad \text{for all } i \in [n],\ r, l \in [k{+}1],\ r \neq l \\
& \begin{bmatrix} 1 & (\operatorname{diag} Y)^T \\ \operatorname{diag} Y & Y \end{bmatrix} \in \mathcal{CP}^{n(k+1)}.
\end{aligned} \tag{5.18}$$

Testing whether a given matrix is completely positive is NP-hard [50], so we do not attempt to solve problem (5.18) directly. However, the complete positivity condition in (5.18) can be relaxed to the SDP condition. This approach provides an SDP upper bound on $\alpha_k(G)$ different from the generalized ϑ- and ϑ'-numbers.

5.5 Matrix and vector lifting SDP relaxations

Next, we propose several SDP relaxations that differ from the generalized ϑ- and ϑ'-numbers considered earlier. We present our relaxations in the order from the strongest to the weakest one. We also show which constraints have to be added or removed to make each pair of relaxations equivalent, meaning that from every feasible solution to one problem one can construct a feasible solution to the other with the same objective value. The first bound is a relaxation of copositive problem (5.18). Clearly, every completely positive matrix is doubly non-negative, that is, PSD and non-negative. This fact leads to the following upper bound on $\alpha_k(G)$:

Vector lifting SDP relaxation of problem (5.15)

$$\begin{aligned} \theta_k^1(G) = \max_{Y \in \mathbb{S}^{n(k+1)}} \quad & \sum_{r\in[k]} \sum_{i\in[n]} Y_{ii}^{rr} \\ \text{s.t.} \quad & \sum_{r\in[k+1]} \operatorname{diag} Y^{rr} = e \\ & Y_{ij}^{rr} = 0, \quad \text{for all } \{ij\} \in E,\ r \in [k] \\ & Y_{ii}^{rl} = 0, \quad \text{for all } i \in [n],\ r,l \in [k+1],\ r \neq l \\ & Y \geq 0,\ \begin{bmatrix} 1 & (\operatorname{diag} Y)^T \\ \operatorname{diag} Y & Y \end{bmatrix} \succeq 0. \end{aligned} \tag{5.19}$$

The name of the problem above is motivated by the fact that this problem can be obtained using the classical vector lifting SDP relaxation of problem (5.15) (see e.g., [224] for this approach in graph partitioning). A vector lifting relaxation of (5.15) is based on the same matrix (5.17) that we use for the copositive reformulation (5.18). As a result, every feasible solution to IP formulation (5.15) provides a feasible solution to problem (5.19). Hence $\theta_k^1(G)$ is an upper bound on $\alpha_k(G)$.

In relaxation (5.19), Y is of size $n(k+1)\times n(k+1)$, which is computationally demanding. Therefore we consider a smaller vector lifting SDP relaxation, which is derived from problem (5.12) using the same construction as in (5.17), where X is a feasible solution to problem (5.12).

Vector lifting SDP relaxation of problem (5.12)

$$\theta_k^2(G) = \max_{Y \in \mathbb{S}^{nk}} \sum_{r\in[k]} \sum_{i\in[n]} Y_{ii}^{rr} \tag{5.20}$$

$$\text{s.t.} \quad Y_{ij}^{rr} = 0, \quad \text{for all } \{ij\} \in E,\ r \in [k] \tag{5.21}$$

$$Y_{ii}^{rl} = 0, \quad \text{for all } i \in [n],\ r,l \in [k],\ r \neq l \tag{5.22}$$

$$Y \geq 0,\ \begin{bmatrix} 1 & (\operatorname{diag} Y)^T \\ \operatorname{diag} Y & Y \end{bmatrix} \succeq 0.$$

In problem (5.20) we do not use several types of constraints which seem reasonable for the MkCS but are redundant. First, the constraint $\sum_{r\in[k]} \operatorname{diag} Y^{rr} \leq e$ arises naturally from the last constraint in problem (5.12) but is redundant.

Lemma 5.7. *The constraint $\sum_{r\in[k]} \operatorname{diag} Y^{rr} \leq e$ is redundant for problem* (5.20).

Proof. Let Y be feasible for problem (5.20). For any $i \in [n]$ consider the principal submatrix $\hat{Y}$ of Y of the size $(k{+}1) \times (k{+}1)$ that has the diagonal $[1, y^\top] := [1, Y_{ii}^{11}, ..., Y_{ii}^{kk}]$. Since $Y_{ii}^{rl} = 0$ for $r, l \in [k]$ such that $r \neq l$, the submatrix has the form

$$\hat{Y} = \begin{bmatrix} 1 & y^\top \\ y & \text{diag } y \end{bmatrix},$$

where diag y is the diagonal matrix with y on the diagonal. Since $\hat{Y} \succeq 0$, we have

$$e^\top \text{diag } ye - e^\top yy^\top e = \sum_{r=1}^{k} y_r - (\sum_{r=1}^{k} y_r)^2 = \sum_{r=1}^{k} Y_{ii}^{rr} - (\sum_{r=1}^{k} Y_{ii}^{rr})^2 \geq 0,$$

which implies $\sum_{r\in[k]} Y_{ii}^{rr} \leq 1$. □

Lemma 5.7 follows immediately using the invariance of problem (5.20) under the color permutations. Since the problem is convex and invariant, it is enough to consider only the solutions Y invariant under the color permutations. In Section 5.6 we reformulate problem (5.20) using this type of solutions. The resulting reformulation (5.37) makes the redundance of the constraint $\sum_{r\in[k]} \text{diag } Y^{rr} \leq e$ evident.

Next, the clique constraints $\sum_{i\in C} \text{diag } Y_{ii}^{rr} \leq 1$, where $C \subseteq [n]$ denotes a set of indices of vertices in a clique, are redundant for problem (5.20) for any C. The reason is the equivalence of problem (5.20) to the Lovász ϑ-problem of the Cartesian product of G and a complete graph on k vertices, see Section 5.6.3 for more details. Clique constrains are well known to be redundant for the Lovász ϑ-problem, which is proven, for instance, in Chapter 9 of [135]. Notice that since the clique constraints are redundant for (5.20), they are also redundant for (5.19).

Relaxation (5.20) is less computationally demanding than (5.19) since Y is of the size $nk{\times}nk$. However, (5.20) is in general weaker than the SDP relaxation (5.19), see, for example, [28, 187]. Next, we show that this is the case for our problems and that the relaxations become equivalent after adding several natural inequality constraints to problem (5.20). For $i, j \in [n]$ and $l, r \in [k]$, we add

$$1 - \sum_{r\in[k]} Y_{ii}^{rr} - \sum_{r\in[k]} Y_{jj}^{rr} + \sum_{r\in[k]} \sum_{l\in[k]} Y_{ij}^{rl} \geq 0, \text{ for all } i > j \tag{5.23}$$

$$Y_{ii}^{ll} - \sum_{r\in[k]} Y_{ij}^{rl} \geq 0, \quad \text{for all } i \neq j, \ l. \tag{5.24}$$

These inequalities are based on the reformulation-linearization technique by Sherali and Adams [209]. In particular, inequalities (5.23) are linearizations of the products of pairs of constraints (5.14). Inequalities (5.24) represent multiplication of element-wise non-negativity constraint on X with each individual constraint in (5.14). These inequalities were also used by Rendl and Sotirov [187] in a similar way we use them in this chapter.

Theorem 5.8. *Problem* (5.20) *with additional constraints* (5.23), (5.24) *is equivalent to problem* (5.19).

Proof. First, let Y be feasible for problem (5.19). Then $Y(1{:}kn, 1{:}kn)$ is feasible for problem (5.20) by construction of both problems. Now, let Z be feasible for problem (5.20) such that it also satisfies (5.23), (5.24), and denote $z = \operatorname{diag} Z$. For ease of presentation we write constraints (5.23) and (5.24) as matrix inequalities using the following transformation matrix:

$$M_{tr} = \begin{bmatrix} 1 & -(m^1)^\intercal \\ \dots & \dots \\ 1 & -(m^n)^\intercal \end{bmatrix}, \; m^i(j) = \begin{cases} 1, & j \in \{i, n+i, \dots, n(k-1)+i\} \\ 0, & \text{otherwise.} \end{cases} \tag{5.25}$$

Then inequalities (5.23) can be written as follows:

$$M_{tr} \begin{bmatrix} 1 & z^\intercal \\ z & Z \end{bmatrix} M_{tr}^\intercal \geq 0. \tag{5.26}$$

Inequalities (5.24) can be written as follows:

$$M_{tr} \begin{bmatrix} z^\intercal \\ Z \end{bmatrix} \geq 0. \tag{5.27}$$

We claim that

$$Y = \begin{bmatrix} Z & [\, z \; Z \,] M_{tr}^\intercal \\ M_{tr} \left[\begin{smallmatrix} z^\intercal \\ Z \end{smallmatrix} \right] & M_{tr} \left[\begin{smallmatrix} 1 & z^\intercal \\ z & Z \end{smallmatrix} \right] M_{tr}^\intercal \end{bmatrix} \tag{5.28}$$

is feasible for problem (5.19). First, it follows from (5.26), (5.27) and the feasibility of Z for (5.20) that $Y \geq 0$. Now, define

$$\hat{z} = e - \sum_{r \in [k]} \operatorname{diag} Z^{rr} = M_{tr} \begin{bmatrix} 1 \\ z \end{bmatrix}. \tag{5.29}$$

For each $i \in [n]$, the i^{th} diagonal entry of $M_{tr} \left[\begin{smallmatrix} 1 & z^\intercal \\ z & Z \end{smallmatrix} \right] M_{tr}^\intercal$ equals

$$1 - \sum_{r \in [k]} Z_{ii}^{rr} - \sum_{r \in [k]} Z_{ii}^{rr} + \sum_{r \in [k]} \sum_{l \in [k]} \operatorname{diag} Z_{ii}^{rl} = 1 - \sum_{r \in [k]} Z_{ii}^{rr} = \hat{z}_i,$$

where we use (5.22) for the first equality and (5.29) for the second one. Hence $\sum_{r \in [k+1]} \operatorname{diag} Y^{rr} = e$. Next, we prove that for each $r \in [k]$, the block of $M_{tr} \left[\begin{smallmatrix} z^\intercal \\ Z \end{smallmatrix} \right]$ that corresponds to $Y^{(k+1)r}$ has a zero diagonal. For $r \in [k]$ and $i \in [n]$,

$$Y_{ii}^{(k+1)r} = Y_{ii}^{rr} - \sum_{l \in [k]} Y_{ii}^{rl} = 0.$$

Finally,

$$\begin{bmatrix} 1 & (\operatorname{diag} Y)^T \\ \operatorname{diag} Y & Y \end{bmatrix} = \begin{bmatrix} 1 & z^\intercal & \hat{z}^\intercal \\ z & Z & \left[\begin{smallmatrix} z & Z \end{smallmatrix}\right] M_{tr}^\intercal \\ \hat{z} & M_{tr} \left[\begin{smallmatrix} z^\intercal \\ Z \end{smallmatrix}\right] & M_{tr} \left[\begin{smallmatrix} 1 & z^\intercal \\ z & Z \end{smallmatrix}\right] M_{tr}^\intercal \end{bmatrix} = \begin{bmatrix} I \\ M_{tr} \end{bmatrix} \begin{bmatrix} 1 & z^\intercal \\ z & Z \end{bmatrix} \begin{bmatrix} I & M_{tr}^\intercal \end{bmatrix} \succeq 0$$

since $\left[\begin{smallmatrix} 1 & z^\intercal \\ z & Z \end{smallmatrix}\right] \succeq 0$. □

Notice that the additional vector in the first row and column in both problems (5.19) and (5.20) could be omitted to obtain smaller SDP relaxations, however, these relaxations are trivial with the optimal value at least n. One could use, for example, $Y = \frac{1}{k+1} I$ and $Z = \frac{1}{k} I$ as solutions to the first and the second relaxations, respectively.

Problem (5.20) is still large, especially when k grows. Next, we consider the so-called matrix lifting relaxation to reduce the size of the matrix variables in the relaxations even more, see, for example [52] for more information. The idea behind this type of relaxations is as follows. As before, let X be a solution to problem (5.12) and consider

$$Y = \begin{bmatrix} I_k \\ X \end{bmatrix} \begin{bmatrix} I_k & X^\intercal \end{bmatrix} = \begin{bmatrix} I_k & X^\intercal \\ X & XX^\intercal \end{bmatrix}.$$

Linearizing the block $XX^\intercal$, we obtain the matrix lifting relaxation of problem (5.12):

Matrix lifting SDP relaxation of problem (5.12)

$$\theta_k^3(G) = \max_{Z \in \mathbb{S}^n, X \in \mathbb{R}^{n \times k}} \langle I, Z \rangle \tag{5.30}$$

$$\begin{aligned} \text{s.t.} \quad & Z_{ij} = 0 \text{ for } \{ij\} \in E \\ & Z_{ii} \le 1 \text{ for } i \in [n] \\ & Z_{ii} = \sum_{r \in [k]} X_{ir} \text{ for } i \in [n] \qquad (5.31) \\ & Z \ge 0,\ X \ge 0 \\ & \begin{bmatrix} I_k & X^\intercal \\ X & Z \end{bmatrix} \succeq 0. \end{aligned}$$

The new relaxation is not stronger than the previous ones.

Theorem 5.9. *Problem* (5.30) *is equivalent to* (5.20) *without constraint* (5.22).

Proof. First, let Y be feasible for problem (5.20) without constraint (5.22). We show that $Z := \sum_{r \in [k]} Y^{rr}$ and $X := [\operatorname{diag} Y^{11}, \ldots, \operatorname{diag} Y^{rr}]$ are feasible for problem (5.30). Constraints of problem (5.20) for Y and Lemma 5.7 imply that all but the last constraints of (5.30) hold for (Z, X). To prove that the last constraint holds too, we use the Schur complement. First, for every $r \in [k]$ and a principal submatrix $\begin{bmatrix} 1 & (\text{diag } Y^{rr})^\intercal \\ \text{diag } Y^{rr} & Y^{rr} \end{bmatrix}$ of $\begin{bmatrix} 1 & (\text{diag } Y)^\intercal \\ \text{diag } Y & Y \end{bmatrix}$ we have that

$$Y^{rr} - \text{diag } Y^{rr}(\text{diag } Y^{rr})^\intercal \succeq 0,$$

from where it follows

$$\begin{aligned} Z - XX^\intercal &= \sum_{r\in[k]} Y^{rr} - \sum_{r\in[k]} \text{diag } Y^{rr}(\text{diag } Y^{rr})^\intercal \\ &= \sum_{r\in[k]} (Y^{rr} - (\text{diag } Y^{rr}(\text{diag } Y^{rr})^\intercal) \succeq 0. \end{aligned}$$

Now, let (Z, X) be feasible for problem (5.30). Denote by $X(:, r)$ the r^{th} column of X. We show that $Y^{rr} := \frac{1}{k}Z$ and $Y^{rl} := \frac{1}{k}XX^\intercal$ for all $r, l \in [k]$, $r \neq l$ are feasible for problem (5.20) without constraint (5.22). Constraints of problem (5.30) for (Z, X) imply that all constraints of problem (5.20) hold for Y, except for the SDP constraint and (5.22). Next, we prove that the SDP constraint of problem (5.20) holds for Y. First we use the Schur complement and the last constraint of problem (5.30) to obtain:

$$0 \preceq Z - XX^\intercal = kY^{rr} - XX^\intercal. \tag{5.32}$$

Next, from constraint (5.31) it follows

$$\text{diag } Y^{rr} = \frac{1}{k}\text{diag } Z = \frac{1}{k} \sum_{r\in[k]} X(:, r). \tag{5.33}$$

Finally, consider any $w \in \mathbb{R}^{nk}$ and for $r \in [k]$ denote $w^r = w_{(r-1)n+1:rn}$. Denote $y = \text{diag } Y$, then

$$\begin{aligned} w^\intercal(Y - yy^\intercal)w &= \sum_{r,l\in[k]} (w^r)^\intercal Y^{rl} w^l - \left(\sum_{r=1}^{k} (w^r)^\intercal \text{diag } Y^{rr}\right)^2 \\ &\overset{(5.32)}{\geq} \frac{1}{k} \sum_{r,l\in[k]} (w^r)^\intercal XX^\intercal w^l - \left(\sum_{r=1}^{k} (w^r)^\intercal \text{diag } Y^{rr}\right)^2 \\ &\overset{(5.33)}{=} \frac{1}{k} \sum_{r,l\in[k]} (w^r)^\intercal \sum_{m=1}^{k} X(:, m)X(:, m)^\intercal w^l - \frac{1}{k^2}\left(\sum_{r=1}^{k} (w^r)^\intercal \sum_{m=1}^{k} X(:, m)\right)^2 \\ &= \frac{1}{k} \sum_{m=1}^{k} \left(\sum_{r=1}^{k} (w^r)^\intercal X(:, m)\right)^2 - \frac{1}{k^2}\left(\sum_{m=1}^{k}\sum_{r=1}^{k} (w^r)^\intercal X(:, m)\right)^2 \\ &= \frac{1}{k^2}\left[\|e\|^2 \sum_{m=1}^{k} \left(\sum_{r=1}^{k} (w^r)^\intercal X(:, m)\right)^2 - \left(\sum_{m=1}^{k} 1 \sum_{r=1}^{k} (w^r)^\intercal X(:, m)\right)^2\right] \geq 0, \end{aligned}$$

where e is the vector of all ones of length k, and the last inequality follows from the Cauchy-Schwarz inequality. Hence the SDP constraint in (5.20) holds for Y. $\square$

Notice that constraints $Y_{ii} \leq 1$ for $i \in [n]$ are redundant when $k=1$. A reasonable question is how to strengthen problem (5.30). At the end of Section 5.6 we consider several types of constraints that could be useful, and now we mention several types of constraints that do not help. First, we could use the fact that for the optimal solutions of problems (5.1) and (5.12), $X^\intercal X = \operatorname{diag} z$ for some $z > 0$. Thus $\operatorname{diag} z - X^\intercal X \succeq 0$ and, by the Schur complement, $\left[\begin{smallmatrix} I_n & X \\ X^\intercal & \operatorname{diag} z \end{smallmatrix}\right] \succeq 0$. However, the latter constraint is redundant: one can always take $z = X^\intercal e$, then $\operatorname{diag} z - X^\intercal X$ is a diagonally dominant matrix, given (5.31) and constraints on Y. In the sequel we show that constraints $X_{ir} + X_{jr} \leq 1$ for $\{ij\} \in E, r \in [k]$ and $k \geq 1$ are redundant too.

Proposition 5.10. *Constraints*

$$X_{ir} + X_{jr} \leq 1 \text{ for } \{ij\} \in E,\ r \in [k] \tag{5.34}$$

are redundant for problem (5.30).

Proof. Let (X, Y) be a feasible solution to problem (5.30).
First let $k=1$, and therefore $m=1$. From (5.31) we have that $Y_{ii} = X_{im}, Y_{jj} = X_{jm}$. Therefore from the last constraint of problem (5.30):

$$\begin{bmatrix} 1 & X_{im} & X_{jm} \\ X_{im} & Y_{ii} & 0 \\ X_{jm} & 0 & Y_{jj} \end{bmatrix} \succeq 0 \Longrightarrow Y_{ii}Y_{jj} - Y_{ii}X_{jm}^2 - Y_{jj}X_{im}^2 = Y_{ii}Y_{jj}(1 - X_{im} - X_{jm}) \geq 0.$$

The inequality above together with non-negativity of Y implies that $X_{im} + X_{jm} \leq 1$.

Now let $k \geq 2$. Denote the average over all column permutations of X by $\bar{X}$. Let $P_{ij} \in \mathbb{S}^{k+n}$ be the matrix that corresponds to the permutation of the i^{th} and the j^{th} entries in a vector in $\mathbb{R}^{k+n}$. Since problem (5.30) is invariant under the permutations of colors, for all $i, j \in [k]$, the matrix $P_{ij} \left[\begin{smallmatrix} I_k & X^\intercal \\ X & Y \end{smallmatrix}\right] P_{ij}^\intercal$ is feasible for problem (5.30). Since the problem is also convex, $(\bar{X}, Y)$ is a feasible solution to it as well. Recall that $Sym(k)$ is the group of permutations on k elements. Then for every $i, j \in [n]$, $r \in [k]$,

$$\bar{X}_{ir} + \bar{X}_{jr} = \tfrac{1}{k!} \sum_{\pi \in Sym(k)} (X_{i\pi(r)} + X_{j\pi(r)}) = \tfrac{1}{k!} \sum_{r=1}^{k} (k-1)!(X_{ir} + X_{jr}) = \tfrac{Y_{ii}+Y_{jj}}{k} \leq 1.$$

Hence there is a feasible solution with the same objective value as for (X, Y) but with constraints (5.34) satisfied. □

Finally, we could try the reformulation-linearization technique, similarly to constraints (5.23) and (5.24). In this case, for all $i \neq j$, $l \in [k]$ by construction of Y and X in problem (5.30), we obtain $(1 - Y_{ii})(1 - Y_{jj}) \geq 0$ and $X_{i,k}^2(1 - Y_{jj}) \geq 0$. However, we cannot linearize these constraints naturally for problem (5.30).

5.6 Symmetry reduction on colors

Next, we exploit the invariance of the MkCS under the permutations of the colors to reduce the sizes the vector and matrix lifting SDP relaxations. We begin with a simple example of using the invariance of the MkCS. Consider a trivial SDP relaxation based on the optimal solution $X \in \{0,1\}^{n\times k}$ to problem (5.12) and $Y = XX^\intercal$.

$$\begin{aligned}\theta_k^4(G) = \max_{Y\in\mathbb{S}^{n\times n}} \ & \langle I, Y\rangle \\ \text{s.t.} \ & Y_{ij} = 0 \text{ for } \{ij\} \in E \\ & Y_{ii} \le 1 \text{ for } i \in [n] \\ & Y \succeq 0.\end{aligned} \tag{5.35}$$

The optimal value of the problem above equals n using, for example, $Y = I$ as a solution. Problem (5.35) can be obtained from problem (5.20) if we omit the first row and column of this problem and do the symmetry reduction on colors, as we describe further in this section. Thus using the symmetry we confirm that the additional row and column are crucial for our vector lifting SDP relaxations.

Now, we show how to reduce the size of the largest SDP constraints of problem (5.20) and problem (5.30) to $(n+1)\times(n+1)$. Vector lifting relaxations are usually not strictly feasible, which hampers numerical stability of the interior point method that we use to solve SDPs. The symmetry-reduced problems are strictly feasible which is an advantage for our solution method. We begin with a lemma related to strict feasibility:

Lemma 5.11. *Let $n \ge 1, k \ge 1$, and let e be the vector of all ones of the size n. Then*

$$M := \begin{bmatrix} k & \frac{1}{(n+1)}e^\intercal \\ \frac{1}{(n+1)}e & \frac{1}{(n+1)}I \end{bmatrix} \succ 0. \tag{5.36}$$

Proof. From $M \succ 0$, using the Schur complement, we have that $M \succ 0$ if and only if $k(n+1)I - ee^\intercal \succ 0$. For any $x \in \mathbb{R}^n \setminus \{0\}$

$$\begin{aligned} x^\intercal(k(n+1)I - ee^\intercal)x &= k(n+1)\|x\|^2 - (x^\intercal e)^2 \\ &= (k + (k-1)n)\|x\|^2 + (n\|x\|^2 - (x^\intercal e)^2) \\ &\ge (k + (k-1)n)\|x\|^2 \\ &> 0. \end{aligned}$$

The latter inequality holds since $k \ge 1$. □

Now, we reduce problem (5.20) with constraints (5.23) and (5.24) using the following result, the proof of which can be found, for instance, in [80].

Theorem 5.12 (Lemma 2.8 in [80]). *Let $Y \in \mathbb{R}^{kn}$ be a block matrix that consists of k^2 blocks of the size $n\times n$. Let Y have a matrix $A \in \mathbb{S}^n$ as its diagonal blocks, and a matrix $B \in \mathbb{S}^n$ as its non-diagonal blocks, i.e.*

$$Y = \underbrace{\begin{bmatrix} A & B & \dots & B \\ B & A & \dots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \dots & A \end{bmatrix}}_{k \text{ blocks}} = I \otimes A + (J - I) \otimes B.$$

Then $Y \succeq 0$ if and only if $A - B \succeq 0$ and $A + (k-1)B \succeq 0$.

Theorem 5.13. *Problem* (5.19) *and problem* (5.20) *with additional constraints* (5.23), (5.24) *are equivalent to the following problem:*

$$\begin{aligned} \theta_k^5(G) = \max_{Z,X\in\mathbb{S}^{n\times n}} \quad & \langle I, Z\rangle && (5.37)\\ \text{s.t.} \quad & Z_{ij} = 0, \quad \text{for } \{ij\} \in E \\ & X_{ii} = 0, \quad \text{for } i \in [n] && (5.38)\\ & Z \geq 0, \ \ X \geq 0 \\ & Z - X \succeq 0 && (5.39)\\ & \begin{bmatrix} 1 & (\text{diag} Z)^\intercal \\ \text{diag} Z & Z + (k-1)X \end{bmatrix} \succeq 0. && (5.40)\\ & 1 - Z_{ii} - Z_{jj} + Z_{ij} + (k-1)X_{ij} \geq 0, \quad \text{for } i,j \in [n], i > j && (5.41)\\ & Z_{ii} - Z_{ij} - (k-1)X_{ij} \geq 0, \quad \text{for } i,j \in [n], i \neq j, && (5.42) \end{aligned}$$

and the latter problem is strictly feasible.

Proof. First, (5.19) and (5.20) with additional constraints (5.23), (5.24) are equivalent by Theorem 5.8. We show that (5.20) with (5.23), (5.24) is equivalent to (5.37). Let Y be a feasible solution to problem (5.20) with additional constraints (5.23), (5.24). If we permute the color labels, we permute the "columns" and "rows" of blocks in Y. For instance, permuting color r and color l results in permuting blocks Y^{pl} and Y^{pr}, for all $p \in [k]$, and then permuting blocks Y^{lp} and Y^{rp}, for all $p \in [k]$.
By construction problem (5.20) with constraints (5.23), (5.24) is convex and invariant under the color permutations. Therefore $\bar{Y}$, the average over all color permutations of Y, is feasible for problem (5.20). By construction, $\bar{Y}$ has the form

$$\bar{Y} = \frac{1}{k} \underbrace{\begin{bmatrix} Z & X & \dots & X \\ X & Z & \dots & X \\ \vdots & \vdots & \ddots & \vdots \\ X & X & \dots & Z \end{bmatrix}}_{k \text{ blocks}} = \tfrac{1}{k} I \otimes Z + \tfrac{1}{k}(J - I) \otimes X. \qquad (5.43)$$

Problem (5.20) with constraints (5.23), (5.24) can be restricted without loss of generality to Y of the form above, which results in problem (5.37). Indeed, the objective and all linear constraints of problem (5.37) are obtained by rewriting the objective and the corresponding linear constraints of problem (5.20) with (5.23), (5.24) for Y defined in (5.43). Now, consider the SDP constraint

$$\begin{bmatrix} 1 & (\text{diag } Y)^T \\ \text{diag } Y & Y \end{bmatrix} \succeq 0,$$

which, by the Schur complement, is equivalent to $Y - \text{diag } Y(\text{diag } Y)^\intercal \succeq 0$. We have

$$\begin{aligned} Y - \text{diag } Y(\text{diag } Y)^\intercal = \tfrac{1}{k} I \otimes (Z - \tfrac{1}{k}\text{diag } Z(\text{diag } Z)^\intercal) \\ + \tfrac{1}{k}(J - I) \otimes (X - \tfrac{1}{k}\text{diag } Z(\text{ diag } Z)^\intercal). \end{aligned}$$

Hence by Theorem 5.12, the SDP constraint holds if and only if

$$Z - X \succeq 0, \; Z + (k-1)X - \text{diag } Z(\text{diag } Z)^\intercal \succeq 0. \tag{5.44}$$

To show strict feasibility of (5.37), we construct a feasible solution (Z, X) to this problem that strictly satisfies all inequalities and SDP constraints. Let $A_{\bar{G}}$ be the adjacency matrix of the complement of G and define M as in (5.36). Then $M \succ 0$ by Lemma 5.11. Since also $\frac{1}{k(n+1)} I \succ 0$, there exists $0 < \varepsilon < \frac{1}{k(n+1)(k+1)}$ such that

$$\tfrac{1}{k} M + \varepsilon \begin{bmatrix} 0 & 0^\intercal \\ 0 & A_{\bar{G}} + (k-1)(J - I) \end{bmatrix} \succ 0 \;\text{ and }\; \tfrac{1}{k(n+1)} I + \varepsilon A_{\bar{G}} - \varepsilon(J - I) \succ 0. \tag{5.45}$$

Define $Z = \frac{1}{k(n+1)} I + \varepsilon A_{\bar{G}}$, $X = \varepsilon(J - I)$. By the choice of ε and (5.45), (Z, X) strictly satisfies all constraints of problem (5.37), except for possibly (5.41). To see that constraints (5.41) are strictly satisfied too, notice that these constraints are valid for all $i, j \in [n], i > j$. Hence they appear in the problem when $n \geq 2$. Therefore, for all $i, j \in [n], i > j$ and the chosen (Z, X), we have

$$1 - Z_{ii} - Z_{jj} + Z_{ij} + (k-1)X_{ij} \geq 1 - \tfrac{2}{k(n+1)} = \tfrac{kn+k-2}{k(n+1)} \geq \tfrac{1}{k(n+1)}.$$

□

Next, we reduce the matrix lifting relaxation (5.30).

Theorem 5.14. *Problem* (5.30) *is equivalent to the following problem:*

$$\begin{aligned} \theta_k^6(G) = \max_{Z \in \mathbb{S}^{n \times n}} \; & \langle I, Z \rangle \\ \text{s.t. } & Z_{ij} = 0 \textit{ for } \{ij\} \in E \\ & Z_{ii} \leq 1 \textit{ for } i \in [n] \\ & Z \geq 0 \\ & \begin{bmatrix} k & \text{diag} Z^\intercal \\ \text{diag} Z & Z \end{bmatrix} \succeq 0, \end{aligned} \tag{5.46}$$

and the latter problem is strictly feasible.

Proof. Let (Z, X) be a feasible solution to problem (5.30). Since problem (5.30) is convex and invariant under the permutations of the colors, $(\bar{X}, Y)$, where $\bar{X}$ is the average over all column permutations of X, is feasible for problem (5.30). By construction, all columns of $\bar{X}$ are equal to each other. Therefore it is enough to consider the solutions (Z, X), such that all the columns of X are equal to each other. Denote a column of X by x, then the problem reduces to

$$\begin{aligned} \max_{Z\in\mathbb{S}^{n\times n}} \quad & \langle I, Z\rangle \\ \text{s.t.} \quad & Z_{ij} = 0 \text{ for } \{ij\} \in E \\ & Z_{ii} \le 1 \text{ for } i \in [n] \\ & Z_{ii} = kx_i \text{ for } i \in [n] \\ & Z \ge 0, \begin{bmatrix} I_k & ex^\intercal \\ xe^\intercal & Z \end{bmatrix} \succeq 0, \end{aligned}$$

where $e \in \mathbb{R}^k$ is the vector of all ones. Use the Schur complement and the third constraint to rewrite the last constraint:

$$Z \succeq 0, \; Z - xe^\intercal ex^\intercal = Z - kxx^\intercal = Z - \tfrac{1}{k}\text{diag}Z(\text{diag}Z)^\intercal \succeq 0.$$

To show strict feasibility, consider M defined in (5.36). $M \succ 0$ by Lemma 5.11. Let $A_{\bar{G}}$ be the adjacency matrix of the complement of G. Then there exists $\varepsilon > 0$ such that $M + \varepsilon\left[\begin{smallmatrix} 0 & 0^\intercal \\ 0 & A_{\bar{G}} \end{smallmatrix}\right] \succ 0$. Therefore matrix $Z = \frac{1}{(n+1)}I + \varepsilon A_{\bar{G}}$ is a strictly feasible solution to problem (5.46) by construction. □

Theorem 5.15. *Problem* (5.46) *is equivalent to problem* (5.37) *without constraints* (5.38)*,* (5.41)*,* (5.42)*.*

Proof. First, let (Z, X) be a feasible solution to (5.37) without constraints (5.38), (5.41), (5.42). We claim that Z is feasible for problem (5.46), and the objective values of the problems evaluated in the corresponding solutions are equal. First, $Z_{ii} \le 1$ follows from the SDP constraint of problem (5.37). Other linear constraints in (5.46) are feasible as they form a subset of the linear constraints of (5.37). Now, consider the SDP constraint in (5.46). From (5.39) and (5.40),

$$Z - \tfrac{1}{k}\text{diag }Z(\text{diag }Z)^\intercal \overset{(5.40)}{\succeq} Z - \tfrac{1}{k}Z - \tfrac{k-1}{k}X = \tfrac{k-1}{k}\Big(Z - X\Big) \overset{(5.39)}{\succeq} 0.$$

Since $\text{diag }Z(\text{diag }Z)^\intercal \succeq 0$, the above implies that $Z \succeq 0$. Therefore the SDP constraint of problem (5.46) follows by the Schur complement.

Now, let Z be a feasible solution to (5.46). We claim that $(Z, \frac{1}{k}\text{diag }Z(\text{diag }Z)^\intercal)$ is a feasible solution to problem (5.37) without constraints (5.38), (5.41), (5.42), and the objective values of the problems evaluated in the corresponding solutions are equal.

The linear constraints of problem (5.37), besides (5.38), (5.41), (5.42), are clearly satisfied by $(Z, \frac{1}{k}\text{diag } Z(\text{diag } Z)^\intercal)$. The SDP constraint of problem (5.46) implies $Z \succeq 0$ and

$$Z - X = Z - \tfrac{1}{k}\text{diag } Z(\text{diag } Z)^\intercal \succeq 0,$$

by the Schur complement. Therefore SDP constraint (5.39) is satisfied. Finally,

$$Z + (k-1)X = Z + \tfrac{(k-1)}{k}\text{diag } Z(\text{diag } Z)^\intercal \succeq 0$$

and

$$Z + (k-1)X - \text{diag } Z(\text{diag } Z)^\intercal = Z + (k-1)X - kX = Z - X \succeq 0,$$

which implies SDP constraint (5.40). □

If we omit constraints (5.41), (5.42), Theorem 5.15 is an analogue of Theorem 5.9 for the corresponding symmetry-reduced problems. Using symmetry, we reduce the vector and matrix lifting relaxations whose size depends on k to smaller problems whose size does not depend on k. At the end, the symmetry-reduced problems have almost the same matrix size as the smaller ϑ-number (5.9) and ϑ'-number (5.11) problems. Namely, problems (5.9) and (5.11) have two SDP constraints of size $n\times n$, the first reduced problem (5.37) has two SDP constraints of sizes $n\times n$ and $(n+1)\times(n+1)$, and the weaker reduced problem (5.46) has one SDP constraint of size $(n+1)\times(n+1)$. Overall, problems (5.9), (5.11) and (5.46) may still be substantially smaller than problem (5.37) since the latter has a large number of linear constraints (5.41), (5.42). Notice that if we do not use these constraints, we obtain the symmetry-reduced analog of problem (5.20).

The final relaxation we consider is the symmetry-reduced version of problem (5.19).

Corollary 5.16. *Problem* (5.19) *and problem* (5.37) *are equivalent to the following problem:*

$$\begin{aligned}
\theta_k^7(G) = \max_{Z,X,A,B\in\mathbb{S}^{n\times n}} \ & \langle I, Z\rangle \\
\text{s.t. } & Z_{ij} = 0 \textit{ for } \{ij\} \in E \\
& X_{ii} = 0, \quad \textit{for } i \in [n] \\
& B_{ii} = 0, \quad \textit{for } i \in [n] \\
& Z_{ii} + A_{ii} = 1 \textit{ for } i \in [n] \\
& Z \geq 0,\ X \geq 0,\ A \geq 0,\ B \geq 0 \\
& Z - X \succeq 0 \\
& \begin{bmatrix} k & \sqrt{k}(\text{diag}Z)^\intercal & (\text{diag}A)^\intercal \\ \sqrt{k}\ \text{diag}Z & Z + (k-1)X & B \\ \text{diag}A & B & A \end{bmatrix} \succeq 0.
\end{aligned} \tag{5.47}$$

Proof. First, problem (5.19) and problem (5.37) are equivalent by Theorem 5.8 and Theorem 5.13. Now, consider a feasible solution Y to problem (5.19). The problem is invariant under permutations of the k colors, which correspond to the first k^2 blocks of Y. Further we use the same approach as in the proof of Theorem 5.13, so the details are omitted. By invariance and convexity of the problem, it is enough to consider the solutions of the form

$$Y = \frac{1}{k}\begin{bmatrix} 1 & (e \otimes \operatorname{diag} Z)^{\mathsf{T}} & (\operatorname{diag} A)^{\mathsf{T}} \\ e \otimes \operatorname{diag} Z & I \otimes Z + (J - I) \otimes X & e \otimes B \\ \operatorname{diag} A & (e \otimes B)^{\mathsf{T}} & A \end{bmatrix}.$$

We obtain problem (5.47) using the Schur complement and Theorem 5.12. □

We close this section with Figure 5.1 that shows the relations among all SDP relaxations in this chapter. Equalities in Figure 5.1 represent equivalence between problems.

Figure 5.1 – Relations among SDP upper bounds on the MkCS. The strongest relaxation with the smallest size of SDP constraints is highlighted in bold.

$$\begin{array}{ccccccccc}
 & & & & & & \vartheta'_k(G) & \overset{\text{construct.}}{\leq} & \vartheta_k(G) \\
 & & & & & & \text{Conj.5.17} \vee\!| & ? & \\
 & & \theta^1_k(G) & \overset{\text{Thm.5.8}}{\leq} & \theta^2_k(G) & \overset{\text{Thm.5.9}}{\leq} & \theta^3_k(G) & \overset{\text{construct.}}{\leq} & \theta^4_k(G) \\
 & & \|\ \text{Thm.5.13} & & & & \|\text{Thm.5.14} & & \\
\alpha_k(G) & \overset{\text{construct.}}{\leq} & \boldsymbol{\theta^5_k(G)} & & \overset{\text{Thm.5.15}}{\leq} & & \theta^6_k(G) & & \\
 & & \|\ \text{Cor.5.16} & & & & & & \\
 & & \theta^7_k(G) & & & & & &
\end{array}$$

The figure shows that the relationship between $\theta^6_k(G)$ (5.46) and $\vartheta'_k(G)$ (5.11) is to be established. However, our numerical experiments in Section 5.7 suggest the following result.

Conjecture 5.17. *The upper bound* $\theta^6_k(G)$ (5.46) *is at least as good as the upper bound* $\vartheta'_k(G)$ (5.11)*.*

5.6.1 Boolean quadric polytope inequalities

Our strongest SDP having the smallest size of SDP constraints is problem (5.37). To further strengthen this problem, one can add inequalities from the boolean quadric polytope (BQP) (see, e.g, Padberg [167]). Namely, let X be a feasible solution to binary problem (5.12) and consider $Y = \text{vec}\, X(\text{vec}\, X)^\intercal$. Then for all $i, j, p \in [nk]$ the following BQP inequalities are valid for Y:

$$0 \le Y_{i,j} \le Y_{i,i} \tag{5.48}$$

$$Y_{i,i} + Y_{j,j} \le 1 + Y_{i,j} \tag{5.49}$$

$$Y_{i,p} + Y_{j,p} \le Y_{p,p} + Y_{i,j} \tag{5.50}$$

$$Y_{i,i} + Y_{j,j} + Y_{p,p} \le Y_{i,j} + Y_{i,p} + Y_{j,p} + 1. \tag{5.51}$$

By construction, the BQP inequalities are valid for problem (5.12), and adding them to problem (5.20) provides a possibly stronger upper bound than problem (5.20) alone.

Doing the symmetry reduction on colors, we consider only those feasible solutions Y to problem (5.20) that can be written as in (5.43). Therefore for all $i, j, p \in [n], i \neq j \neq p$, the following valid inequalities can be added to (5.37):

$$0 \le Z_{i,j} \le Z_{i,i}, \qquad 0 \le X_{i,j} \le Z_{i,i} \tag{5.52}$$

$$Z_{i,i} + Z_{j,j} \le k + Z_{i,j}, \qquad Z_{i,i} + Z_{j,j} \le k + X_{i,j} \tag{5.53}$$

$$\begin{aligned} &X_{i,p} + X_{j,p} \le Z_{p,p} + X_{i,j}, && Z_{i,p} + Z_{j,p} \le Z_{p,p} + Z_{i,j} \\ &X_{i,p} + X_{j,p} \le Z_{p,p} + Z_{i,j}, && X_{i,p} + Z_{j,p} \le Z_{p,p} + X_{i,j} \\ &Z_{i,p} + X_{j,p} \le Z_{p,p} + X_{i,j} && \end{aligned} \tag{5.54}$$

$$\begin{aligned} &Z_{i,i} + Z_{j,j} + Z_{p,p} \le X_{i,j} + X_{i,p} + X_{j,p} + k \\ &Z_{i,i} + Z_{j,j} + Z_{p,p} \le Z_{i,j} + Z_{i,p} + Z_{j,p} + k \\ &Z_{i,i} + Z_{j,j} + Z_{p,p} \le Z_{i,j} + X_{i,p} + X_{j,p} + k \\ &Z_{i,i} + Z_{j,j} + Z_{p,p} \le X_{i,j} + X_{i,p} + Z_{j,p} + k, \\ &Z_{i,i} + Z_{j,j} + Z_{p,p} \le X_{i,j} + Z_{i,p} + X_{j,p} + k. \end{aligned} \tag{5.55}$$

Inequalities (5.52) correspond to inequalities (5.48), inequalities (5.53) correspond to (5.49), and so on.

Some of the BQP constraints (5.52)– (5.55) are redundant for problem (5.37). First, inequalities (5.52) follow from inequalities (5.42). Also, inequalities (5.53) are redundant since the 2-clique constraints are redundant for problems (5.19),(5.20). As a result, in numerical experiments we use inequalities of the type (5.54) and (5.55). Notice that the first inequality in each of the two sets (5.54) and (5.55) is only valid for $k \ge 3$.

Next, one could also look at the triangle inequalities for Y in (5.17): $Y_{ij}+Y_{jp}-Y_{ip} \le 1$ for any $i, j, p \in [n]$. These inequalities follow from (5.50), the fact that $Y_{ii} \le 1$, and the non-negativity of Y. Therefore we do not consider them in our numerical experiments.

5.6.2 Symmetry reductions for other partition problems

Notice that a k-colorable subgraph of a graph corresponds to a partition of this graph's vertices into $k+1$ subsets, i.e., k stable sets and the rest of the vertices. Therefore we could consider the k-colorable subgraph problem as a partition problem. Some other partition problems are also invariant under the permutations of the subsets, e.g., the k-equipartition problem (see, e.g., [102, 186, 224] for the problem formulation) and the max-k-cut problem (see, e.g., [186, 212] for the problem formulation). Therefore one could apply the symmetry reduction used in this section to those problems.

Interestingly, in contrast to MkCS, the vector and matrix lifting relaxations are equivalent for the max-k-cut and k-equipartition problems. de Klerk et al. [36] have proven the equivalence for the max-k-cut, and Sotirov [211] has proven the equivalence for the k-equipartition. In their proof, de Klerk et al. [36] exploit the mentioned invariance of the max-k-cut problem under permutations of the subset. We show that using the symmetry reduction with respect to this invariance is a short way to illustrate the equivalence.

We denote relaxations of the max-k-cut problem on graph G by $c_k(G)$ and relaxations of the k-equipartition by $p_k(G)$. The vector lifting relaxations of both problems are particular cases of the relaxation of the general graph partition problem by Wolkowicz and Zhao [224].

Let L be the Laplacian of G. The symmetry-reduced versions of vector lifting relaxations are

$$
\begin{aligned}
c_k^v(G) = \max_{Z,X\in\mathbb{S}^n}\ & \tfrac{1}{2}\langle L, Z\rangle \\
\text{s.t.}\ & X_{ii} = 0, \quad \text{for } i \in [n] \\
& Z_{ii} = 1, \quad \text{for } i \in [n] \\
& Z \geq 0,\ X \geq 0 \\
& Z - X \succeq 0 \\
& Z + (k-1)X - J \succeq 0.
\end{aligned}
\qquad
\begin{aligned}
p_k^v(G) = \min_{Z,X\in\mathbb{S}^n}\ & \tfrac{1}{2}\langle L, Z\rangle \\
\text{s.t.}\ & X_{ii} = 0, \quad \text{for } i \in [n] \\
& Z_{ii} = 1, \quad \text{for } i \in [n] \\
& Z \geq 0,\ X \geq 0 \\
& Z - X \succeq 0 \\
& Z + (k-1)X - J \succeq 0 \\
& Ze = \tfrac{n}{k}e.
\end{aligned}
$$

Next, we look at the matrix lifting relaxations. For the k-equipartition, reducing the classical matrix lifting SDP relaxation results in well-known relaxation by Karisch and Rendl [102], which is equivalent to relaxation by Sotirov [212]. Similarly, for the max-k-cut, reducing the classical matrix lifting SDP relaxation results in the relaxation by van Dam and Sotirov [215], which is equivalent to the relaxation by Frieze and Jerrum [67].

$$
\begin{aligned}
c_k^m(G) = \max_{Z \in \mathbb{S}^n} \ & \tfrac{1}{2}\langle L, Z\rangle \\
\text{s.t.} \quad & Z_{ii} = 1, \quad \text{for } i \in [n] \\
& Z \geq 0, \\
& Z - \tfrac{1}{k} J \succeq 0.
\end{aligned}
\qquad
\begin{aligned}
p_k^m(G) = \min_{Z \in \mathbb{S}^n} \ & \tfrac{1}{2}\langle L, Z\rangle \\
\text{s.t.} \quad & Z_{ii} = 1, \quad \text{for } i \in [n] \\
& Z \geq 0 \\
& Z - \tfrac{1}{k} J \succeq 0 \\
& Ze = \tfrac{n}{k} e.
\end{aligned}
$$

To show that the vector and matrix lifting relaxations are equivalent, observe that for both the k-equipartition and max-k-cut one can construct feasible solutions to the matrix lifting relaxation from the vector lifting relaxation with the same objective value, and the other way round. First, from a feasible solution Z to each of the symmetry-reduced matrix-lifting relaxations we obtain a feasible solution $(Z, \frac{1}{k-1}(J-Z))$ to the corresponding symmetry-reduced vector lifting relaxation. For the opposite direction, a feasible solution (Z, X) to each of the symmetry-reduced vector lifting relaxations provides a feasible solution Z to the corresponding matrix lifting relaxation, as in the case of MkCS. Hence the vector and matrix lifting relaxations are equivalent.

5.6.3 The MkCS as the maximum stable set problem

In this section we show that the MkCS problem on a graph G can also be considered as the stable set problem on the Cartesian product of the complete graph on k vertices and G. Then, we show how is the Schrijver's ϑ'-number on the Cartesian product of mentioned two graphs related to the vector lifting relaxation (5.20).

We denote by $K_k = (V_k, E_k)$, where $V_k = [k]$ the complete graph on k vertices. The Cartesian product $K_k \square G$ of graphs K_k and $G = (V, E)$ is a graph with the vertex set $V_k \times V$ and the edge set $E_\square$ where two vertices (u, i) and (v, j) are adjacent if $u = v$ and $(i, j) \in E$ or $i = j$ and $(u, v) \in E_k$. The following result shows that the MkCS problem on G corresponds to the stable set problem on $K_k \square G$.

Theorem 5.18. *Let $G = (V, E)$, and let K_k be the complete graph on k vertices. Then $\alpha_k(G) = \alpha(K_k \square G)$.*

Proof. First, if $S_1, \ldots, S_k$ are disjoint stable sets in G, then $1 \times S_1, \ldots, k \times S_k$ is a stable set in $K_k \square G$. On the other hand, let S be a stable set in $K_k \square G$ of the largest cardinality. Then S can be partitioned into $S_1, \ldots, S_k$ such that $S_1 = 1 \times \hat{S}_1, \ldots, S_k = k \times \hat{S}_k$, $\hat{S}_1 \subseteq V, \ldots, \hat{S}_k \subseteq V$. Moreover, $\hat{S}_1, \ldots, \hat{S}_k$ are disjoint since $u \in \hat{S}_l \cap \hat{S}_p$ for some $l, p \in [k]$ with $l \neq p$ implies that there is an edge between (l, u) and (p, u) that is also in the stable set S. Hence $\hat{S}_1, \ldots, \hat{S}_k$ are disjoint stable sets in G. $\square$

Let us now compare the bounds for the two problems. The Schrijver's ϑ'-number on $K_k \Box G$ is as follows

$$\begin{aligned}
\vartheta'(K_k \Box G) = \max_{Y \in \mathbb{S}^{nk}} \quad & \langle J, Y \rangle \\
\text{s. t.} \quad & Y_{ij}^{rr} = 0, \quad \text{for all } \{ij\} \in E,\ r \in [k] \\
& Y_{ii}^{rl} = 0, \quad \text{for all } i \in [n],\ r, l \in [k],\ r \neq l \\
& \langle I, Y \rangle = 1, \\
& Y \succeq 0 \\
& Y \geq 0,
\end{aligned} \tag{5.56}$$

where Y is of the size $nk \times nk$. The above SDP relaxation follows directly from (5.11), (5.10), and the definition of $K_k \Box G$. We show below the following interesting result:

$$\vartheta'(K_k \Box G) = \theta_k^2(G) \geq \alpha_k(G),$$

where $\theta_k^2(G)$ is the optimal value of the vector lifting relaxation (5.20). To see this, first, notice that any feasible solution to the vector lifting relaxation (5.20) provides a feasible solution to problem (5.56). In particular let Y be feasible for (5.20), then it readily follows that $\hat{Y} = \frac{1}{\langle I, Y \rangle} Y$ is feasible for (5.56), see also [74]. Moreover, the SDP constraint in the vector lifting relaxation (5.20) implies

$$\langle J, \hat{Y} \rangle = \tfrac{1}{\langle I, Y \rangle} \langle J, Y \rangle \geq \tfrac{1}{\langle I, Y \rangle} \langle J, \operatorname{diag} Y \operatorname{diag} Y^\intercal \rangle = \langle I, Y \rangle.$$

For the opposite direction, we refer to Galli and Letchford [69] that showed if $\hat{Y}$ is feasible for (5.56), then there is a feasible solution Y for (5.20) such that $\langle I, Y \rangle \geq \langle J, \hat{Y} \rangle$, where

$$Y_{ij} = \hat{Y}_{ij} \frac{\sum_{m=1}^{kn} \hat{Y}_{im} \sum_{m=1}^{kn} \hat{Y}_{jm}}{\langle J, \hat{Y} \rangle \hat{Y}_{ii} \hat{Y}_{jj}} \text{ for } i \neq j, \text{ and } Y_{ii} = \frac{\left(\sum_{j=1}^{kn} \hat{Y}_{ij} \right)^2}{\langle J, \hat{Y} \rangle \hat{Y}_{ii}}, \text{ for } i \in [kn].$$

Note that the authors of [69] and [74] consider the SDP relaxations ϑ-number that do not include the non-negativity constraints on the matrix variables, in contrast to our relaxations (5.20) and (5.56).

5.7 Numerical results

In this section we compare the performance of all SDP bounds considered in this chapter with the existing results on MkCS. We are aware of two papers with numerical results on MkCS for general graphs and k: Campêlo and Corrêa [30], Januschowski and Pfetsch [93]. In the former paper the authors present an IP formulation of the

MkCS by representatives and implement a Lagrangian decomposition of that formulation. As a result, they obtain upper and lower bounds on the size of the maximum k-colorable subgraph. In the latter paper the authors propose a branch-and-cut method that accounts for both the symmetry with respect to the color permutations and the inner graph symmetry. They present the results only for the cases where they were able to solve the problem to optimality, and we can use only these graphs for comparison. We test our relaxations on the graphs from these papers where the number of vertices is not larger than 200[1]. This size is the largest one we can handle using up to 15 GB of RAM without exploiting the structure of the graphs. All computations were done in MATLAB R2018b with Yalmip [128] on a computer with two processors Intel® Xeon® Gold 6126 CPU @ 2.60 GHz and 512 GiB of RAM. SDP programs were solved with MOSEK Version 8.0.0.80.

Tables 5.1, 5.2 show our numerical results. In Table 5.1 we present the results for the graphs from [30]. These graphs are complements of the graphs that were used as benchmarks in the Second DIMACS Implementation Challenge for the max-clique problem [98]. In Table 5.2 we present the results for the graphs from [93]. Graphs marked by [02] were used as benchmarks in the COLOR02 symposium [99]. Other graphs are benchmarks from the Second DIMACS Challenge [98] or their complements. All graphs in the tables that end with "-c" are complements of the original graphs.

Now, we briefly describe the graphs. We begin with DIMACS graphs. "genA_pB-C" are artificially generated graphs with A vertices, edge density B and large, known embedded clique of the size C. "kellerA" with n vertices are graphs used to establish the results on Keller's conjecture in $\mathbb{R}^n$ [34, 113]. The vertices of these graphs are n-tuples of $0, 1, 2, 3$. Two vertices are adjacent iff there is a position at which the difference of the corresponding components is 2 modulo 4 and if there is a further position at which the corresponding components are different. "sanA_pB_C" and "sanrA_pB_C" are randomly generated graph instances that have A vertices and edge density B. "c-fatA-B" are graphs with A vertices. "Ca.b" are random graphs on a vertices with edge probability b. "hammingA-B" are Hamming graphs on A-bit words with an edge iff the two words are at least hamming distance B apart. "JohnsonA-B-C" are Johnson graphs generated by binary vectors of length A and weight B, with two vertices adjacent iff the Hamming distance between them is at least C. The last two graph types are vertex-transitive.

Next, consider the COLOR02 graphs. "Myciel" are graphs based on the Mycielski transformation. They are triangle free, but the coloring number increases in prob-

[1] We do not consider the graph "brock200.2" from [30] since it is not clear whether this graph or its complement was used in [30]. The density of the graph is 50%, so the bounds for the graph and the complement could be similar.

lem size so that the graphs can have arbitrary large gaps between $\omega(G)$ and $\chi(G)$. The “FullIns” and “Insertions” graphs are a generalization of the Mycielski graphs. “DSJc.n.p” are standard random graphs where an edge between two vertices appears with probability p “QueenA-A” graph is a graph with vertices that correspond to the squares of the $A \times A$ chess board and are connected by an edge if the corresponding squares are in the same row, column, or diagonals (according to the queen move rule at the chess game).

In our computations, we use the symmetry-reduced versions from Section 5.6 since they are equivalent to the versions from Section 5.5, but smaller and strictly feasible. We do not consider problem (5.47) since it is equivalent to problem (5.37) but has larger SDP constraints. To implement problem (5.37) with BQP inequalities (5.54) and (5.55), we use the cutting plane method. We add at most $2n$ BQP inequalities of each type at every iteration of the method and do at most four iterations.

We mark the graphs for which we obtain the optimal values in italic blue. We mark the best bounds for the given graph and k by boldface. All bounds are rounded to the nearest third digit. The reported times are the solver running times.

Table 5.1 – Results for graphs with up to 200 vertices considered by Campêlo and Corrêa [30].

Graph G		gen200_p0.9_55-c		keller4-c		san200_0.7_2-c		san200_0.9_2-c	
n		200		171		200		200	
$\|E\|$		1990		5100		5970		1990	
density, %		10		35		30		10	
Vertex-transitive?		no		no		no		no	
k		2	3	2	3	2	3	2	3
Upper bound [30]	value	109.00	161.60	27.90	41.90	36.00	54.000	117.00	184.00
	time, sec.	*46*	*71*	*27*	*30*	*32*	*28*	*1800*	*46*
$\vartheta_k(G)$ (5.9)	value	103.230	150.52	28.02	42.04	35.99	53.98	109.05	157.29
	time, sec.	*278*	*306*	*62*	*75*	*216*	*251*	*291*	*318*
$\vartheta'_k(G)$ (5.11)	value	102.79	149.70	**26.93**	**40.40**	35.68	53.34	108.68	156.59
	time, sec.	*497*	*472*	*121*	*138*	*373*	*404*	*558*	*488*
$\theta^6_k(G)$ (5.46)	value	100.84	146.22	**26.93**	**40.40**	35.63	53.24	106.61	152.84
	time, sec.	*330*	*309*	*64*	*74*	*197*	*189*	*426*	*307*
$\theta^5_k(G)$ (5.37) w.o.	value	100.52	146.01	**26.93**	**40.40**	35.61	53.24	106.24	152.26
(5.41) and (5.42)	*time, sec.*	*2427*	*3435*	*983*	*1087*	*2315*	*2699*	*3437*	*3878*
$\theta^5_k(G)$ (5.37)	value	**100.36**	145.61	**26.93**	**40.40**	**35.60**	53.240	**106.03**	151.44
	time, sec.	*3330*	*4123*	*1099*	*1298*	*3587*	*3650*	*3271*	*4055*
$\theta^5_k(G)$ (5.37) with	value	**100.36**	**145.35**	**26.93**	**40.40**	**35.60**	**53.22**	**106.03**	**151.16**
(5.54) and (5.55)	*time, sec.*	*14187*	*10893*	*1217*	*1190*	*12066*	*5508*	*13590*	*11889*
Lower bound [30]	value	69	99	20	30	31	43	75	98

Graph G		c-fat200-5-c		***c-fat200-2-c***		C125.9-c		gen200_p0.9_44-c	
n		200		200		125		200	
$\|E\|$		11427		16665		787		1990	
Density, %		57		84		10		10	
Vertex-transitive?		n/a		n/a		no		no	
k		2	3	2	3	2	3	2	3
Upper bound [30]	value	116.00	172.00	**46.00**	**68.00**	79.40	115.60	88.00	132.00
	time, sec.	*<1*	*<1*	*<1*	*<1*	*70*	*125*	*378*	*364*
$\vartheta_k(G)$ (5.9)	value	120.69	181.04	46.33	68.33	75.61	112.86	88.00	132.00
	time, sec.	*36*	*40*	*7*	*7*	*29*	*32*	*330*	*330*
$\vartheta'_k(G)$ (5.11)	value	120.69	181.04	46.33	68.33	75.09	112.18	88.00	132.00
	time, sec.	*61*	*52*	*9*	*8*	*51*	*54*	*507*	*460*
$\theta^6_k(G)$ (5.46)	value	120.69	181.03	**46.00**	**68.00**	74.63	107.26	88.00	131.94
	time, sec.	*35*	*33*	*5*	*5*	*29*	*25*	*258*	*258*
$\theta^5_k(G)$ (5.37) w.o.	value	120.69	181.01	**46.00**	**68.00**	74.41	106.96	88.00	131.90
(5.41) and (5.42)	*time, sec.*	*902*	*1018*	*940*	*726*	*228*	*322*	*2176*	*3013*
$\theta^5_k(G)$ (5.37)	value	120.69	175.99	**46.00**	**68.00**	74.12	105.90	88.00	131.84
	time, sec.	*1199*	*1148*	*879*	*616*	*295*	*342*	*2976*	*3528*
$\theta^5_k(G)$ (5.37) with	value	**119.42**	**175.23**	**46.00**	**68.00**	**74.10**	**105.31**	**87.99**	**131.66**
(5.54) and (5.55)	*time, sec.*	*7649*	*6955*	*3627*	*639*	*1130*	*1806*	*13570*	*16242*
Lower bound [30]	value	116	172	46	68	61	81	64	93

Table 5.2 – Results for graphs with up to 200 vertices considered by Januschowski and Pfetsch [93].

Graph G		[02]1-FullIns_4	[02]1-Insertions_4	[02]4-FullIns_3	[02]5-FullIns_3	*c-fat200-1-c*		*c-fat200-2-c*	
n		93	67	114	154	200		200	
$\|E\|$		593	232	541	792	18366		16665	
Density, %		14	10	8	7	92		84	
Vertex-transitive?		no	no	no	no	n/a		n/a	
k		3	3	3	3	6	7	7	8
$\alpha_k(G)$ [93]		87	63	106	144	72	84	156	178
$\vartheta_k(G)$ (5.9)	value	93.00	**67.00**	114.00	154.00	**72.00**	**84.00**	156.33	178.33
	time, sec.	*6*	*2*	*18*	*85*	*2*	*2*	*7*	*6*
$\vartheta'_k(G)$ (5.11)	value	93.00	**67.00**	114.00	154.00	**72.00**	**84.00**	156.33	178.33
	time, sec.	*8*	*2*	*22*	*100*	*2*	*3*	*10*	*10*
$\theta^6_k(G)$ (5.46)	value	92.59	**67.00**	107.400	145.33	**72.00**	**84.00**	**156.00**	**178.00**
	time, sec.	*11*	*1*	*20*	*89*	*1*	*1*	*5*	*4*
$\theta^5_k(G)$ (5.37) w.o.	value	92.57	**67.00**	107.31	145.25	**72.00**	**84.00**	**156.00**	**178.00**
(5.41) and (5.42)	*time, sec.*	*107*	*14*	*230*	*1047*	*577*	*739*	*615*	*537*
$\theta^5_k(G)$ (5.37)	value	92.43	**67.00**	107.30	145.25	**72.00**	**84.00**	**156.00**	**178.00**
	time, sec.	*103*	*10*	*196*	*713*	*776*	*778*	*620*	*644*
$\theta^5_k(G)$ (5.37) with	value	**91.33**	**67.00**	**107.25**	**145.23**	**72.00**	**84.00**	**156.00**	**178.00**
(5.54) and (5.55)	*time, sec.*	*554*	*34*	*901*	*3551*	*806*	*804*	*638*	*673*

Graph G		*hamming6-4-c*		Johnson8-4-4-c	[02]*DSJC125.9*			c-fat200-1	*gen200_p0.9_44*
n		64		70	125			200	200
$\|E\|$		1312		560	6961			1534	17910
Density, %		65		23	90			8	90
Vertex-transitive?		yes		yes	no			no	no
k		4	5	4	4	5	6	10	4
$\alpha_k(G)$ [93]		16	20	52	16	20	23	180	20
$\vartheta_k(G)$ (5.9)	value	21.33	26.67	**56.00**	**16.00**	**20.00**	23.96	184.67	**20.00**
	time, sec.	*<1*	*<1*	*1*	*<1*	*<1*	*<1*	*300*	*3*
$\vartheta'_k(G)$ (5.11)	value	**16.00**	**20.00**	**56.00**	**16.00**	**20.00**	23.95	184.67	**20.00**
	time, sec.	*<1*	*<1*	*2*	*<1*	*<1*	*1*	*742*	*3*
$\theta^6_k(G)$ (5.46)	value	**16.00**	**20.00**	**56.00**	**16.00**	**20.00**	**23.73**	184.67	**20.00**
	time, sec.	*<1*	*<1*	*<1*	*<1*	*<1*	*<1*	*300*	*2*
$\theta^5_k(G)$ (5.37) w.o.	value	**16.00**	**20.00**	**56.00**	**16.00**	**20.00**	**23.73**	184.67	**20.00**
(5.41) and (5.42)	*time, sec.*	*3*	*3*	*9*	*103*	*92*	*145*	*3612*	*955*
$\theta^5_k(G)$ (5.37)	value	**16.00**	**20.00**	**56.00**	**16.00**	**20.00**	**23.73**	**184.65**	**20.00**
	time, sec.	*3*	*3*	*9*	*104*	*106*	*141*	*5184*	*847*
$\theta^5_k(G)$ (5.37) with	value	**16.00**	**20.00**	**56.00**	**16.00**	**20.00**	**23.73**	**184.65**	**20.00**
(5.54) and (5.55)	*time, sec.*	*4*	*4*	*10*	*115*	*118*	*147*	*5485*	*1002*

Graph G		gen200_p0.9_55	^{02}myciel5		^{02}myciel6	^{02}queen6_6	*san200_0.9_1*	san200_0.9_2	sanr200_0.9
n		200	47		95	36	200	200	200
$\|E\|$		17910	236		755	580	17910	17910	17863
Density, %		90	22		17	92	90	90	90
Vertex-transitive?		no	no		no	no	no	no	no
k		4	4	5	3	6	4	4	4
$\alpha_k(G)$ [93]		17	44	46	83	32	16	16	16
$\vartheta_k(G)$ (5.9)	value	18.22	**47.00**	**47.00**	95.00	35.98	16.10	17.23	17.92
	time, sec.	*5*	*<1*	*<1*	*5*	*<1*	*4*	*4*	*4*
$\vartheta'_k(G)$ (5.11)	value	18.20	**47.00**	**47.00**	95.00	35.98	16.08	17.21	**17.91**
	time, sec.	*6*	*<1*	*<1*	*6*	*<1*	*4*	*5*	*4*
$\theta_k^6(G)$ (5.46)	value	**18.15**	**47.00**	**47.00**	95.00	35.84	16.08	17.21	**17.91**
	time, sec.	*4*	*<1*	*<1*	*4*	*<1*	*3*	*3*	*2*
$\theta_k^5(G)$ (5.37) w.o.	value	**18.15**	**47.00**	**47.00**	95.00	35.84	16.08	17.21	**17.91**
(5.41) and (5.42)	*time, sec.*	*1238*	*1*	*2*	*43*	*<1*	*1103*	*963*	*871*
$\theta_k^5(G)$ (5.37)	value	**18.15**	**47.00**	**47.00**	95.00	**35.81**	16.08	17.21	**17.91**
	time, sec.	*1365*	*1*	*2*	*37*	*<1*	*1282*	*1136*	*888*
$\theta_k^5(G)$ (5.37) with	value	**18.15**	**47.00**	**47.00**	**93.32**	**35.81**	**16.07**	**17.20**	**17.91**
(5.54) and (5.55)	*time, sec.*	*1386*	*4*	*7*	*366*	*3*	*2897*	*2036*	*918*

Table 5.1 shows that for all compared graphs, except “c-fat200-5-c”, SDP bounds are are at least as good as the upper bounds from [30]. In particular, for six out of eight graphs the best SDP bound is strictly better than the upper bound from [30], and for one graph all bounds are tight. Table 5.2 shows that we can find the optimal value for six out of eighteen graphs from [93].

The best bounds, as expected, are provided by the symmetry-reduced vector lifting relaxation $\theta_k^5(G)$ (5.37) with BQP constraints (5.54) and (5.55). However, in many cases the potentially stronger vector lifting relaxations (all modifications of $\theta_k^5(G)$) provide the same bound as cheaper SDP relaxations, especially the matrix lifting relaxation $\theta_k^6(G)$ (5.46). Also, the solution times for the bounds based on $\theta_k^5(G)$ are substantially larger than the times for other relaxations. This is especially remarkable for $\theta_k^5(G)$ with BQP constraints since in this case problem (5.37) has to be solved several times during the cutting places algorithm.

For all graphs the bound $\theta_k^6(G)$, which is the weakest among our relaxations, is at least as good as the bound $\vartheta'_k(G)$, and therefore $\vartheta_k(G)$. The computational time of all three bounds is similar, so $\theta_k^6(G)$ is the most preferable of them in terms of both bound quality and solution time.

5.8 Conclusion and questions for future research

In this work we analyze the existing SDP upper bounds and propose several new SDP upper bounds for the maximum k-colorable subgraph problem. The initial size of our new SDP relaxations depends on the number of colors k. We show how to reduce the sizes of SDP relaxations using symmetry with respect to the color permutations.

The reduction results in several SDP relaxations with at most two SDP constraints of order at most $(n+1)$ for any k and any graph type.

We compute all SDP bounds for graphs considered in [30, 93] with up to 200 vertices. We compare the resulting numbers to the optimal solutions obtained in [93] and to the upper and lower bounds from [30]. We solve the problem for some graphs to optimality and obtain stronger bounds than in [30] for all but one tested graphs. Also, the numerical experiments suggest that the weakest of the new SDP bounds is at least as good as the existing one for all graphs.

The main drawback of all SDP bounds analyzed in this chapter is the fact that solving SDP relaxations becomes too computationally demanding for graphs with more than 200 vertices. Many graphs from real life problems, and some graphs from [30, 93], have inherent symmetry which can be exploited to reduce the size of our SDP problems even further. For small graphs, instead of exploiting the symmetry, we can try to break it by fixing one vertex to be colored or uncolored (see, e.g. ,[187]). That is, we can fix the variable corresponding to that vertex to zero or one. On vertex-transitive graphs, fixing the values of one vertex is enough. For general graphs, breaking the symmetry for each of the vertices will provide a valid upper bound on $\alpha_k(G)$. Applying this approach to the relaxations from Section 5.6, one could improve the resulting bounds.

CHAPTER 6

New bounds for truthful scheduling on two unrelated selfish machines

6.1 Introduction and main results

Scheduling on unrelated parallel machines is a classical discrete optimization problem. In this problem one has to allocate n independent, indivisible tasks to m simultaneously working unrelated machines (not necessarily identical). The goal is to minimize the time to complete all the tasks. This time is called a makespan, and the scheduling problem is also called the minimum makespan problem. Lenstra et al. [124] proved that the problem is NP-complete, even for the case $n = m = 2$, and that a polynomial-time algorithm cannot achieve an approximation ratio less than $\frac{3}{2}$ unless $P = NP$.

We restrict ourselves to the case of $m = 2$ machines. For this case there is a linear-time algorithm by Potts [179] and a polynomial-time algorithm by Shchepin and Vakhania [208] which provide $\frac{3}{2}$-approximations. Both algorithms use linear relaxations of integer programs and rounding techniques. We are interested in truthful scheduling on unrelated machines. For this problem, the earlier mentioned algorithms are not suitable and the best approximation ratio is still unknown.

Truthful scheduling stems from the minimum makespan problem in the setting of algorithmic mechanism design. In this setting, every machine belongs to a rational agent who requires payments for performing tasks and aims to maximize his or her utility. Nisan and Ronen [163] introduced this approach to model interactions on the Internet, such as routing and information load balancing. The minimum makespan problem is one of many optimization problems considered in algorithmic mechanism design. These include, among others, combinatorial auctions (see, e.g., [10], [55] and references therein) and graph theoretic problems, such as the shortest paths tree [75] and the maximum matching problem [216].

To solve the minimum makespan problem in algorithmic mechanism design, one can use an allocation mechanism. An allocation mechanism consists of two algorithms:

one allocates tasks to machines, and the other one allocates payments to agents (the machines' owners). The goal of the mechanism is to minimize the makespan. We consider direct revelation mechanisms. These mechanisms collect the information about running times from each agent and allocate tasks and payments based on this information according to a policy known to the agents in advance. To maximize their utilities, the agents can lie about processing times of their machines. As a result, direct revelation mechanisms may be hard to implement correctly.

A truthful mechanism motivates the agents to tell the right processing times of their machines. That is, telling the truth becomes a dominant strategy for each agent regardless of what the other agents do. This property guarantees that the processing times used to construct the mechanism are correct. Hence the makespan can be estimated in advance without additional assumptions. There is a vast literature on truthful mechanisms [33, 147, 163, 196]. Not all task allocation algorithms correspond to truthful mechanisms. No truthful allocation mechanism is known, for example, for the previously mentioned linear programming (LP) relaxations with rounding by Potts [179] and Shchepin and Vakhania [208]. Therefore the best approximation ratio for truthful scheduling on unrelated machines is one of the hardest fundamental questions in mechanism design.

Saks and Yu [196] showed that a task allocation algorithm corresponds to a truthful mechanism if and only if the algorithm is monotone. Intuitively, a task allocation algorithm is monotone if it assigns a higher load to a machine as long as the running times on this machine decrease (see Section 6.2 for the formal definition of monotonicity). In this chapter we concentrate on monotone task allocation algorithms and do not consider the allocation of payments.

Nisan and Ronen [163] show that no deterministic monotone algorithm can achieve an approximation ratio less than 2, but randomized algorithms can do better in expectation. From here on we say that a randomized allocation algorithm has some property, e.g., monotonicity, if this property holds with probability one according to the distribution of the random bits of the algorithm. Randomized algorithms that are truthful in this sense give rise to universally truthful mechanisms considered in this chapter.

A deterministic algorithm is task-independent when the allocation of any task does not change as long as the processing times of this task stay fixed. Every deterministic monotone allocation algorithm on two machines with a finite approximation ratio is task-independent (Dobzinski and Sundararajan [55]). Therefore, if a given randomized algorithm has a finite expected approximation ratio, this algorithm is task independent with probability one. Hence we can restrict ourselves without loss of generality to monotone and task-independent randomized algorithms to find the best truthful approximation ratio for two machines.

Finally, we restrict our attention to scale-free algorithms. An algorithm is scale-free if scaling all running times by some positive number does not influence the output. Following Lu [130], we note that for $m = 2$, scale-freeness and allocation independence imply that the allocation of each task depends only on this task's running times ratio, which simplifies the analysis. Scale-free algorithms are widely used in the literature, and the latest most efficient algorithms for truthful scheduling on two machines by Chen et al. [32], Lu [130] or Lu and Yu [131] are scale-free. In the sequel we work with monotone, task-independent, scale-free (denoted by MIS) task allocation algorithms. These algorithms have proven to provide good upper bounds on approximation ratios in scheduling [32, 130, 131, 132, 163]. Lu and Yu [130, 131, 132] present a way to construct a payment allocation procedure for MIS algorithms which results in truthful allocation mechanisms.

To conclude the discussion about the properties of algorithms, notice that another, less restrictive, way to define randomized algorithms would be to say that the properties hold in expectation over the random bits of the algorithm. Randomized algorithms that are truthful in this sense correspond to truthful in expectation mechanisms. One can convert LP relaxations with rounding to such mechanisms for certain classes of problems (e.g., for combinatorial auctions), see Azar et al. [10], Elbassioni et al. [62] or Lavi and Swamy [122]. Truthful in expectation mechanisms could perform better in expectation than the universally truthful ones. In this chapter we do not analyze the former type of mechanisms. We refer the reader to Auletta et al. [9] and Lu and Yu [132] for more information on truthfulness in expectation.

Denote by R_n the best worst-case expected approximation ratio of randomized MIS algorithms for the makespan minimization on two machines with n tasks. For simplicity, in the rest of the chapter we call R_n the best approximation ratio. Our approach translates the problem of finding R_n from the context of truthful scheduling to the context of non-linear optimization. This approach was not common until recently when several successful truthful or truthful in expectation mechanisms have been constructed using linear or nonlinear programs [10, 32, 62, 122]. This chapter continues the trend to combine optimization with mechanism design and has the following contributions:

1. A $Min - Max$ formulation for R_n, see (6.19) in Corollary 6.14.
2. A unified approach to construct upper and lower bounds on R_n, see Section 6.4. In formulation (6.19), the outer minimization is over multivariate cumulative distribution functions (CDFs) and the inner maximization is over the positive orthant in two dimensions. This problem is in general not tractable, so we build bounds on the optimal value. The lower bounds are the result of restricting the inner maximization to a finite subset of the positive orthant. To obtain the upper bounds, we restrict the outer minimization to the set of piecewise

constant CDFs. This is a general approach which could work for any $Min{-}Max$ problem that requires optimizing over a set of functions, not necessarily CDFs.

3. New upper and lower bounds on R_n for $n \in \{2, 3, 4\}$ and the task allocation algorithms corresponding to the given upper bounds (see Table 6.1). To our knowledge, the resulting upper bounds are currently the best for all monotone algorithms (not only MIS) on two machines.

Table 6.1 – Bounds on R_n

n	Lower bound		Upper bound	
	existing	**new**	existing	**new**
2	1.505949 [130]	1.5059953	1.5068 [32]	1.5059964
3		1.5076	1.5861 [32]	1.5238
4		1.5195		1.5628

4. Almost tight bounds on R_2 (see Table 6.1).
For $n = 2$ tasks the initial problem (6.19) simplifies to a new problem (6.36), where the outer minimization is over univariate CDFs. We use piecewise rational CDFs to obtain the upper and lower bounds with a gap not larger than 10^{-6}.

The outline of the chapter is as follows. In Section 2 we provide more details about randomized MIS algorithms for two machines, describe results from earlier research and formulate the basic optimization problem for R_n. In Section 3 we exploit the symmetry of this problem to analyze the performance of MIS algorithms and to prove our $Min - Max$ formulation (6.19) for R_n. In Section 4 we construct and compute bounds on the optimal value of the $Min - Max$ problem for several small n. Section 5 analyzes the case with two tasks in more detail to improve the bounds for this case. Section 6 concludes the chapter. Section 7 provides the omitted proofs. All computations are done in MATLAB R2017a on a computer with the processor Intel® Core™ i5-3210M CPU @ 2.5 GHz and 7.7 GiB of RAM. To solve linear programs, we use IBM ILOG CPLEX 12.6.0 solver.

6.2 Preliminaries

Unless otherwise specified, lower-case letters denote numbers, bold lower-case letters denote vectors, and capital letters denote matrices. We use parentheses to denote vectors and brackets to denote intervals, e.g., (x_1, x_2) is a vector and $[x_1, x_2]$ or $[x_1, x_2)$ are intervals.

The input into the minimum makespan problem with m machines and n tasks is a matrix of processing times $T = (T_{ij}),\ i \in [m],\ j \in [n]$. We describe the solution to the problem by a task allocation matrix $X \in \{0,1\}^{m\times n}$, such that $X_{ij} = 1$ if task j is processed on machine i and $X_{ij} = 0$ otherwise. Now, for given X and T, we define the makespan of machine i M_i and the overall makespan M.

$$M_i(X,T) := \sum_{j=1}^{n} X_{ij}T_{ij}, \quad M(X,T) := \max_i M_i(X,T), \tag{6.1}$$

The optimal makespan for T is

$$M^*(T) := \min_{X\in\{0,1\}^{m\times n}} M(X,T). \tag{6.2}$$

For an allocation algorithm $\mathcal{A}$ and an input matrix T, $X^{\mathcal{A},T} \in \{0,1\}^{m\times n}$ denotes the output of $\mathcal{A}$ on T. If $\mathcal{A}$ is randomized, $X^{\mathcal{A},T}$ is a random variable and $M(X^{\mathcal{A},T},T)$ is the expected makespan. The worst-case (expected) approximation ratio of algorithm $\mathcal{A}$ equals the supremum of the ratio $\frac{M(X^{\mathcal{A},T},T)}{M^*(T)}$ over all time matrices T. We refer the reader to Motwani and Raghavan [145] for a comprehensive discussion on randomized algorithms.

6.2.1 The best approximation ratio of randomized MIS algorithms

According to Section 6.1, randomized MIS algorithms are monotone, task-independent and scale-free (MIS) with probability one. Therefore, by fixing the random bits of a randomized MIS algorithm, we obtain a deterministic MIS algorithm with probability one. We provide next a formal description of deterministic MIS algorithms.

A task allocation algorithm is monotone if for every two processing time matrices T and T' which differ only on machine i, $\sum_{j=1}^{n}(X_{ij}^{\mathcal{A},T} - X_{ij}^{\mathcal{A},T'})(T_{ij} - T'_{ij}) \leq 0$ (see [33]). That is, the load of a machine increases as long as the running times on this machine decrease. An algorithm is task-independent if the allocation of a task depends only on its running times. To be precise, for any two time matrices T and T' such that $T_{ij} = T'_{ij}$ for task j and all $i \in [m]$, the allocation of task j is identical, i.e., $X_{ij}^{\mathcal{A},T} = X_{ij}^{\mathcal{A},T'}$, for all $i \in [m]$. An algorithm is scale-free if the multiplication of all running times by the same positive number does not change the allocation. That is, for any $T \in \mathbb{R}_{++}^{m\times n}$ and $\lambda > 0$, the outputs of the algorithm on the inputs T and λT are identical.

Deterministic MIS algorithms for $m = 2$ are characterized by Lu [130]:

Theorem 6.1 (Lu [130]). *All deterministic MIS algorithms for scheduling on two unrelated machines are of the following form. For every task $j \in [n]$, assign a* *threshold* $z_j \in \mathbb{R}_{++}$ *and one of the following two conditions:* $T_{1j} < z_j T_{2j}$ *or* $T_{1j} \leq z_j T_{2j}$*. The task goes to the first machine if and only if the corresponding condition is satisfied.*

Let $\mathcal{C}$ be the class of (randomized) algorithms which randomly assign a threshold z_j and a condition $T_{1j} < z_j T_{2j}$ or $T_{1j} \leq z_j T_{2j}$ to each task j and then proceed as given in Theorem 6.1 for the deterministic case. With probability one a randomized MIS algorithm is a MIS algorithm, and therefore of the form given by Theorem 6.1. Hence, to find the best approximation ratio, it is enough to consider only algorithms in $\mathcal{C}$. Next, we show that to find the best approximation ratio, we can restrict ourselves to a subclass of $\mathcal{C}$.

Let $\mathcal{P}_n$ be the family of Borel probability measures supported on the positive orthant, i.e. $\text{supp}(\mathbb{P}) \subseteq \mathbb{R}^n_{++}$ for all $\mathbb{P} \in \mathcal{P}_n$, where $\text{supp}(\mathbb{P})$ is the support of $\mathbb{P}$. We use $\mathbf{E}_{\mathbb{P}}[\ ]$ and $\mathbf{P}_{\mathbb{P}}[\ ]$ to denote the expectation and the probability over the measure $\mathbb{P}$, respectively. In the sequel we use the notions of a probability measure and the corresponding probability distribution interchangeably. This depends on which notion is more suitable. For a $\mathbb{P} \in \mathcal{P}_n$ we define algorithm $\mathcal{A}^{\mathbb{P}}$ as follows:

Algorithm 6.2. *a monotone, task-independent, scale-free task allocation algorithm* $\mathcal{A}^{\mathbb{P}}$ *for 2 machines*

Input: processing time matrix $T = (T_{ij}) \in \mathbb{R}^{2\times n}_{++}$

Output: allocation $X \in \{0,1\}^{2\times n}$

1. Draw a vector of thresholds $(z_1, z_2, \ldots, z_n)$ according to $\mathbb{P}$
2. For each task $j = 1, 2, \ldots, n$ do
3. If $\frac{T_{1j}}{T_{2j}} < z_j$: $X_{1j} \leftarrow 1,\ X_{2j} \leftarrow 0$
4. Else: $X_{1j} \leftarrow 0,\ X_{2j} \leftarrow 1$
5. Output X

Denote the family of all algorithms of the form above by $\mathcal{A}^{\mathcal{P}_n}$:

$$\mathcal{A}^{\mathcal{P}_n} := \{\mathcal{A}^{\mathbb{P}} \ :\ \mathbb{P} \in \mathcal{P}_n\}.$$

Consider a measure $\mathbb{P} \in \mathcal{P}_n$ and the corresponding algorithm $\mathcal{A}^{\mathbb{P}} \in \mathcal{A}^{\mathcal{P}_n}$. Let $X^{\mathbb{P},T} \in \{0,1\}^{2\times n}$ be a randomized allocation produced by $\mathcal{A}^{\mathbb{P}}$ on time matrix T. For every $\mathbb{P}$ we define the expected makespan of $\mathcal{A}^{\mathbb{P}}$ on T as

$$M(\mathbb{P}, T) = \mathbf{E}_{\mathbb{P}} \max\left\{ \sum_{j\in[n]} T_{2j} X^{\mathbb{P},T}_{2j},\ \sum_{j\in[n]} T_{1j} X^{\mathbb{P},T}_{1j} \right\}. \tag{6.3}$$

Recall that $M^*(T)$, defined in (6.2), denotes the optimal makespan for T. Let $R_n(\mathbb{P}, T)$ be the expected approximation ratio of $\mathcal{A}^{\mathbb{P}}$ on T and $R_n(\mathbb{P})$ be the worst-case approximation ratio:

$$R_n(\mathbb{P}, T) = \frac{M(\mathbb{P}, T)}{M^*(T)}, \tag{6.4}$$

$$R_n(\mathbb{P}) = \sup_{T \in \mathbb{R}^{2\times n}_{++}} R_n(\mathbb{P}, T) \tag{6.5}$$

It could happen that for some $\mathbb{P}$ the ratio $R_n(\mathbb{P}, T)$ is unbounded in T. We do not consider these cases as we know that $R_n \leq 1.5861$ (see Section 6.2.2). To avoid technicalities, we work on $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$ so that the supremum $\sup_{T \in \mathbb{R}^{2\times n}_{++}} R_n(\mathbb{P}, T)$ is always defined.

When a tie $\frac{T_{1j}}{T_{2j}} = z_j$ occurs for some $j \in [n]$, algorithms from family $\mathcal{A}^{\mathcal{P}_n}$ send task j to the second machine. In general, an algorithm in $\mathcal{C}$ could send the task to the first or the second machine. Next, we show that this behavior at the ties does not affect the worst-case performance:

Theorem 6.3. *For a given number of tasks n, let $\mathbb{P} \in \mathcal{P}_n$ and define*

$$\mathcal{T} := \Big\{T \in \mathbb{R}^{2\times n}_{++} : \mathbf{P}_{z\sim\mathbb{P}}\Big[z_j = {}^{T_{1j}}/_{T_{2j}}\Big] = 0 \text{ for all } j \in [n]\Big\}. \tag{6.6}$$

Let $R_n(\mathbb{P})$ be defined as in (6.5)*, then*

$$R_n(\mathbb{P}) = \sup_{T \in \mathcal{T}} R_n(\mathbb{P}, T).$$

To prove Theorem 6.3, we introduce some additional notation. First, $\rho(T)$ denotes the vector of the running time ratios

$$\rho(T) := \Big(\tfrac{T_{11}}{T_{21}}, \tfrac{T_{12}}{T_{22}}, \dots, \tfrac{T_{1n}}{T_{2n}}\Big). \tag{6.7}$$

Second, for $\mathbf{z} \in \mathbb{R}^n_{++}$ we denote by $\mathcal{A}^{\mathbf{z}}$ the algorithm in $\mathcal{A}^{\mathcal{P}_n}$ with the thresholds fixed at $\mathbf{z}$. Finally, for $T \in \mathbb{R}^{2\times n}_{++}$ we let $X^{\mathbf{z},T}$ and $M(\mathbf{z}, T)$ be the output and the makespan of $\mathcal{A}^{\mathbf{z}}$ on T, respectively. We begin with a lemma.

Lemma 6.4. *For a given number of tasks n, let $\mathbb{P} \in \mathcal{P}_n$. For every time matrix $T \in \mathbb{R}^{2\times n}_{++}$, there exists a sequence of time matrices $\{T_k\}_{k>0} \subset \mathbb{R}^{2\times n}_{++}$ such that $\mathbf{P}_{z\sim\mathbb{P}}\Big[z_j = {}^{T_{k,1j}}/_{T_{k,2j}}\Big] = 0$ for all $j \in [n]$ and k, and*

$$R_n(\mathbb{P}, T) = \lim_{k\to\infty} R_n(\mathbb{P}, T_k).$$

Proof. Consider a sequence of non-negative numbers $\{\epsilon_k\}$ such that

1. $\mathbf{P}_{z\sim\mathbb{P}}\Big[z_j = \rho(T)_j + \epsilon_k\Big] = 0,\ \forall\ k \in \mathbb{N},\ j \in [n]$
2. $\lim\limits_{k\to\infty} \epsilon_k = 0$

A sequence with these properties exists since for every $j \in [n]$ the case $\mathbf{P}_{z\sim\mathbb{P}}\Big[z_j = a\Big] > 0$ is possible for countably many $a \in \mathbb{R}_{++}$ only. Next, we build a sequence of time matrices $\{T_k\}_{k>0},\ T_k = (T_{k,ij})$:

$$T_{k,1j} = T_{1j} + \epsilon_k T_{2j} \quad \text{and} \quad T_{k,2j} = T_{2j} \quad \text{for all } j \in [n].$$

Notice that $T = \lim_{k\to\infty} T_k$. By adding $\epsilon_k T_{2j}$ to every task on the first machine, we ensure that $\mathbf{P}_{z\sim\mathbb{P}}\Big[z_j = {T_{k,1j}}/{T_{k,2j}}\Big] = 0$ for all $j \in [n]$ and k. For each k and all $j \in [n],\ i \in \{1, 2\}$, we have $T_{ij} \leq T_{k,ij}$. So $M^*(T) \leq M^*(T_k) \leq M(X^*, T_k)$, where X^* is the optimal allocation for T. X^* is finite (binary, in particular) and $T = \lim_{k\to\infty} T_k$. Combining this with (6.1), we see that $M(X^*, T_k)$ tends to $M^*(T)$ when k tends to infinity. Therefore

$$\lim_{k\to\infty} M^*(T_k) = M^*(T). \tag{6.8}$$

For every time matrix T_k and task j, consider the event $B_{k,j}$: "$\rho(T)_j < z_j \leq \rho(T_k)_j$". Let $B_k = \bigcup_{j=1}^n B_{k,j}$ and let B_k^c be the complement of B_k. When B_k happens, outcomes of $\mathcal{A}^{\mathbf{z}}$ on T and T_k are different, otherwise they are the same. By construction of $\mathcal{A}^{\mathbf{z}}$, $M(\mathbf{z}, T)$ is finite for any T. Hence

$$\mathbf{E}_{\mathbf{z}\sim\mathbb{P}}\Big[M(\mathbf{z}, T_k) - M(\mathbf{z}, T) \mid B_k^c\Big] = 0, \quad \text{for all} \quad k \in \mathbb{N}. \tag{6.9}$$

For any $j \in [n]$ we have $\rho(T_k)_j \to \rho(T)_j^+$. Since the CDF of $\mathbb{P}$ is continuous from the right,

$$\lim_{k\to\infty} \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}\Big[z_j \leq \rho(T_k)_j\Big] = \mathbf{P}_{z\sim\mathbb{P}}\Big[z_j \leq \rho(T)_j\Big],$$

which implies

$$\begin{aligned}\lim_{k\to\infty} \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}[B_k] = \lim_{k\to\infty} \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}\left[\bigcup_{j=1}^n B_{k,j}\right] &\leq \sum_{j=1}^n \lim_{k\to\infty} \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}\Big[B_{k,j}\Big] \\ &= \sum_{j=1}^n \lim_{k\to\infty} \left(\mathbf{P}_{\mathbf{z}\sim\mathbb{P}}\Big[z_j \leq \rho(T_k)_j\Big] - \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}\Big[z_j \leq \rho(T)_j\Big]\right) = 0.\end{aligned} \tag{6.10}$$

Finally, for any $k \in \mathbb{N}$

$$\begin{aligned} M(\mathbb{P}, T_k) - M(\mathbb{P}, T) &= \mathbf{E}_{\mathbf{z}\sim\mathbb{P}}\Big[M(\mathbf{z}, T_k) - M(\mathbf{z}, T) \mid B_k\Big]\mathbf{P}_{\mathbf{z}\sim\mathbb{P}}[B_k] \\ &\quad + \mathbf{E}_{\mathbf{z}\sim\mathbb{P}}\Big[M(\mathbf{z}, T_k) - M(\mathbf{z}, T) \mid B_k^c\Big]\Big(1 - \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}[B_k]\Big) \\ &\overset{(6.9)}{=} \mathbf{E}_{\mathbf{z}\sim\mathbb{P}}\Big[M(\mathbf{z}, T_k) - M(\mathbf{z}, T) \mid B_k\Big]\mathbf{P}_{\mathbf{z}\sim\mathbb{P}}[B_k] \\ &\leq |T|_1 \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}[B_k] \end{aligned}$$

Thus by (6.10),

$$\lim_{k\to\infty} M(\mathbb{P}, T_k) = M(\mathbb{P}, T), \tag{6.11}$$

and

$$R_n(\mathbb{P}, T) = \frac{M(\mathbb{P}, T)}{M^*(T)} \overset{(6.8),(6.11)}{=} \lim_{k\to\infty} \frac{M(\mathbb{P}, T_k)}{M^*(T_k)} = \lim_{k\to\infty} R_n(\mathbb{P}, T_k).$$

□

Proof of Theorem 6.3 . Recall that $\mathcal{T}$ is defined in (6.6). By Lemma 6.4, for every $T \in \mathbb{R}^{2\times n}_{++}$ there exists a sequence of time matrices $\{T_k\}_{k>0} \subset \mathcal{T}$ such that

$$R_n(\mathbb{P}, T) = \lim_{k\to\infty} R_n(\mathbb{P}, T_k) \leq \sup_{T\in\mathcal{T}} R_n(\mathbb{P}, T).$$

Hence

$$R_n(\mathbb{P}) = \sup_{T\in\mathbb{R}^{2\times n}_{++}} R_n(\mathbb{P}, T) \leq \sup_{T\in\mathcal{T}} R_n(\mathbb{P}, T).$$

The opposite inequality holds since $\mathcal{T} \subset \mathbb{R}^{2\times n}_{++}$. □

Theorem 6.3 has the following implication:

Corollary 6.5. *The best approximation ratio over all randomized MIS algorithms is the best approximation ratio over all algorithms in $\mathcal{A}^{\mathcal{P}_n}$.*

By Corollary (6.5), the best approximation ratio over all randomized MIS algorithms is

$$R_n = \inf_{\mathbb{P}\in\mathcal{P}_n} R_n(\mathbb{P}). \tag{6.12}$$

Later in the chapter we show that $R_n(\mathbb{P})$ is invariant under permutations of the tasks for every $\mathbb{P} \in \mathcal{P}_n$ (see Theorem 6.6). Therefore, to compute R_n using (6.12), we can restrict the optimization to the distributions $\mathbb{P} \in \mathcal{P}_n$ invariant under permutations of the random variables corresponding to the thresholds $(z_1, \ldots, z_n)$ (see Theorem 6.8). For such distributions, $R_n(\mathbb{P})$ is determined by the worst-case performance of $\mathcal{A}^{\mathbb{P}}$ on each pair of the tasks. See Section 6.3.2, and Theorem 6.11 in particular, for more detail. This means that for every $T \in \mathbb{R}^{2\times n}_{++}$, one can find the expected approximation ratio of $\mathcal{A}^{\mathbb{P}}$ by applying to all pairs of tasks the algorithm $\mathcal{A}^{\mathbb{P}_2}$, where $\mathbb{P}_2$ is the bivariate marginal distribution of $\mathbb{P}$ (by invariance, this distribution is the same for all pairs of thresholds). We use this property of family $\mathcal{A}^{\mathcal{P}_n}$ to construct problem (6.19) for R_n, which is one of the main results of this chapter.

6.2.2 Connection to the current knowledge on monotone algorithms

The best approximation ratio for all monotone task allocation algorithms is not known. For deterministic algorithms with n tasks and m machines, when n and m tend to infinity, the ratio lies in the interval $[1+\phi,\ m]$, where ϕ is the golden ratio. The upper bound is due to Nisan and Ronen [163], and the lower bound is due to Koutsoupias and Vidali [106]. To compute this lower bound, the authors use a matrix of processing times where the numbers of rows and columns tend to infinity. If n or m is finite, the lower bound may be different. Koutsoupias and Vidali [106] present bounds for several finite time matrices as well. For randomized algorithms, the best approximation ratio lies in the interval $[2-\frac{1}{m},\ 0.83685m]$. The lower bound is due to Mu'alem and Schapira [147] who use Yao's minimax principle (for some details on this principle see, e.g., Motwani and Raghavan [145]). The upper bound is due to Lu and Yu [131]. The gap between the bounds grows with m, and the case with the smallest number of machines, $m=2$, has gained much attention in the literature.

For $m=2$, Nisan and Ronen [163] have shown that the best approximation ratio of deterministic monotone algorithms is equal to 2, for any finite n. The ratio for randomized monotone algorithms lies in the interval $[1.5, 1.5861]$. Chen et al. [32] compute the upper bound using an algorithm from family $\mathcal{A}^{\mathcal{P}_n}$ with independently distributed thresholds. The lower bound is the earlier mentioned bound by Mu'alem and Schapira [147]. There exist tighter lower bounds for certain cases. Lu [130] shows that algorithms from family $\mathcal{A}^{\mathcal{P}_n}$ (and thus, by Corollary 6.5, all randomized MIS algorithms) cannot achieve a ratio better than $\frac{25}{16}$ $(=1.5625)$ for sufficiently large n. Chen et al. [32] prove that algorithm $\mathcal{A}^{\mathbb{P}}$ cannot do better than 1.5852 when $\mathbb{P}$ is a product measure, i.e., when the thresholds are independent random variables.

The cases with $m=2$ and small $n>2$ are not well studied. Chen et al. [32] present upper bounds for some n. Finding these bounds requires solving some non-convex optimization problems. However, the numerical method used by Chen et al. [32] does not guaranty global optimality.

The case with $m=2,\ n=2$ is the simplest one, but even for this case the best approximation ratio is unknown. The ratio for algorithms from family $\mathcal{A}^{\mathcal{P}_2}$ lies in the interval $[1.505949, 1.5068]$. The upper bound is due to Chen et al. [32], the lower bound is computed by Lu [130] using Yao's minimax principle. Notice that Lu [130] states that the lower bound is 1.506, but we repeated the calculations from this chapter and obtained the number 1.505949. Thus, when reporting results, we use this number as the currently best lower bound. We improve this bound and show that $|R_2 - 1.505996| < 10^{-6}$, in particular, $R_2 < 1.506$.

From the description above, one can see that the existing bounds for truthful scheduling on unrelated machines are obtained using ad hoc procedures or (for the lower bounds) Yao's minimax principle. This chapter develops a unified approach to construct upper and lower bounds on R_n for any fixed n and provides an alternative to Yao's minimax principle for the construction of lower bounds. As a result, we improve the bounds for truthful scheduling for $m = 2$ and $n \in \{2, 3, 4\}$.

Next, we compare our approach to the existing methods for upper bounds in [10, 32, 62, 122] that use optimization. First, our method for upper bounds generalizes the approach by Chen et al. [32]. This generalization considers a broader class of algorithms and thus provides stronger upper bounds. The methods in [10, 62, 122] allow constructing truthful in expectation algorithms for some algorithmic mechanism design problems. These methods use LP relaxations of the corresponding integer problems. The minimum makespan problem on unrelated machines is not among the problems for which the approaches in [10, 62, 122] are guaranteed to work. Our approach is fundamentally different. First, we use the tools from continuous optimization to obtain possibly non-linear, but tractable approximations. Second, our algorithms are universally truthful. Next, our method works for the minimum makespan problem on unrelated machines. An open question for further research is for which other problems this technique could be useful.

Finally, we obtain new bounds only for small $n \leq 4$ because of the growing size of the lower bound optimization problems. However, one can make these problems less computationally demanding by using, for example, the column generation technique (see, e.g., Gilmore and Gomory [71]). With a more efficient lower bound computation, one could obtain new bounds for $n > 4$ with our approach since the solution to the upper bound problem is a relatively simple construction based on the solution to the lower bound problem, as described in Section 6.4.

6.3 Using the symmetry of the problem to obtain a new formulation for the best approximation ratio

In this section we exploit the fact that problem (6.12) is invariant under permuting the tasks and the machines to simplify formulation (6.12) and obtain formulation (6.19) in Section 6.3.2.

6.3.1 Using the symmetry of the problem

Recall that by Definition 1.1, the group $Sym(n)$ acts on a given time matrix from the right T by permuting its columns. Now, to reflect row permutations in T we define another group called S_{inv}. The group S_{inv} consists of the identity action ${}^{id}.$ and the action ${}^{inv}.$ which takes element-wise reciprocals of any vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_{++}$:

$$^{id}\mathbf{x} = \mathbf{x}, \quad ^{inv}\mathbf{x} = \left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}\right).$$

Now, we define the action of $S_{inv} \times Sym(n)$ on $\mathcal{P}_n$. Given $\mathbb{P} \in \mathcal{P}_n$, $\gamma \in S_{inv}$, $\pi \in Sym(n)$ and a random variable $\mathbf{z} \sim \mathbb{P}$, we consider the transformation $\mathbf{z} \to {}^{\gamma}\mathbf{z}\pi$. We define ${}^{\gamma}\mathbb{P}\pi \in \mathcal{P}_n$ as the distribution of ${}^{\gamma}\mathbf{z}\pi$. Next, we prove that problem (6.12) is convex and invariant under the action of $S_{inv} \times Sym(n)$ on $\mathbb{P}$. As a result, to find the infimum in (6.12), it is enough to optimize over the distributions $\mathbb{P}$ invariant under the action of $S_{inv} \times Sym(n)$. This approach is regularly used in convex programming, see Dobre and Vera [53], Gatermann and Parrilo [70] or de Klerk et al. [37].

Given distributions $\mathbb{P}_1, \ldots, \mathbb{P}_k \in \mathcal{P}_n$, and weights $\alpha_i \geq 0$ for all $i \in [k]$ such that $\sum_{i=1}^k \alpha_i = 1$, we define the convex combination $\sum_{i=1}^k \alpha_i \mathbb{P}_i \in \mathcal{P}_n$ as the distribution where we draw from P_i, $i \in [k]$ with probability α_i. The construction of $\sum_{i=1}^k \alpha_i \mathbb{P}_i$ and definitions (6.3),(6.4) imply that

$$R_n\left(\sum_{i=1}^k \alpha_i \mathbb{P}_i, T\right) = \sum_{i=1}^k \alpha_i R_n(\mathbb{P}_i, T).$$

Therefore, using (6.5), we have

$$R_n\left(\sum_{i=1}^k \alpha_i \mathbb{P}_i\right) \leq \sum_{i=1}^k \alpha_i R_n(\mathbb{P}_i), \tag{6.13}$$

that is, $R_n(\mathbb{P})$ is convex in $\mathbb{P}$. Now, we show the invariance of $R_n(\mathbb{P})$ under the action of $S_{inv} \times Sym(n)$.

Theorem 6.6. *For any given number of tasks n, $\mathbb{P} \in \mathcal{P}_n$, $\gamma \in S_{inv}$ and $\pi \in Sym(n)$,*

$$R_n(\mathbb{P}) = R_n({}^{\gamma}\mathbb{P}\pi).$$

To prove Theorem 6.6, let σ_{id} be the identity action of S_2 and σ_{swap} be the non-identity action of S_2. We say that that the group $S_2 \times Sym(n)$ acts on a matrix in $\mathbb{R}^{2\times n}$ by permuting the two rows and the n columns of this matrix, see Definition 1.1. Namely, for $A \in \mathbb{R}^{2\times n}$ and $(\sigma, \pi) \in S_2 \times Sym(n)$, the action of $(\sigma, \pi) \in S_2 \times Sym(n)$ on A is

$$\sigma A \pi := (A_{\sigma i, \pi j}).$$

We show that for any $T \in \mathbb{R}^{2\times n}_{++}$ the optimal makespan $M^*(T)$ is invariant under the actions of $S_2 \times Sym(n)$ on T. Moreover, the expected makespan $M(\mathbb{P}, T)$ is invariant under the actions of $S_2 \times Sym(n)$ on T and $S_{inv} \times Sym(n)$ on $\mathbb{P}$. These two results together imply Theorem 6.6.

Lemma 6.7. $M^*(T) = M^*(\sigma T \pi)$ *for all* $(\sigma, \pi) \in S_2 \times Sym(n),\ T \in \mathbb{R}^{2\times n}_{++}$

Proof. Consider a time matrix T and actions $\pi \in Sym(n)$, $\sigma \in S_2$. Let $X^* = (X^*_{ij})$ be an optimal allocation matrix for T. Then $\sigma T\pi = (T_{\sigma i,\pi j})$, $\sigma X^* \pi = (X^*_{\sigma i,\pi j})$. This implies

$$M^*(\sigma T\pi) \leq \max_i \Big\{ \sum_{j\in[n]} T_{\sigma i,\pi j} X^*_{\sigma i,\pi j} \Big\} = \max_i \Big\{ \sum_{j\in[n]} T_{ij} X^*_{ij} \Big\} = M^*(T).$$

Analogously, for the time matrix $\sigma T\pi$ and actions $\pi^{-1} \in Sym(n)$, $\sigma^{-1} \in S_2$, we obtain

$$M^*(T) \leq M^*(\sigma T\pi).$$

□

Proof of Theorem 6.6. Let $\mathbb{P} \in \mathcal{P}_n$, $\mathbf{z} \in \mathbb{R}^n_{++}$ and $T \in \mathbb{R}^{2\times n}_{++}$. Since we are interested in $R_n(\mathbb{P})$, by Lemma 6.4 we can assume without loss of generality that T is such that $\mathbf{P}_{z\sim\mathbb{P}}\Big[z_j = {T_{1j}}/{T_{2j}}\Big] = 0$ for all $j \in [n]$. Consider $(\gamma,\pi) \in S_{inv}\times Sym(n)$ and $\mathbf{y} = {}^{\gamma}\mathbf{z}\pi$. Let $\sigma = \sigma_{id}$ if $\gamma =^{id}$. and $\sigma = \sigma_{swap}$ if $\gamma =^{inv}$. Then $\mathcal{A}^{\mathbf{z}}$ sends task j to machine i on T if and only if $\mathcal{A}^{\mathbf{y}}$ sends task πj to machine σi on $\sigma T\pi$. As a result, $T_{ij}X^{\mathbf{z},T}_{ij} = T_{\sigma i\pi j}X^{\mathbf{y},\sigma T\pi}_{\sigma i\pi j}$ for all i,j and

$$\begin{aligned} M\Big(\mathbf{z},T\Big) &= \max_i \left\{ \sum_{j\in[n]} T_{ij}X^{\mathbf{z},T}_{ij} \right\} = \max_i \left\{ \sum_{j\in[n]} T_{\sigma i\pi j}X^{\mathbf{y},\sigma T\pi}_{\sigma i\pi j} \right\} \\ &= \max_i \left\{ \sum_{j\in[n]} T_{ij}X^{\mathbf{y},\sigma T\pi}_{ij} \right\} = M\Big(\mathbf{y},\sigma T\pi\Big). \end{aligned}$$

Therefore

$$M(\mathbb{P},T) = \mathbf{E}_{\mathbf{z}\sim\mathbb{P}} M\Big(\mathbf{z},T\Big) = \mathbf{E}_{\mathbf{y}\sim{}^{\gamma}\mathbb{P}\pi} M\Big(\mathbf{y},\sigma T\pi\Big) = M({}^{\gamma}\mathbb{P}\pi,\sigma T\pi), \tag{6.14}$$

Combining this with Lemma 6.7, obtain

$$R_n(\mathbb{P},T) = R_n({}^{\gamma}\mathbb{P}\pi,\sigma T\pi).$$

By Theorem 6.3,

$$R_n(\mathbb{P}) = \sup_{T\in\mathcal{T}} R_n(\mathbb{P},T) = \sup_{T\in\mathcal{T}} R_n({}^{\gamma}\mathbb{P}\pi,\sigma T\pi) = R_n({}^{\gamma}\mathbb{P}\pi),$$

where $\mathcal{T}$ is defined in (6.6). □

Theorem 6.8. *For any given number of tasks n,*

$$R_n = \inf_{\mathbb{P}\in\mathcal{P}_n} R_n(\mathbb{P}) \text{ such that } \mathbb{P} \text{ is invariant under the action of } S_{inv}\times Sym(n). \tag{6.15}$$

Proof. As problem (6.15) has a smaller feasibility set than problem (6.12), the optimal value of problem (6.15) is not smaller than R_n. To prove the opposite inequality, we show that for any distribution $\mathbb{P} \in \mathcal{P}_n$ there is a distribution $\mathbb{Q} \in \mathcal{P}_n$ invariant under the action of $S_{inv} \times Sym(n)$ such that $R_n(\mathbb{Q}) \leq R_n(\mathbb{P})$. Given $\mathbb{P} \in \mathcal{P}_n$, take $\alpha_i = \frac{1}{2(n!)}$ for $i \in [2(n!)]$ and consider the convex combination

$$\mathbb{Q} := \frac{1}{2(n!)} \sum_{(\gamma,\pi)\in S_{inv}\times Sym(n)} {}^{\gamma}\mathbb{P}\pi.$$

By construction, $\mathbb{Q}$ has the required invariance property and

$$R_n(\mathbb{Q}) \overset{(6.13)}{\leq} \frac{1}{2(n!)} \sum_{(\gamma,\pi)\in S_{inv}\times Sym(n)} R_n({}^{\gamma}\mathbb{P}\pi) \overset{\text{Theorem 6.6}}{=} \frac{1}{2(n!)} \sum_{(\gamma,\pi)\in S_{inv}\times Sym(n)} R_n(\mathbb{P})$$
$$= R_n(\mathbb{P}).$$

□

6.3.2 New formulation for the best approximation ratio

From Theorem 6.8, problem (6.12) is invariant under permuting the tasks and the machines. In the sequel we exploit the invariance under permuting the tasks only. First, this simplifies the presentation. Second, in our numerical computations using the invariance under the two types of permutations produced the same bounds as using invariance under task permutations only.

Let $\mathcal{C}_n \subset \mathcal{P}_n$ be the family of probability measures invariant under the actions of $Sym(n)$:

$$\mathcal{C}_n = \{\mathbb{P} \in \mathcal{P}_n \mid \mathbb{P} = \mathbb{P}\pi, \quad \text{for all} \quad \pi \in Sym(n)\}. \tag{6.16}$$

In the rest of the chapter we restrict the optimization to the distributions from $\mathcal{C}_n$.

Corollary 6.9. *For any given number of tasks n,*

$$R_n = \inf_{\mathbb{P}\in\mathcal{C}_n} R_n(\mathbb{P}). \tag{6.17}$$

Proof. The Corollary follows from Theorem 6.8 and

$$\{\mathbb{P} \in \mathcal{P}_n \mid \mathbb{P} = {}^{\gamma}\mathbb{P}\pi, \quad \text{for all} \quad (\gamma,\pi) \in S_{inv} \times Sym(n)\} \subset \mathcal{C}_n \subset \mathcal{P}_n.$$

□

Proposition 6.10 next is straightforward but crucial for our analysis.

Proposition 6.10. *Let $\mathbb{P} \in \mathcal{C}_n$. Then $\mathbb{P}$ has a cumulative distribution function (CDF) invariant under permutations of the variables. Moreover, for $0 < k < n$, all k-variate marginal distributions are identical. In particular, $\mathbb{P}$ is a joint distribution of n identically distributed random variables.*

By Proposition 6.10, if $\mathbb{P} \in \mathcal{C}_n$, then all univariate marginal distributions of $\mathbb{P}$ are identical and all bivariate marginal distributions of $\mathbb{P}$ are identical. Denote the corresponding univariate and bivariate CDFs by $F_{\mathbb{P}}$ and $H_{\mathbb{P}}$, respectively, and define

$$\phi^{\mathbb{P}}(x,y) = 1 + y - \min\left\{1,\ 1 - {}^1\!/\!_x + y\right\} F_{\mathbb{P}}(x) - yF_{\mathbb{P}}(y) + \min\left\{1 + {}^1\!/\!_x,\ 1 + y\right\} H_{\mathbb{P}}(x,y) \tag{6.18}$$

First, we present a result by Chen et al. [32], which follows from Lu and Yu [132]

Theorem 6.11 (Chen et al. [32])**.** *For any given number of tasks n, $\mathbb{P} \in \mathcal{C}_n$, and $T \in \mathbb{R}^{2\times n}_{++}$,*

$$R_n(\mathbb{P}, T) \leq \max_{j,k\in[n]} \phi^{\mathbb{P}}\left(\frac{T_{1j}}{T_{2j}}, \frac{T_{1k}}{T_{2k}}\right).$$

Notice that this upper bound is defined by only two tasks out of n. Using Theorem 6.11, we obtain the following formulation for $R_n(\mathbb{P})$:

Theorem 6.12. *For any given number of tasks n, and $\mathbb{P} \in \mathcal{C}_n$,*

$$R_n(\mathbb{P}) = \sup_{x,y\in\mathbb{R}_{++}} \phi^{\mathbb{P}}(x,y).$$

Proof. By Theorem 6.11, $R_n(\mathbb{P}) \leq \sup_{x,y\in\mathbb{R}_{++}} \phi^{\mathbb{P}}(x,y)$. Next, we prove the opposite inequality. Consider $\mathbb{P} \in \mathcal{C}_n$. We start with the case $n = 2$ and proceed similarly to Lu and Yu [132]. Denote the bivariate marginal distribution of $\mathbb{P}$ by $\mathbb{P}_2$. Consider a time matrix $T \in \mathbb{R}^{2\times 2}_{++}$ and denote $\frac{T_{11}}{T_{21}}$ by x, $\frac{T_{12}}{T_{22}}$ by y. Construct the following matrix T_0:

	Task 1	Task 2
Machine 1	1	y
Machine 2	${}^1\!/\!_x$	1

The expected makespan of $\mathcal{A}^{\mathbb{P}_2}$ on this instance is $M(\mathbb{P}_2, T_0)$:

$$\begin{aligned}
M(\mathbb{P}_2, T_0) &= \mathbf{P}_{\mathbf{z}\sim\mathbb{P}_2}\Big[z_1 > x,\ z_2 > y\Big](1+y) + \mathbf{P}_{\mathbf{z}\sim\mathbb{P}_2}\Big[z_1 > x,\ z_2 \le y\Big] \\
&\quad + \mathbf{P}_{\mathbf{z}\sim\mathbb{P}_2}\Big[z_1 \le x,\ z_2 > y\Big] \max\Big\{{}^1\!/\!_x,\ y\Big\} \\
&\quad + \mathbf{P}_{\mathbf{z}\sim\mathbb{P}_2}\Big[z_1 \le x,\ z_2 \le y\Big]\Big(1 + {}^1\!/\!_x\Big) \\
&= \Big[1 - H(x,y) - \Big(F(y) - H(x,y)\Big) - \Big(F(x) - H(x,y)\Big)\Big](1+y) \\
&\quad + \Big(F(y) - H(x,y)\Big) + \Big(F(x) - H(x,y)\Big)\max\Big\{{}^1\!/\!_x, y\Big\} \\
&\quad + H(x,y)\Big(1 + {}^1\!/\!_x\Big) \\
&= 1 + y - F(x)\Big(1 + y - \max\Big\{{}^1\!/\!_x,\ y\Big\}\Big) - yF(y) \\
&\quad + H(x,y)\Big(y - \max\Big\{{}^1\!/\!_x,\ y\Big\} + 1 + {}^1\!/\!_x\Big) \\
&= 1 + y - \min\Big\{1,\ 1 - {}^1\!/\!_x + y\Big\}F(x) - yF(y) \\
&\quad + \min\Big\{1 + {}^1\!/\!_x, 1 + y\Big\}H(x,y) \\
&= \phi^{\mathbb{P}}(x,y).
\end{aligned}$$

Denote the minimum makespan on T_0 by M^*. By construction $M^* \le 1$, hence

$$R_2(\mathbb{P}_2) \ge R_2(\mathbb{P}_2, T_0) = \frac{M(\mathbb{P}_2, T_0)}{M^*} \ge M(\mathbb{P}_2, T_0) = \phi^{\mathbb{P}}(x,y).$$

This holds for all $x, y \in \mathbb{R}_{++}$, therefore

$$R_2(\mathbb{P}_2) \ge \sup_{x,y\in\mathbb{R}_{++}} \phi^{\mathbb{P}}(x,y).$$

Now, fix $n > 2$. Choose a small $\epsilon > 0$ and consider the following time matrix T_ϵ:

	Task 1	Task 2	Task 3	...	Task n
Machine 1	1	y	ϵ	...	ϵ
Machine 2	${}^1\!/\!_x$	1	ϵ	...	ϵ

The expected makespan of $\mathcal{A}^{\mathbb{P}}$ on this instance, $M(\mathbb{P}, T_\epsilon)$, satisfies

$$M(\mathbb{P}_2, T_0) \le M(\mathbb{P}, T_\epsilon) \le M(\mathbb{P}_2, T_0) + (n-2)\epsilon,$$

and the optimal makespan, $M^*(T_\epsilon)$, satisfies

$$M^* \le M^*(T_\epsilon) \le M^* + (n-2)\epsilon.$$

Using the result for $n = 2$,

$$R_n(\mathbb{P}) \ge \lim_{\epsilon\to 0} R_n(\mathbb{P}, T_\epsilon) = \lim_{\epsilon\to 0} \frac{M(\mathbb{P}, T_\epsilon)}{M^*(T_\epsilon)} = \frac{M(\mathbb{P}_2, T_0)}{M^*} \ge \phi^{\mathbb{P}}(x,y),$$

which holds for all $x, y \in \mathbb{R}_{++}$. □

Remark 6.13. *Theorem 6.12 implies that the worst-case approximation ratio for n tasks and $\mathbb{P} \in \mathcal{P}_n$ is the worst-case approximation ratio for two tasks and the bivariate marginal distribution of $\mathbb{P}$.*

The next corollary is the main result of this section, and we use it throughout the rest of the chapter.

Corollary 6.14. *For any given number of tasks n,*

$$R_n = \inf_{\mathbb{P} \in \mathcal{C}_n} \sup_{x,y \in \mathbb{R}_{++}} \phi^{\mathbb{P}}(x, y). \tag{6.19}$$

Proof. The result follows from Corollary 6.9 and Theorem 6.12. □

Corollary 6.15. *$R_{n+1} \geq R_n$ for all $n \geq 2$.*

Proof. The result follows from Corollary 6.14. □

6.4 Upper and lower bounds on the best approximation ratio

To find R_n using (6.19), one needs to optimize over a family of distributions, which is computationally intractable. Therefore we construct upper and lower bounds on the optimal value of the problem. The idea is to restrict the attention to some subset of feasible distributions or some subset of $\mathbb{R}^2_{++}$, over which it is easier to solve problem (6.19).

1. For the lower bound, we take a finite set $\mathcal{S} \subset \mathbb{R}_{++}$ and find the supremum in (6.19) for $x, y \in \mathcal{S}$ only. A conventional approach to lower bounds is to propose several good-guess time matrices T, use these matrices to build a randomized instance of the minimum makespan problem and apply Yao's minimax principle. Our approach is different as we evaluate randomized algorithms on deterministic instances.

2. For the upper bound, we find a good-guess distribution $\mathbb{P}$ and solve the inner maximization problem for this distribution. The distribution is built using the solution to the lower bound problem for $n \in \{2, 3, 4\}$. For $n = 2$ we propose a more efficient approach in Section 6.5.

6.4.1 Characterizing CDFs

To implement the ideas above, we have to optimize over distributions. For this purpose we represent a distribution via its CDF. To characterize CDFs, we follow

Nelsen [155]. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that $x_i \leq y_i$ for all $i \in [n]$, we define the n-box $\mathcal{B}_{xy} := \Big[x_1, y_1\Big] \times \Big[x_2, y_2\Big] \times \cdots \times \Big[x_n, y_n\Big]$. The set of vertices of $\mathcal{B}_{xy}$ is $V_{xy} = \Big\{x_1, y_1\Big\} \times \Big\{x_2, y_2\Big\} \times \cdots \times \Big\{x_n, y_n\Big\}$. The sign of vertex $\mathbf{b} \in V_{xy}$ is defined by

$$\operatorname{sgn}(\mathbf{b}) := \begin{cases} 1, & \text{if } b_i = x_i \text{ for an even number of entries } i \\ -1, & \text{if } b_i = x_i \text{ for an odd number of entries } i. \end{cases}$$

Given a set $D \subseteq \mathbb{R}$, define $\overline{D} := \Big(D \cup \{0\} \cup \{\infty\}\Big)$. A function $G : \overline{\mathbb{R}}^n \to \mathbb{R}$ is called n-increasing on $\overline{D}^n$ when

$$\sum_{\mathbf{b} \in V_{xy}} \operatorname{sgn}(\mathbf{b}) G(\mathbf{b}) \geq 0, \text{ for all } \mathbf{x} \leq \mathbf{y},\ \mathbf{x}, \mathbf{y} \in \overline{D}^n \tag{6.20}$$

Remark 6.16 (Chapter 2.1 in Nelsen [155])**.** *For $n > 1$, the fact that G is n-increasing does not necessarily imply that G is non-decreasing in each argument, and the other way round.*

The following family of functions captures the concept of CDF.

Definition 6.17. *Let $\mathcal{S} \subseteq \mathbb{R}_{++}$. $\mathcal{G}_n(\mathcal{S})$ is the family of functions $G : \overline{\mathcal{S}}^n \to [0, 1]$ satisfying the conditions below.*

1. *G is right continuous on $\overline{\mathcal{S}}^n$*
2. *G is n-increasing on $\overline{\mathcal{S}}^n$*
3. *$G(\mathbf{z}) = 0$ for all $\mathbf{z}$ in $\overline{\mathcal{S}}^n$ such that at least one of $z_i = 0$*
4. *$G(\infty, \ldots, \infty) = 1$*

Theorem 6.18 (Definition 2.10.8. in Nelsen [155])**.** *A function $G : \overline{\mathbb{R}}^n_{++} \to [0, 1]$ is a CDF of some $\mathbb{P} \in \mathcal{P}_n$ if and only if $G \in \mathcal{G}_n(\mathbb{R}_{++})$.*

6.4.2 Formulation of upper and lower bounds

To construct the upper bound, we restrict the inner maximization in problem (6.19) *to a subset of* $\mathbb{R}_{++}$. *We do this using the next lemma.*

Lemma 6.19. *Let $\mathcal{S} \subseteq \mathbb{R}_{++}$ be a **finite** set. Then $g \in \mathcal{G}_n(\mathcal{S})$ if and only if there exists $G \in \mathcal{G}_n(\mathbb{R}_{++})$ such that $g = G|_{\overline{\mathcal{S}}^n}$. That is, g is a restriction of G to $\overline{\mathcal{S}}^n$.*

Proof. If there is $G \in \mathcal{G}_n(\mathbb{R}_{++})$ such that $g = G|_{\overline{\mathcal{S}}^n}$, then $g \in \mathcal{G}_n(\mathcal{S})$ by definition of $\mathcal{G}_n(\mathbb{R}_{++})$. On the other hand, let $g \in \mathcal{G}_n(\mathcal{S})$ and consider a number $a > \max\{s : s \in \mathcal{S}\}$. Let $\mathcal{S}_a = \mathcal{S} \cup \{a\}$ and define a new function $\hat{g} : \overline{\mathcal{S}_a}^n \to [0, 1]$ such that $\hat{g}(\mathbf{z}) = g(\mathbf{z})$ for $\mathbf{z} \in \overline{\mathcal{S}}^n$. For $\mathbf{z} \notin \overline{\mathcal{S}}^n$, construct a new vector $\mathbf{y}$ by replacing all occurrences of

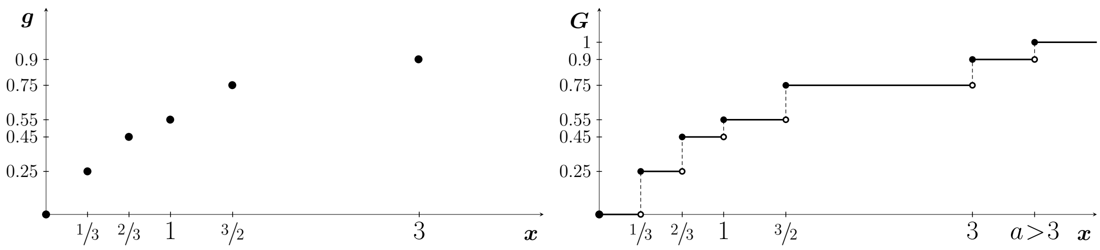
a in $\mathbf{z}$ with ∞ and define $\hat{g}(\mathbf{z}) = g(\mathbf{y})$. Consider the following piecewise constant function:

$$G(z_1, \ldots, z_n) := \hat{g}\Big(\max_{x \in \overline{\mathcal{S}}_a}\{x : x \le z_1\}, \ldots, \max_{x \in \overline{\mathcal{S}}_a}\{x : x \le z_n\}\Big). \tag{6.21}$$

It is straightforward to show that $G \in \mathcal{G}_n(\mathbb{R}_{++})$ and $g = G|_{\overline{\mathcal{S}}^n}$. See Figure 6.1 for an illustration of the case $n = 1$. $\square$

Remark 6.20. *The choice of a in the proof of Lemma 6.19 is free and might influence our upper bound computations in Section 6.4.3.*

Figure 6.1 – $\mathcal{S} = \{\frac{1}{3}, \frac{2}{3}, 1, \frac{3}{2}, 3\}$, $n = 1$. The left plot: a function $g \in \mathcal{G}_n(\mathcal{S})$.
The right plot: a function $G \in \mathcal{G}_n(\mathbb{R}_{++})$.
Notice that G is a CDF and g is the restriction of G to $\overline{\mathcal{S}}^n$.



For a finite $\mathcal{S} \subset \mathbb{R}_{++}$ and $g \in \mathcal{G}_n(\mathcal{S})$, we define the restriction of the objective in problem (6.19) to $\mathcal{S}$:

$$\begin{aligned}\phi^g(x,y) = 1 + y - \min\Big\{1,\ 1 - 1/x + y\Big\} g(x, \infty, \ldots, \infty) - y g(y, \infty, \ldots, \infty) \\ + \min\Big\{1 + 1/x,\ 1 + y\Big\} g(x, y, \infty, \ldots, \infty) \quad \text{for all } x, y \in \overline{\mathcal{S}}.\end{aligned} \tag{6.22}$$

By Lemma 6.19, $\phi^g = \phi^{\mathbb{P}}|_{\overline{\mathcal{S}}^2}$ for some $\mathbb{P} \in \mathcal{P}_n$.

Theorem 6.21. *Given a number of tasks n, for any $\mathbb{P} \in \mathcal{C}_n$ and finite $\mathcal{S} \subset \mathbb{R}_{++}$, we have*

$$R_n(\mathbb{P}) \ge R_n \ge R_n(\mathcal{S}) := \inf_{g \in \mathcal{G}_n(\mathcal{S})} \sup_{x,y \in \mathcal{S}} \Big\{\phi^g(x,y) : g(\mathbf{z}) = g(\mathbf{z}\pi) \textit{ for all } \pi \in Sym(n), \mathbf{z} \in \overline{\mathcal{S}}^n\Big\} \tag{6.23}$$

Proof. The first inequality follows immediately from Corollary 6.14. Now, we prove the second inequality. Every $\mathbb{P} \in \mathcal{C}_n$ has a CDF $G_{\mathbb{P}} \in \mathcal{G}_n(\mathbb{R}_{++})$ invariant under permutations of the variables by Proposition 6.10. At the same time, every such invariant $G \in \mathcal{G}_n(\mathbb{R}_{++})$ corresponds to some $\mathbb{P}_G \in \mathcal{C}_n$. Combining this with Corollary 6.14, we obtain

$$\begin{aligned} R_n &= \inf_{G \in \mathcal{G}_n(\mathbb{R}_{++})} \sup_{x,y \in \mathbb{R}_{++}} \{\phi^{\mathbb{P}_G}(x,y) : G(\mathbf{z}) = G(\mathbf{z}\pi) \text{ for all } \pi \in Sym(n),\ \mathbf{z} \in \overline{\mathbb{R}}^n_{++}\} \\ &\geq \inf_{G \in \mathcal{G}_n(\mathbb{R}_{++})} \sup_{x,y \in \mathcal{S}} \{\phi^{\mathbb{P}_G}(x,y) : G(\mathbf{z}) = G(\mathbf{z}\pi) \text{ for all } \pi \in Sym(n),\ \mathbf{z} \in \overline{\mathbb{R}}^n_{++}\} \\ &= \inf_{g \in \mathcal{G}_n(\mathcal{S})} \sup_{x,y \in \mathcal{S}} \{\phi^{g}(x,y) : g(\mathbf{z}) = g(\mathbf{z}\pi) \text{ for all } \pi \in Sym(n),\ \mathbf{z} \in \overline{\mathcal{S}}^n\}. \end{aligned}$$

The last equality holds by Lemma 6.19. Notice that if $G \in \mathcal{G}_n(\mathbb{R}_{++})$ is invariant under permutations of the variables, then so is the $g := G|_{\overline{\mathcal{S}}^n}$. On the other hand, if $g \in \mathcal{G}_n(\mathcal{S})$ is invariant under permutations of the variables, then so is the G defined in (6.21).

□

6.4.3 The implementation and numerical results

Let $S \subset \mathbb{R}_{++}$. To compute the lower bound $R_n(\mathcal{S})$ from (6.23), we use the epigraph form of the optimization problem for $R_n(\mathcal{S})$:

$$\begin{aligned} R_n(\mathcal{S}) = \inf_{g \in \mathcal{G}_n(\mathcal{S}),\ t \in \mathbb{R}} \quad & t \\ \text{s.t.} \quad & \phi^g(x,y) \leq t \qquad \text{for all } x,y \in \mathcal{S} \\ & g(\mathbf{z}) = g(\mathbf{z}\pi) \qquad \text{for all } \pi \in Sym(n),\ \mathbf{z} \in \overline{\mathcal{S}}^n \end{aligned} \tag{6.24}$$

The optimization variable in the problem above is g. This variable is a vector in $\mathbb{R}^{(|\mathcal{S}|+2)^n}$ which represents a function $g \in \mathcal{G}_n(\mathcal{S})$. We slightly abuse the notation and do not use a bold symbol for g to underline that g corresponds to a function with a finite support. Family $\mathcal{G}_n(\mathcal{S})$ is an infinite family of functions g, and each of these functions is defined on a finite set $\overline{\mathcal{S}}^n$ with cardinality $(|\mathcal{S}|+2)^n$. For the purpose of optimization, this means that we consider all vectors $g \in \mathbb{R}^{(|\mathcal{S}|+2)^n}$ which satisfy the four conditions in Definition 6.17 and the invariance property. All mentioned conditions are linear, and there are finitely many of them. Therefore the optimization over the infinite family of functions $\mathcal{G}_n(\mathcal{S})$ can be written as a finite LP. We use the invariance of g (the second constraint) and Conditions 3-4 in Definition 6.17 to reduce the number of variables in problem (6.24) (the size of g as a vector). To ensure that g corresponds to an n-increasing function as specified in (6.20), it is enough to consider $\mathbf{x}, \mathbf{y} \in \overline{\mathcal{S}}^n$ such that x_i, y_i are consecutive points in $\mathcal{S}$ for all $i \in [n]$. This reduces the number of constraints in problem (6.24).

To compute the upper bound $R_n(\mathbb{P})$ using formulation (6.23), we first construct a good-guess distribution $\mathbb{P}$. Given a set $\mathcal{S}$ and the solution g to the lower bound problem (6.23) on $\mathcal{S}$, we use the distribution $\mathbb{P}_g$ which corresponds to the CDF

(6.21) based on g. To construct this CDF, we choose a number $a > \max\{s : s \in \mathcal{S}\}$, as explained in the proof of Lemma 6.19. In the rest of this section we work with $\mathcal{S}_a = \mathcal{S} \cup \{a\}$. To solve (6.23) for $\mathbb{P}_g$, we define the following set of intervals:

$$\mathcal{I}_{\mathcal{S}} = \Big\{I_1, \ldots, I_{|\mathcal{S}|+2}\Big\} = \Big\{[0, s_1), \ldots, [s_{|\mathcal{S}|}, a), [a, \infty)\Big\},$$

This set of intervals covers $\mathbb{R}_+$, therefore by (6.23)

$$R_n(\mathbb{P}_g) = \sup_{x,y\in\mathbb{R}_{++}} \phi^{\mathbb{P}_g}(x, y) = \max_{I_i,I_j\in\mathcal{I}_{\mathcal{S}}} \left\{ \sup_{x\in I_i, y\in I_j} \phi^{\mathbb{P}_g}(x, y) \right\}. \tag{6.25}$$

We solve the inner maximization problem in (6.25) for each pair $i, j \in [|\mathcal{S}| + 2]$. The expression for $\phi^{\mathbb{P}_g}$ (6.18) for the case $xy \geq 1$ is different from the case $xy < 1$. To simplify the computations when the line $xy = 1$ crosses the rectangle $I_i \times I_j$, we restrict our attention to $\mathcal{S}$ of a particular type. Consider a collection of $k-1$ positive real numbers $r_1 < r_2 < \cdots < r_{k-1} < 1$, let

$$\mathcal{S}_k = \{s_1,\ s_2, \ldots, s_{2k-1}\} = \Big\{r_1,\ r_2, \ldots, r_{k-1},\ 1,\ \tfrac{1}{r_{k-1}}, \ldots, \tfrac{1}{r_2},\ \tfrac{1}{r_1}\Big\}. \tag{6.26}$$

For any $a > \frac{1}{r_1}$, $\mathcal{S}_{k,a} = \mathcal{S}_k \cup \{a\}$ subdivides $\mathbb{R}_+$ into $2k+1$ intervals $\mathcal{I}_{\mathcal{S}_k}$. First, consider a pair of intervals $I_i, I_j \in \mathcal{I}_{\mathcal{S}_k}$ such that $i \notin \{1, 2k+1\}$ and $j \notin \{1, 2k+1\}$ (i.e., neither of them are the first or the last interval). Due to the choice of $\mathcal{S}_k$, the line $xy = 1$ crosses the rectangle $I_i \times I_j$ if and only if $i+j = 2k+1$. Denote the bivariate marginal CDF and the univariate marginal CDF of $\mathbb{P}_g$ by H_g and F_g, respectively. Then

$$\phi^{\mathbb{P}_g}(x, y) = \begin{cases} 1+y-F_g(x)-yF_g(y)+\Big(1+{}^1\!/\!_x\Big)H_g(x, y) & i+j\geq 2k+1,\ xy\geq 1 \\ 1+y\Big(1-F_g(x)-F_g(y)+H_g(x, y)\Big) & \\ \quad -\Big(1-{}^1\!/\!_x\Big)F_g(x)+H_g(x, y) & i+j\leq 2k+1,\ xy<1. \end{cases} \tag{6.27}$$

We construct the CDF of $\mathbb{P}_g$ using (6.21), therefore

$$\begin{aligned} H_g(x, y) &= g(s_{i-1}, s_{j-1}, \infty, \ldots, \infty) \ \text{ for all } (x, y) \in I_i \times I_j \\ F_g(x) &= g(s_{i-1}, \infty, \ldots, \infty) \ \text{ for all } x \in I_i. \end{aligned} \tag{6.28}$$

That is, the marginal CDFs are constant on $I_i \times I_j$ and I_i, respectively. As the range of a CDF is $[0, 1]$, we conclude that for $x \in I_i, y \in I_j$, $\phi^{\mathbb{P}_g}(x, y)$ is non-increasing in x and non-decreasing in y. The latter holds since for any $\mathbb{P} \in \mathcal{P}_2$ invariant under S_2 and for any $x, y \in \mathbb{R}$,

$$H_{\mathbb{P}}(x, y) = F_{\mathbb{P}}(x) + F_{\mathbb{P}}(y) - 1 + \mathbf{P}_{\mathbf{z}\sim\mathbb{P}}[z_1 > x, z_2 > y] \geq \max\Big\{0, F_{\mathbb{P}}(x) + F_{\mathbb{P}}(y) - 1\Big\}. \tag{6.29}$$

Hence the optimal value of the inner maximization problem in (6.25) can be obtained by first substituting the CDFs (6.28) into the function (6.27) and then substituting

$x = s_{i-1}$, $y = s_j$. Note that this optimum is not attained. For the case $i+j = 2k+1$, i.e., when the line $xy = 1$ crosses the rectangle $I_i \times I_j$, the result holds due to the choice of $\mathcal{S}_k$.

When $i \in \{1, 2k+1\}$ or $j \in \{1, 2k+1\}$, the function $\phi^{\mathbb{P}_g}(x, y)$ simplifies, and we solve such cases separately. The solution approach resembles the one from the previous paragraph.

In numerical experiments we use uniform finite sets

$$\mathcal{S}_k^u = \left\{\tfrac{1}{k}, \tfrac{2}{k}, \dots, \tfrac{k-1}{k},\ 1,\ \tfrac{k}{k-1}, \dots, k\right\}. \tag{6.30}$$

Table 6.2 shows the best obtained bounds and the k we use to compute these bounds.

Table 6.2 – Bounds for the case of two machines and $2, 3, 4$ tasks based on Theorem 6.21

n	Lower bound		Upper bound		k
	Current	New	Current	New	
2	1.505949 [130]	1.505980	1.5068 [32]	1.5093	250
3		1.5076	1.5861 [32]	1.5238	50
4		1.5195		1.5628	20

We round the lower bounds $R_n(\mathcal{S}_k^u)$ down and the upper bounds $R_n(\mathbb{P}_g)$ up. We verify all upper bounds with exact arithmetics using the MATLAB symbolic package and the following procedure. First, we obtain the optimal solution g to problem (6.24) and round the elements of the set $\mathcal{S}_k^u$ and the number a to the 8^{th} digit. Next, we transform the rounded values into rational numbers and compute $R_n(\mathbb{P}_g)$ as a rational number. By Lemma 6.19 and Theorem 6.21, the rounded g provides the algorithm $\mathcal{A}^{\mathbb{P}_g}$ with the worst-case approximation ratio $R_n(\mathbb{P}_g)$.

The upper bound for $n = 2$ in Table 6.2 is worse than the best existing upper bound. We improve our result and obtain a new best exiting upper bound for $n = 2$ in the next section.

6.5 More precise bounds for two tasks

In this section we analyze the case with $n = 2$ tasks and $m = 2$ machines in more detail. Now, to obtain an upper bound, we do not simply use some good-guess distribution as we did before, but we optimize over a subset of $\mathcal{C}_n$ (6.16). Moreover, as a side result of this optimization, we obtain a non-uniform set $\mathcal{S}_k$ which produces a better lower bound than the one from Table 6.2 in the previous section.

Problem (6.19) simplifies for $n = 2$. Given $F \in \mathcal{G}_1(\mathbb{R}_{++})$, define

$$H(x,y) := \max\Big\{0,\ F(x) + F(y) - 1\Big\}. \tag{6.31}$$

H is a copula, i.e., there is $\mathbb{P}_{H,F} \in \mathcal{P}_2$ for which H is its CDF and F is its marginal CDF. See Nelsen [155] for the detailed description of copulas and their properties. Moreover, by construction $\mathbb{P}_{H,F} \in \mathcal{C}_2$. Therefore we can rewrite problem (6.19) using univariate marginal CDF's.

Theorem 6.22. *Consider family $\mathcal{G}_1(\mathbb{R}_{++})$ from Definition 6.17.*

$$\begin{aligned} R_2 = \inf_{F\in\mathcal{G}_1(\mathbb{R}_{++})} \sup_{x,y\in\mathbb{R}_{++}} & \ 1 + y - \min\Big\{1,\ 1 - {}^1\!/\!_x + y\Big\}F(x) - yF(y) \\ & + \min\Big\{1 + {}^1\!/\!_x,\ 1+y\Big\}\max\Big\{0,\ F(x)+F(y)-1\Big\}. \end{aligned} \tag{6.32}$$

Proof. Given $F \in \mathcal{G}_1(\mathbb{R}_{++})$, let

$$\begin{aligned} \phi^F(x,y) = 1 + y - \min\Big\{1,\ 1 - {}^1\!/\!_x + y\Big\}F(x) - yF(y) \\ + \min\Big\{1 + {}^1\!/\!_x,\ 1+y\Big\}\max\Big\{0,\ F(x)+F(y)-1\Big\}. \end{aligned}$$

Consider any $\mathbb{P} \in \mathcal{C}_2$ with the univariate CDF $F_{\mathbb{P}}$. Define $\phi^{\mathbb{P}}(x,y)$ as in (6.18). From (6.29) for all $x, y \in \mathbb{R}_{++}$,

$$\phi^{\mathbb{P}}(x,y) \geq \phi^{F_{\mathbb{P}}}(x,y).$$

Thus

$$R_2 \overset{(6.19)}{=} \inf_{\mathbb{P}\in\mathcal{C}_2} \sup_{x,y\in\mathbb{R}_{++}} \phi^{\mathbb{P}}(x,y) \geq \inf_{\mathbb{P}\in\mathcal{C}_2} \sup_{x,y\in\mathbb{R}_{++}} \phi^{F_{\mathbb{P}}}(x,y) \geq \inf_{F\in\mathcal{G}_1(\mathbb{R}_{++})} \sup_{x,y\in\mathbb{R}_{++}} \phi^F(x,y).$$

On the other hand, for all $F \in \mathcal{G}_1(\mathbb{R}_{++})$ there is copula H from (6.31) with the corresponding distribution $\mathbb{P}_{H,F} \in \mathcal{C}_2$. Hence

$$\begin{aligned} R_2 = \inf_{\mathbb{P}\in\mathcal{C}_2} \sup_{x,y\in\mathbb{R}_{++}} \phi^{\mathbb{P}}(x,y) &\leq \inf_{F\in\mathcal{G}_1(\mathbb{R}_{++})} \sup_{x,y\in\mathbb{R}_{++}} \phi^{\mathbb{P}_{H,F}}(x,y) \\ &= \inf_{F\in\mathcal{G}_1(\mathbb{R}_{++})} \sup_{x,y\in\mathbb{R}_{++}} \phi^F(x,y). \end{aligned}$$

□

Remark 6.23. *Nelsen [155] shows that for $n > 2$ the function*

$$G(x_1,\ldots,x_n) = \max\Big\{0,\ \sum_{i=1}^n F(x_i) - n + 1\Big\}$$

is not a CDF. We do not see other suitable n-variate CDFs which would have a bivariate margin H from (6.31). *As a result, the proof of Theorem 6.22 fails for $n > 2$.*

6.5.1 New upper bound for two tasks

To compute a new upper bound on R_2, we restrict the set of functions in problem (6.32) to the family of piecewise rational univariate CDFs. We say that a function is piecewise rational if it can be written as a fraction where both the numerator and the denominator are polynomials. The domain of each CDF is subdivided into pieces by $\mathcal{S}_k$ defined in (6.26). Given $\mathcal{S}_k$, we introduce a collection of intervals:

$$\begin{aligned}\mathcal{I}_{\mathcal{S}_k} &= \Big\{I_1, I_2, \dots, I_k, I_{k+1}, \dots, I_{2k-1}, I_{2k}\Big\}\\ &= \Big\{[0, r_1), [r_1, r_2), \dots, [r_{k-1}, 1), \Big[1, {}^1\!/\!_{r_{k-1}}\Big), \dots, \Big[{}^1\!/\!_{r_2}, {}^1\!/\!_{r_1}\Big), \Big[{}^1\!/\!_{r_1}, +\infty\Big)\Big\}. \qquad (6.33)\end{aligned}$$

Remark 6.24. *We build the intervals using the points from $\mathcal{S}_k$ only. This is different from Section 6.4.3 where we use an additional number $a > \max \mathcal{S}_k$ to construct the intervals.*

Given $\mathcal{I}_{\mathcal{S}_k}$ and a family of continuous functions $\mathcal{F}$, we consider CDFs which "piecewisely" belong to $\mathcal{F}$.

Definition 6.25. *$\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)$ is a family of functions $F : \mathbb{R}_{++} \to [0, 1]$ such that*

$$F(x) = \begin{cases} f_1(1/x) & x \in I_1\\ f_2(1/x) & x \in I_2\\ \dots & \\ f_k(1/x) & x \in I_k\\ 1 - f_k(x) & x \in I_{k+1}\\ \dots & \\ 1 - f_2(x) & x \in I_{2k-1}\\ 1 - f_1(x) & x \in I_{2k}, \end{cases} \qquad (6.34)$$

$f_1(x) = 0$, $f_i(x) \in \mathcal{F}$, $f_i(x)$ is non-decreasing, $f_k(1) \le 0.5$, $0 \le f_i(x_i) \le f_{i+1}(x_i)$ for all $i < k$. By construction, F is a CDF and thus $\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k) \subset \mathcal{G}_1(\mathbb{R}_{++})$. Restricting the minimization in problem (6.32) to $\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)$ provides an upper bound on R_2. We use the symmetry of F to simplify the expression for this bound.

Proposition 6.26. *Define $\mathcal{S}_k$ as in (6.26) and consider family $\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)$ from Definition 6.25. R_2 is not larger than*

$$\begin{aligned}R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big) &= \inf_{F \in \mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)} \sup_{x,y \in \mathbb{R}_{++}} 1 + y - \min\Big\{1,\ 1 - 1/x + y\Big\}F(x) - yF(y) \qquad (6.35)\\ &\qquad + \min\Big\{1 + 1/x,\ 1 + y\Big\} \max\Big\{0,\ F(x) + F(y) - 1\Big\}\\ &= \inf_{F \in \mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)} \sup_{x,y \in \mathbb{R}_{++},\ xy \ge 1} y - 1/x + \Big(1 + 1/x - y\Big)F(y) + 1/x F(x). \qquad (6.36)\end{aligned}$$

Proof. Let $F \in \mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)$. Problem (6.35) is a restriction of problem (6.32) to a smaller set of functions. Therefore the former problem defines an upper bound on R_2.

Next, we show the equality between (6.35) and (6.36). Denote by ϕ^F the objective of problem (6.35). We claim that for every $x, y \in \mathbb{R}_{++}$,

$$\phi^F(x,y) \le \sup_{x,y\in\mathbb{R}_{++},\ xy\ge 1} \phi^F(x,y).$$

Let $\hat{x} > 0, \hat{y} > 0$. First, consider the case $\hat{x}\hat{y} \ge 1$. By construction of (6.34), $\hat{x}\hat{y} \ge 1$ implies $F(\hat{x}) + F(\hat{y}) \ge F(\hat{x}) + F\left(1/\hat{x}\right) \ge 1$. Then

$$\begin{aligned}\phi^F(\hat{x},\hat{y}) &= 1 + \hat{y} - F(\hat{x}) - \hat{y}F(\hat{y}) + \left(1 + 1/\hat{x}\right)\left(F(\hat{x}) + F(\hat{y}) - 1\right)\\ &= \hat{y} - 1/\hat{x} + \left(1 + 1/\hat{x} - \hat{y}\right)F(\hat{y}) + 1/\hat{x}F(\hat{x}).\end{aligned} \tag{6.37}$$

Now, let $\hat{x}\hat{y} < 1$. By construction of $\mathcal{I}_{\mathcal{S}_k}$ in (6.33), there are $I_i, I_j \in \mathcal{I}_{\mathcal{S}_k}$ such that $\hat{x} \in I_i,\ \hat{y} \in I_j$. The set $\mathcal{S}_k$ is finite, therefore there is a sequence $\{(x_t,y_t)\}_{t=1}^{\infty}$ such that for all t the following holds: $x_t \in I_i,\ y_t \in I_j,\ x_t \to \hat{x}^+,\ y_t \to \hat{y}^+,\ x_t y_t < 1$ and $x_t, y_t \notin \mathcal{S}_k$. For all $x \in \mathbb{R}_{++} \setminus \mathcal{S}_k$ we have $F(x) + F\left(1/x\right) = 1$. Hence for all t

$$\begin{aligned}\phi^F(x_t,y_t) &= 1 + y_t - \left(1 - 1/x_t + y_t\right)F(x_t) - y_tF(y_t)\\ &= y_t\left(1 - F(y_t)\right) + \left(1 - 1/x_t + y_t\right)\left(1 - F(x_t)\right) + 1/x_t - y_t\\ &= 1/x_t - y_t + \left(1 - 1/x_t + y_t\right)F\left(1/x_t\right) + y_tF\left(1/y_t\right)\\ &\overset{(6.37)}{=} \phi^F\left(1/y_t, 1/x_t\right).\end{aligned} \tag{6.38}$$

Finally, F is right continuous in $(\hat{x},\hat{y})$, and so is $\phi^F(\hat{x},\hat{y})$. Since $x_t \to \hat{x}^+, y_t \to \hat{y}^+$,

$$\phi^F(\hat{x},\hat{y}) = \lim_{t\to\infty}\phi^F(x_t,y_t) \overset{(6.38)}{=} \lim_{t\to\infty}\phi^F\left(1/y_t, 1/x_t\right) \le \sup_{x,y\in\mathbb{R}_{++},\ xy>1}\phi^F\left(x,y\right).$$

The last inequality follows from $\left(1/y_t\right)\left(1/x_t\right) > 1$. □

6.5.2 Implementing the new upper bound for two tasks

In this subsection, for a given $\mathcal{S}_k$, we choose $\mathcal{F}$ in Definition 6.25 to be the family of linear univariate functions

$$\mathcal{F} := \{c^0 + c^1x:\ c^0, c^1 \in \mathbb{R}\}. \tag{6.39}$$

Then $\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)$ includes, in particular, the CDFs from [32, 130, 131, 163] where the authors use piecewise functions with domains subdivided into 2, 4 or 6 intervals. We observe that the upper bounds are better when the domains are subdivided more times or when each piece has a more complex form than a constant function, i.e., when c^1 can be non-zero. We improve upon the existing upper bounds by using a larger number of pieces and letting c^1 be non-zero in each piece. Define

$$\phi^F(x,y) := y - 1/x + \left(1 + 1/x - y\right)F(y) + 1/xF(x). \tag{6.40}$$

Let $\mathcal{X} := \Big\{(x, y) \in \mathbb{R}^2_{++}\colon xy \geq 1\Big\}$. Consider two formulations for $R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big)$ which are equivalent to problem (6.36)

$$R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big) = \inf_{F\in\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k),\ t} \quad t \tag{6.41}$$

$$\text{s.t.} \quad \phi^F(x, y) \leq t,\ \text{for all } (x, y) \in \mathcal{X}$$

$$= \inf_{F\in\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k),\ t} \quad t \tag{6.42}$$

$$\text{s.t.} \quad \sup_{x\in I_i, y\in I_j} \phi^F(x, y) \leq t,\ \text{for all } i + j \geq 2k + 1.$$

The second equality follows from the equivalence of problems (6.36) and (6.35) since for $\mathcal{S}_k$ defined in (6.26) $xy \geq 1$ implies $x \in I_i,\ y \in I_j$ with $i + j \geq 2k + 1$. We use problems (6.41) and (6.42) to approximate $R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big)$ with high precision. Namely, we use relaxations of problem (6.41) to compute lower bounds on $R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big)$, and we use feasible solutions to problem (6.42) to compute upper bounds on $R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big)$.

For $F \in \mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)$ satisfying (6.34) and (6.39), the variables in problem (6.41) are $\Big(t, \{c_i^0\}_{i=1}^k, \{c_i^1\}_{i=1}^k\Big)$. This problem is LP with infinitely many constraints: each $(x, y) \in \mathcal{X}$ induces a linear constraint. Such problems can be well approximated using the cutting-plane approach introduced by Kelley [100]. Namely, we start with a finite set $\mathcal{Y} \subset \mathcal{X}$ and restrict the set of constraints in (6.41) to its finite subset generated by $(x, y) \in \mathcal{Y}$. As a result, we obtain a finite linear problem. Denote its optimal solution by $(\underline{F},\ \underline{t})$. Then $\underline{t}$ is a lower bound on $R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big)$.

Next, we substitute $\underline{F}$ in (6.42) and find a feasible t. We compute the supremum for each pair $i, j \in [2k]$ with $i + j \geq 2k + 1$ using the Karush-Kuhn-Tucker (KKT) conditions. For each pair of intervals the inner maximization problem in (6.42) either is linearly constrained or satisfies the Mangasarian-Fromovitz constraint qualification. Therefore the optimum is among the KKT points, see, e.g., Section 3 in Eustaquio et al. [63] for more details. All problems are simple and have similar structure. Therefore we do not write the KKT conditions explicitly, but consider all possible critical points from the first order conditions and from the boundary conditions. This set contains all the KKT points, and thus the optimal one. At the end we choose the point (x, y) with the maximal value of $\phi^{\underline{F}}(x, y)$ among the critical points. We describe the procedure of computing the critical points for function (6.40) at the end of this subsection. Let $\bar{t}$ be the maximum of $\phi^{\underline{F}}(x, y)$ over all pairs of intervals. The solution $(\underline{F}, \bar{t})$ is feasible for problem (6.42). Thus $\bar{t}$ is an upper bound on $R_2\Big(\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k)\Big)$. Let (x^*, y^*) be a point such that $\phi^{\underline{F}}(x^*, y^*) = \bar{t}$. If $|\bar{t} - \underline{t}| > 10^{-8}$, we proceed from the beginning by restricting problem (6.41) to the updated set $\mathcal{Y} \leftarrow \mathcal{Y} \cup \{(x^*, y^*)\}$. Otherwise we stop.

To obtain numerical results, we use uniform sets $\mathcal{S}_k^u$ of the form (6.30). We work with family $\mathcal{C}_{\mathcal{F}}(\mathcal{S}_k^u)$ from Definition 6.25, where the underlying family of functions $\mathcal{F}$ is defined in (6.39). We initialize the cutting-plane procedure with $\mathcal{Y} = \Big\{(x, y) : x, y \in \mathcal{S}_k^u,\ xy \geq 1\Big\}$. The best obtained upper bound is indicated in bold in Table 6.3, it is stronger than the currently best upper bound 1.5068.

Table 6.3 – Upper bounds on the best approximation ratio for 2 tasks

k	5	10	16	50	100
Upper bound on R_2	1.5174	1.5096	1.5066	1.5060	**1.5059964**

We verify the upper bound 1.5059964 using exact arithmetics in a similar way as we do it for the upper bounds in Table 6.2 of Section 6.4.3.

Possible critical points computation for function (6.40).

Next, we show how to find possible critical points for the function

$$\phi^F(x, y) = y - {}^1\!/\!_x + \Big(1 + {}^1\!/\!_x - y\Big)F(y) + {}^1\!/\!_x F(x),$$

with F defined in (6.34) using (6.39) on $I_i \times I_j$ such that $i, j \in [2k],\ i + j \geq 2k + 1$. By construction, in each interval F is differentiable, and ϕ^F is differentiable on $I_i \times I_j$ (the expression simplifies for $i = 1$ and $j = 2k$). The first derivatives of the function $\phi^F(x, y)$ are:

$$\frac{\partial \phi^F(x, y)}{\partial x} = \frac{1}{x^2}\left(1 - F(x) - F(y) + \frac{\partial F(x)}{\partial x}x\right) \quad \text{and}$$
$$\frac{\partial \phi^F(x, y)}{\partial y} = 1 - F(y) + \Big(1 + {}^1\!/\!_x - y\Big)\frac{\partial F(y)}{\partial y}.$$

Next, we consider the three possible cases for i, j. For each of these cases we substitute $F(x)$, $F(y)$ from (6.34) and find the analytical solution to the system $\frac{\partial \phi^F(x,y)}{\partial x} = 0$, $\frac{\partial \phi^F(x,y)}{\partial y} = 0$. For this purpose we use Wolfram|Alpha [223]. The obtained solution is denoted by $(x^*,\ y^*)$.

Case 1. $k < i \leq 2k,\ 1 \leq j \leq k.$

In this case $F(x) = 1 - c_i^0 - c_i^1 x,\ F(y) = c_j^0 + {}^{c_j^1}\!/\!_y$ and $y \in (0, 1)$. Hence

$$\frac{\partial \phi^F(x, y)}{\partial x} = \frac{1}{x^2}\left(c_i^0 - c_j^0 - \frac{c_j^1}{y}\right), \quad \frac{\partial \phi^F(x, y)}{\partial y} \geq 0.$$

The latter holds since $F(y) \leq 1$, $F(y)$ is non-decreasing by construction, and $y \in (0, 1)$. The sign of the derivative with respect to x does not depend on x. The function $\phi^F(x, y)$ is non-decreasing in y and is either non-increasing or non-decreasing in

x. We not know this in advance, so we use the set $\Big\{(s_{i-1}, s_j), (s_i, s_j)\Big\}$ as possible critical points.

Case 2. $k < i \leq 2k,\ k < j \leq 2k.$

In this case $F(x) = 1 - c_i^0 - c_i^1 x,\ F(y) = 1 - c_j^0 - c_j^1 y$ and $y \geq 1$. Hence

$$\frac{\partial \phi^F(x,y)}{\partial x} = {}^1\!/_{x^2}\Big(c_i^0 - 1 + c_j^0 + c_j^1 y\Big), \quad \frac{\partial \phi^F(x,y)}{\partial y} = c_j^0 - \Big(1 + {}^1\!/_x - 2y\Big)c_j^1.$$

As in ***Case 1***, the sign of the derivative with respect to x does not depend on x, hence $\phi^F(x,y)$ is non-increasing or non-decreasing in x. We do not know this in advance, so we start with the set $\{s_{i-1}, s_i\} \times \{s_{j-1}, s_j, y^*\}$ as possible critical points. If $i+j = 2k+1$ (that is, the line $xy = 1$ crosses the rectangle $I_i \times I_j$), we additionally consider the point $(\frac{1}{y^*}, y^*)$. We check all pairs for feasibility and exclude the infeasible ones.

Case 3. $1 \leq i \leq k,\ k < j \leq 2k.$

In this case $F(x) = c_i^0 + {}^{c_i^1}\!/_x,\ F(y) = 1 - c_j^0 - c_j^1 y$, and

$$\frac{\partial \phi^F(x,y)}{\partial x} = \frac{1}{x^2}\left(- c_i^0 + c_j^0 + c_j^1 y - \frac{2c_i^1}{x}\right), \quad \frac{\partial \phi^F(x,y)}{\partial y} = c_j^0 - \Big(1 + {}^1\!/_x - 2y\Big)c_j^1.$$

The sign of the derivatives is unknown, so we start with the set $\{s_{i-1}, s_i, x^*\} \times \{s_{j-1}, s_j, y^*\}$ as possible critical points. If $i + j = 2k + 1$ (that is, the line $xy = 1$ crosses the rectangle $I_i \times I_j$), we additionally consider the set $\Big\{(x^*, \frac{1}{x^*}), (\frac{1}{y^*}, y^*)\Big\}$. We check all resulting pairs for feasibility and exclude the infeasible ones.

When $x \in I_1$ or $y \in I_{2k}$, $\phi^F(x,y)$ simplifies by construction of (6.34). In our computations we analyze these situations separately.

6.5.3 New lower bound for two tasks

The cutting-plane approach from Section 6.5.1 generates not only the upper bound with the corresponding CDF, but also the set of points $\mathcal{Y}$. Using $\mathcal{Y}$, we build a new set $\mathcal{S}_k^*$ of the form (6.26), which is not uniform as in (6.30). We consider all $(x, y) \in \mathcal{Y}$ involved in the binding constraints of problem (6.41) at the last cutting-plane iteration. Next, we take the corresponding x, y and their reciprocals, order ascending, round to the 8^{th} digit and obtain $\mathcal{S}_k^*$ with $k = 82$. For this set the lower bound $R_n(\mathcal{S}_k^*)$ from problem (6.24) is 1.5059953, which is stronger than our lower bound from Table 6.2. As a result, the lower and upper bounds become very close to each other: $|R_2 - 1.505996| \leq 10^{-6}$.

6.6 Conclusion and questions for further research

We consider randomized MIS algorithms to the minimum makespan problem on two unrelated parallel selfish machines. We propose a new $Min - Max$ formulation (6.19) to find R_n, the best approximation ratio over randomized MIS algorithms. The minimization in (6.19) goes over distributions and the maximization goes over $\mathbb{R}^2_{++}$. The problem is generally intractable. Therefore we build upper and lower bounds on the optimal value. To obtain the lower bound, we solve the initial problem on a finite subset of $\mathbb{R}^2_{++}$. Using the resulting solution, we construct a piecewise constant cumulative distribution function (CDF) for which the worst-case performance is easy to estimate. In this way, we obtain the upper bound on R_n. We implement this approach and find new bound for $n \in \{2, 3, 4\}$ tasks.

For $n = 2$ the best CDF is a known function of univariate margins (copula). We parametrize these margins as piecewise rational functions of degree at most one. The resulting upper bound problem (6.36) is a linear semi-infinite problem. We solve it by the cutting-plane approach. This approach provides the upper bound 1.5059964 and the CDF for which the algorithm achieves this bound.

As a side result of the cutting-plane approach, we obtain the lower bound 1.5059953, so $|R_2 - 1.505996| \leq 10^{-6}$. This work leaves several questions for future research. First, our approach could be made more numerically efficient to provide better bounds for $m = 2$ machines. For example, column generation could solve lower bound problem (6.23) on denser grids and for the larger number of tasks n. Parametrizing distributions of more than two variables could improve the results for the upper bound problem (6.23). Second, we work with $m = 2$, machines but there are algorithms for $m > 2$ machines with similar properties, e.g., by Lu and Yu [132]. It would be interesting to see how our approach works in the case of more than two machines. Finally, the piecewise- and pointwise- constructions could be suitable for other problems with optimization over low dimensional functions, including other problems from algorithmic mechanism design.

Bibliography

[1] L. Addario-Berry, W. Kennedy, A. D. King, Z. Li, and B. Reed. Finding a maximum-weight induced $k-$partite subgraph of an $i-$triangulated graph. *Discrete Applied Mathematics*, 158(7):765 – 770, 2010. Third Workshop on Graph Classes, Optimization, and Width Parameters Eugene, Oregon, USA, October 2007.

[2] A. A. Ahmadi and G. Hall. On the construction of converging hierarchies for polynomial optimization based on certificates of global positivity. *ArXiv preprint 1709.09307*, 2018.

[3] A. A. Ahmadi and A. Majumdar. DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–5, March 2014.

[4] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5(1):13–51, 1995.

[5] G. Andrews, R. Askey, and R. Roy. *Special functions.* Cambridge University Press, 1999.

[6] M. F. Anjos and J. B. Lasserre. *Handbook on semidefinite, conic and polynomial optimization*, volume 166. Springer Science & Business Media, 2011.

[7] V. V. Arestov and A. G. Babenko. On Delsarte scheme of estimating the contact numbers. *Proc. of the Steklov Inst. of Math.*, 219:36–65, 1997.

[8] N. Arima, S. Kim, and M. Kojima. A quadratically constrained quadratic optimization model for completely positive cone programming. *SIAM J. Optim.*, 23(4): 2320–2340, 2013.

[9] V. Auletta, G. Christodoulou, and P. Penna. Mechanisms for scheduling with single-bit private values. *Theory Comput. Syst.*, 57(3):523–548, 2015.

[10] Y. Azar, M. Hoefer, I. Maor, R. Reiffenhäuser, and B. Vöcking. Truthful mechanism design via correlated tree rounding. *Math. Program.*, 163(1):445–469, 2017.

[11] C. Bachoc and F. Vallentin. Semidefinite programming bounds for spherical codes. In *2007 IEEE International Symposium on Information Theory*, pages 1801–1805, 2007.

[12] C. Bachoc and F. Vallentin. New upper bounds for kissing numbers from semidefinite programming. *J. Amer. Math. Soc.*, 21(3):909–924, 2008.

[13] C. Bachoc, G. Nebe, F. de Oliveira Filho, and F. Vallentin. Lower bounds for measurable chromatic numbers. *Geometric and Functional Analysis*, 19(3):645–661, 2009.

[14] L. Bai, J. Mitchell, and J.-S. Pang. On conic QPCCs, conic QCQPs and completely positive programs. *Math. Program.*, 159(1-2):109–136, 2016.

[15] A. Barvinok. *A course in convexity.* American Mathematical Society, 2002.

[16] M. Bentert, R. van Bevern, and R. Niedermeier. Inductive k-independent graphs and c-colorable subgraphs in scheduling: a review. *Journal of Scheduling*, Dec 2018.

[17] C. Berg and E. Porcu. From Schoenberg coefficients to Schoenberg functions. *Constructive Approximation*, 45(2):217–241, Apr 2017.

[18] R. Berr and T. Wörmann. Positive polynomials on compact sets. *manuscripta mathematica*, 104(2):135–143, Feb 2001.

[19] G. Blekherman, P. A. Parrilo, and R. R. Thomas, editors. *Semidefinite Optimization and Convex Algebraic Geometry.* SIAM, 2013.

[20] S. Bochner. Hilbert distances and positive definite functions. *Ann. of Math.*, 42: 647–656, 1941.

[21] I. Bomze and F. Jarre. A note on Burer's copositivity representation of mixed binary QPs. *Optimization Letters*, 4(3):465–472, 2010.

[22] I. M. Bomze. Copositive Optimization. In C. A. Floudas and P. M. Pardalos, editors, *Encyclopedia of Optimization, Second Edition*, pages 561–564. Springer, 2009. ISBN 978-0-387-74758-3.

[23] I. M. Bomze. Copositive optimization – recent developments and applications. *European Journal of Operational Research*, 216(3):509–520, 2012.

[24] I. M. Bomze and M. L. Overton. Narrowing the difficulty gap for the celis–dennis–tapia problem. *Mathematical Programming*, 151(2):459–476, 2015.

[25] I. M. Bomze, M. Dür, E. De Klerk, C. Roos, A. J. Quist, and T. Terlaky. On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization*, 18(4):301–320, 2000.

[26] S. Bundfuss and M. Dür. Algorithmic copositivity detection by simplicial partition. *Linear Algebra and its Applications*, 428:1511–1523, 2008.

[27] S. Bundfuss and M. Dür. An adaptive linear approximation algorithm for copositive programs. *SIAM Journal on Optimization*, 20(1):30–53, 2009.

[28] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.*, 120(2):479–495, 2009.

[29] S. Burer and H. Dong. Representing quadratically constrained quadratic programs as generalized copositive programs. *Operations Research Letters*, 40(3):203–206, 2012.

[30] M. Campêlo and R. C. Corrêa. A combined parallel lagrangian decomposition and cutting-plane generation for maximum stable set problems. *Electronic Notes in Discrete Mathematics*, 36:503–510, 2010. ISCO 2010 - International Symposium on Combinatorial Optimization.

[31] M. R. Celis, J. E. Dennis, and R. A. Tapia. A trust region strategy for nonlinear equality constrained optimization. In *Numerical Optimization*, pages 71–82. SIAM, 1985.

[32] X. Chen, D. Du, and L. Zuluaga. Copula-based randomized mechanisms for truthful scheduling on two unrelated machines. *Theory Comput. Syst.*, 57(3):753–781, 2015.

[33] G. Christodoulou, E. Koutsoupias, and A. Vidali. A lower bound for scheduling mechanisms. *Algorithmica*, 55(4):729–740, 2009.

[34] K. Corrádi and S. Szabó. A combinatorial approach for Keller's conjecture. *Periodica Mathematica Hungarica*, 21(2):95–100, Jun 1990.

[35] E. de Klerk and D. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM J. Optim.*, 12(4):875–892, 2002.

[36] E. de Klerk, D. Pasechnik, and J. Warners. On approximate graph colouring and MAX-k-CUT algorithms based on the ϑ-function. *Journal of Combinatorial Optimization*, 8(3):267–294, Sep 2004.

[37] E. de Klerk, D. Pasechnik, and A. Schrijver. Reduction of symmetric semidefinite programs using the regular $*$-representation. *Math. Program.*, 109(2):613–624, 2007.

[38] D. de Laat. Moment methods in extremal geometry. *Ph.D. thesis, Delft University of Technology*, 2016.

[39] D. de Laat and F. Vallentin. A semidefinite programming hierarchy for packing problems in discrete geometry. *Math. Program.*, 151(2):529–553, 2015.

[40] D. de Laat, F. M. de Oliveira Filho, and F. Vallentin. Upper bounds for packings of spheres of several radii. *Forum of Mathematics*, Sigma 2, e23, 2014.

[41] D. de Laat, F. C. Machado, F. M. de Oliveira Filho, and F. Vallentin. k-point semidefinite programming bounds for equiangular lines. *ArXiv preprint 1812.06045*, 2018.

[42] E. DeCorte. The eigenvalue method for extremal problems on infinite vertex-transitive graphs. *Ph.D. thesis, Delft University of Technology*, 2015.

[43] E. DeCorte, F. M. de Oliveira Filho, and F. Vallentin. Complete positivity and distance-avoiding sets, 2018.

[44] P. Delsarte, J. Goethals, and J. Seidel. Spherical codes and designs. *Geom. Dedicata*, 6:363–388, 1977.

[45] J. Demmel, J. Nie, and V. Powers. Representations of positive polynomials on non-compact semi-algebraic sets via KKT ideals. *J. Pure Appl. Algebra*, 209(1):189–200, 2007.

[46] P. Dickinson and J. Povh. On an extension of Pólya's Positivstellensatz. *Journal of Global Optimization*, 61(4):615–625, Apr 2015.

[47] P. J. Dickinson. A new certificate for copositivity. *Linear algebra and its applications*, 2019.

[48] P. J. Dickinson and J. Povh. A new approximation hierarchy for polynomial conic optimization. *Preprint*, 2018. URL http://www.optimization-online.org/DB_HTML/2013/06/3925.html.

[49] P. J. C. Dickinson. The copositive cone, the completely positive cone and their generalisations. *Ph.D. thesis, University of Groningen*, 2013.

[50] P. J. C. Dickinson and L. Gijben. On the computational complexity of membership problems for the completely positive cone and its dual. *Computational Optimization and Applications*, 57(2):403–415, Mar 2014.

[51] P. J. C. Dickinson, G. Eichfelder, and J. Povh. Erratum to: On the set-semidefinite representation of nonconvex quadratic programs over arbitrary feasible sets. *Optimization Letters*, 7(6):1387–1397, Aug 2013.

[52] Y. Ding, D.Ge, and H. Wolkowicz. On equivalence of semidefinite relaxations for quadratic matrix programming. *Mathematics of Operations Research*, 36(1):88–104, 2011.

[53] C. Dobre and J. Vera. Exploiting symmetry in copositive programs via semidefinite hierarchies. *Math. Program.*, 151(2):659–680, 2015.

[54] C. Dobre, M. Dur, L. Frerick, and F. Vallentin. A copositive formulation for the stability number of infinite graphs. *Math. Program.*, 160(1):65–83, 2016.

[55] S. Dobzinski and M. Sundararajan. On characterizations of truthful mechanisms for combinatorial auctions and scheduling. *Proceedings of the 9th ACM Conference on Electronic Commerce*, EC '08:38 – 47, 2008.

[56] H.-D. Doebner, V. K. Dobrev, and J. Hilgert. *Lie Theory and its Applications in Physics II*, chapter Functional-Analytic Methods, pages 1–402.

[57] A. C. Doherty and S. Wehner. Convergence of SDP hierarchies for polynomial optimization on the hypersphere. *ArXiv preprint 1210.5048*, 2013.

[58] H. Dong. Symmetric tensor approximation hierarchies for the completely positive cone. *SIAM J. Optim.*, 23(3):1850–1866, 2013.

[59] M. Dressler, S. Iliman, and T. de Wolff. A Positivstellensatz for sums of nonnegative circuit polynomials. *SIAM Journal on Applied Algebra and Geometry*, 1(1):536–555, 2017.

[60] M. Dür. Copositive programming – a survey. *University of Groningen, Johann Bernoulli Institute for Mathematics and Computer Science, 2010. ISBN 9783642125973. Relation: http://www.rug.nl/informatica/onderzoek/bernoulli Rights: University of Groningen, Johann Bernoulli Institute for Mathematics and Computer Science.*, 2010.

[61] G. Eichfelder and J. Povh. On the set-semidefinite representation of nonconvex quadratic programs over arbitrary feasible sets. *Optimization Letters*, 6(2):1–14, 2012.

[62] K. Elbassioni, K. Mehlhorn, and F. Ramezani. Towards more practical linear programming-based techniques for algorithmic mechanism design. *Theory Comput. Syst.*, 59(4):641–663, 2016.

[63] R. G. Eustaquio, E. W. Karas, and A. A. Ribeiro. Constraint qualifications for nonlinear programming. *Conference proceedings*, 2008. http://paginapessoal.utfpr.edu.br/eustaquio/my-research-interests/kkterabio.pdf.

[64] K. Fan. On a theorem of Weyl concerning eigenvalues of linear transformations. i.

Proceedings of the National Academy of Sciences of the United States of America, 35 (11):652–655, 1949.

[65] P. Fouilhoux and A. R. Mahjoub. Polyhedral results for the bipartite induced subgraph problem. *Discrete Applied Mathematics*, 154(15):2128 – 2149, 2006. International Symposium on Combinatorial Optimization CO'02.

[66] P. Fouilhoux and A. R. Mahjoub. Solving VLSI design and DNA sequencing problems using bipartization of graphs. *Computational Optimization and Applications*, 51(2): 749–781, Mar 2012.

[67] A. Frieze and M. Jerrum. Improved approximation algorithms for MAX k-CUT and MAX BISECTION. In E. Balas and J. Clausen, editors, *Integer Programming and Combinatorial Optimization*, pages 1–13, Berlin, Heidelberg, 1995. Springer Berlin Heidelberg.

[68] A. M. Frisch, B. Hnich, Z. Kiziltan, I. Miguel, and T. Walsh. Propagation algorithms for lexicographic ordering constraints. *Artificial Intelligence*, 170(10):803 – 834, 2006.

[69] L. Galli and A. N. Letchford. On the Lovasz's ϑ-function and some variants. *Discrete Optimization*, 25:159–174, 2017.

[70] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *J. Pure Appl. Algebra*, 192(1):95–128, 2004.

[71] P. C. Gilmore and R. E. Gomory. A linear programming approach to the cutting-stock problem. *Operations Research*, 9:849–859, 1961.

[72] M. X. Goemans and D. P. Williamson. .879 approximationn algorithms for MAX CUT and MAX 2SAT. In *Proceedings of the Twenty-sixth Annual ACM Symposium on Theory of Computing*, STOC '94, pages 422–431, New York, NY, USA, 1994. ACM. ISBN 0-89791-663-8.

[73] P. H. Groetzner. A method for completely positive and nonnegative matrix factorization. *Ph.D. thesis, Trier University*, 2018.

[74] G. Gruber and F. Rendl. Computational experience with stable set relaxations. *SIAM Journal on Optimization*, 13(4):1014–1028, 2003.

[75] L. Gualà and G. Proietti. Exact and approximate truthful mechanisms for the shortest paths tree problem. *Algorithmica*, 49(3):171–191, 2007.

[76] J. C. Guella and V. A. Menegatto. Schoenberg's theorem for positive definite functions on products: A unifying framework. *Journal of Fourier Analysis and Applications*, 25(4):1424–1446, Aug 2019.

[77] J. C. Guella, V. A. Menegatto, and A. P. Peron. An extension of a theorem of Schoenberg to products of spheres. *Banach J. Math. Anal.*, 10(4):671–685, 10 2016.

[78] F. Guo, L. Wang, and G. Zhou. Minimizing rational functions by exact Jacobian SDP relaxation applicable to finite singularities. *J. Glob. Optim.*, 58:261–284, 2014.

[79] F. Guo, C. Wang, and L. Zhi. Semidefinite representations of noncompact convex sets. *SIAM Journal on Optimization*, 25(1):377–395, 2015.

[80] N. Gvozdenović and M. Laurent. The operator Ψ for the chromatic number of a graph. *SIAM J. Optim.*, 19:572–591, 2008.

[81] M. M. Halldórsson, J. Y. Halpern, L. E. Li, and V. S. Mirrokni. On spectrum sharing games. In *Proceedings of the Twenty-third Annual ACM Symposium on Principles of Distributed Computing*, PODC '04, pages 107–114, New York, NY, USA, 2004. ACM. ISBN 1-58113-802-4.

[82] D. Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific J. Math.*, 132(1):35–62, 1988.

[83] G. Hardy, J. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, New York, second edition, 1988.

[84] D. Henrion and J. Lasserre. GloptiPoly: global optimization over polynomials with Matlab and SeDuMi. *ACM Trans. Math. Software*, 29(2):165–194, 2003. ISSN 0098-3500.

[85] A. Hertz, R. Montagné, and F. Gagnon. Constructive algorithms for the partial directed weighted improper coloring problem. *Journal of Graph Algorithms and Applications*, 20(2):159–188, 2016.

[86] A. Hertz, R. Montagné, and F. Gagnon. Online algorithms for the maximum k-colorable subgraph problem. *Computers & Operations Research*, 91:209 – 224, 2018.

[87] R. B. Holmes. *Geometric Functional Analysis and its Applications*. Springer-Verlag New York, 1975.

[88] F. Hüffner. Algorithm engineering for optimal graph bipartization. In S. Nikoletseas, editor, *Experimental and Efficient Algorithms*, pages 240–252, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[89] A. Huq. *The complexity of extended formulations*. PhD thesis, Georgia Institute of Technology, 2016.

[90] D. Husemöller. *Fibre Bundles*. Springer-Verlag New York, 1994.

[91] S. Iliman and T. de Wolff. Amoebas, nonnegative polynomials and sums of squares supported on circuits. *Research in the Mathematical Sciences*, 3(1):9, Mar 2016.

[92] M. F. Iqbal, F. Ahmed, M. Aqeel, and S. Ahmad. Approximation hierarchies for copositive tensor cone. *arXiv preprint 1806.01063*, 2018.

[93] T. Januschowski and M. E. Pfetsch. Branch-Cut-and-Propagate for the Maximum k-Colorable Subgraph Problem with Symmetry. In T. Achterberg and J. C. Beck, editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, pages 99–116, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[94] T. Januschowski and M. E. Pfetsch. The maximum k-colorable subgraph problem and orbitopes. *Discrete Optimization*, 8(3):478–494, 2011.

[95] V. Jeyakumar, J.-B. Lasserre, and G. Li. On polynomial optimization over non-compact semi-algebraic sets. *Journal of Optimization Theory and Applications*, 163 (3):707–718, 2014.

[96] V. Jeyakumar, S. Kim, G. M. Lee, and G. Li. Solving global optimization problems with sparse polynomials and unbounded semialgebraic feasible sets. *J. Glob. Optim*, 65(2):175–190, 2016.

[97] D. Jibetean and E. de Klerk. Global optimization of rational functions: a semidefinite programming approach. *Mathematical Programming*, 106(1):93–109, 2006.

[98] D. S. Johnson and M. A. Trick. Cliques, coloring, and satisfiability: Second DIMACS implementation challenge, october 11-13, 1993. *DIMACS Series in Discrete Math. and Th. Comp. Science. AMS*, 26:653–657, 1996. URL http://wwwdimacs.dimacs.rutgers.edu/pub/challenge/graph/benchmarks/clique/.

[99] D. S. Johnson, A. Mehrotra, and M. A. Trick. COLOR02/03/04: Graph coloring and its generalizations, 2002. URL https://mat.tepper.cmu.edu/COLOR04/.

[100] J. E. K. Jr. The cutting-plane method for solving convex programs. *SIAM J. Appl. Math.*, 8:703 – 712, 1960.

[101] A. Kabatiansky and V. Levenshtein. On bounds for packings on a sphere and in space. *Problems of Information Transmission*, 14(1):1–17, 1978.

[102] S. E. Karisch and F. Rendl. Semidefinite programming and graph equipartition. In P. M. Pardalos and H. Wolkowicz, editors, *Topics in Semidefinite and Interior-Point Methods*, pages 77–95. Fields Institute Communications, vol. 18, American Math. Soc., 1998.

[103] A. R. Karlin, C. Mathieu, and C. Nguyen. Integrality gaps of linear and semi-definite programming relaxations for knapsack. In *Integer Programming and Combinatoral Optimization Lecture Notes in Computer Science*, volume 6655, pages 301–314. Springer-Verlag, 2011.

[104] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller, J. W. Thatcher, and J. D. Bohlinger, editors, *Complexity of Computer Computations: Proceedings of a symposium on the Complexity of Computer Computations, held March 20–22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, and sponsored by the Office of Naval Research, Mathematics Program, IBM World Trade Corporation, and the IBM Research Mathematical Sciences Department*, pages 85–103. Springer US, Boston, MA, 1972.

[105] T. Koornwinder. Yet another proof of the addition formula for Jacobi polynomials. *Journal of Mathematical Analysis and Applications*, 61(1):136 – 141, 1977.

[106] E. Koutsoupias and A. Vidali. A lower bound of $1 + \phi$ for truthful scheduling mechanisms. *Algorithmica*, 66:211 – 223, 2013.

[107] J.-L. Krivine. Anneaux préordonnés. *Journal d'analyse mathématique*, 12(1):307–326, 1964.

[108] X. Kuang, B. Ghaddar, J. Naoum-Sawaya, and L. F. Zuluaga. Alternative LP and SOCP hierarchies for ACOPF problems. *IEEE Transactions on Power Systems*, 32(4):2828–2836, July 2017.

[109] O. Kuryatnikova and J. C. Vera. Approximating the cone of copositive kernels to estimate the stability number of infinite graphs. *Electronic Notes in Discrete Mathematics*, 62:303 – 308, 2017. LAGOS'17 – IX Latin and American Algorithms, Graphs and Optimization.

[110] O. Kuryatnikova and J. C. Vera. Positive semidefinite approximations to the cone of copositive kernels. *ArXiv preprint 1812.00274*, 2018.

[111] O. Kuryatnikova and J. C. Vera. Generalizations of Schoenberg's theorem on positive definite kernels. *ArXiv preprint 1904.02538*, 2018.

[112] O. Kuryatnikova and J. C. Vera. New bounds for truthful scheduling on two unrelated selfish machines. *Theory of Computing Systems*, May 2019. URL https://doi.org/10.1007/s00224-019-09927-x.

[113] J. Lagarias and P. Shor. Keller's cube-tiling conjecture is false in high dimensions. *Bull. AMS (New Series)*, 27(2):279–283, Oct 1992.

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

[115] J. Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity. *SIAM Journal on Optimization*, 17(3):822–843, 2006.

[116] J. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM Journal on Optimization*, 19(4):1995–2014, 2009. ISSN 1052-6234 (print), 1095-7189 (electronic).

[117] J. Lasserre. A new look at nonnegativity on closed sets and polynomial optimization. *SIAM Journal on Optimization*, 21(3):864–885, 2011. doi: 10.1137/100806990.

[118] J. B. Lasserre. An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM Journal on Optimization*, 12(3):756–769, 2002.

[119] J. B. Lasserre. Semidefinite programming vs. LP relaxations for polynomial programming. *Mathematics of Operations Research*, 27(2):347–360, 2002.

[120] J. B. Lasserre. New approximations for the cone of copositive matrices and its dual. *Math. Program.*, 144(1):265–276, Apr 2014.

[121] J. B. Lasserre, K.-C. Toh, and S. Yang. A bounded degree SOS hierarchy for polynomial optimization. *EURO Journal on Computational Optimization*, 5(1):87–117, Mar 2017.

[122] R. Lavi and C. Swamy. Truthful and near-optimal mechanism design via linear programming. *J. ACM*, 58(6):1–25, 2011.

[123] K. Lee, N. Funabiki, and Y. Takefuji. A parallel improvement algorithm for the bipartite subgraph problem. *IEEE transactions on neural networks*, 3(1):139–145, 1992. ISSN 1045-9227. doi: 10.1109/72.105427.

[124] J. Lenstra, D. Shmoys, and É. Tardos. Approximation algorithms for scheduling unrelated parallel machines. *Math. Program.*, 46:259 – 271, 1990.

[125] V. Levenshtein. On bounds for packing in n-dimensional Euclidean space. *Soviet Math. Dokl.*, 20:417–421, 1979.

[126] J. Lewis and M. Yannakakis. The node-deletion problem for hereditary properties is NP-complete. *Journal of Computer and System Sciences*, 20:219–230, 1980.

[127] R. Lippert, R. Schwartz, G. Lancia, and S. Istrail. Algorithmic strategies for the single nucleotide polymorphism haplotype assembly problem. *Briefings in Bioinformatics*, 3(1):23–31, 2002.

[128] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[129] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25 (1):1–7, 1979.

[130] P. Lu. On 2-player randomized mechanisms for scheduling. *Internet and Network Economics: 5th International Workshop, WINE 2009, Rome, Italy, December 14-18, 2009. Proceedings*, pages 30–41, 2009.

[131] P. Lu and C. Yu. An improved randomized truthful mechanism for scheduling unrelated machines. *Dans Proceedings of the 25th Annual Symposium on the Theoretical Aspects of Computer Science - STACS 2008, Bordeaux : France (2008)*, 2008.

[132] P. Lu and C. Yu. Randomized truthful mechanisms for scheduling unrelated machines. *Internet and Network Economics: 4th International Workshop, WINE 2008, Shanghai, China, December 17-20, 2008. Proceedings*, pages 402 – 413, 2008.

[133] C. Lund and M. Yannakakis. The approximation of maximum subgraph problems. In A. Lingas, R. Karlsson, and S. Carlsson, editors, *Automata, Languages and Programming*, pages 40–51, Berlin, Heidelberg, 1993. Springer Berlin Heidelberg.

[134] Z. Luo and L. Qi. Completely positive tensors: properties, easily checkable subclasses, and tractable relaxations. *SIAM Journal on Matrix Analysis and Applications*, 37 (4):1675–1698, 2016.

[135] A. S. M. Grötschel, L. Lovász. *Geometric Algorithms and Combinatorial Optimization.* Wiley, New York, 1988.

[136] F. C. Machado and F. M. de Oliveira Filho. Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry. *Experimental Mathematics*, 0:1–8, 2017.

[137] M. Marek-Sadowska. An unconstrained topological via minimization problem for two-layer routing. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 3(3):184–190, July 1984.

[138] F. Margot. Symmetry in integer linear programming. In M. Jünger, T. Liebling, D. Naddef, G. Nemhauser, W. Pulleyblank, G. Reinelt, G. Rinaldi, and L. Wolsey, editors, *50 Years of Integer Programming 1958–2008*, pages 647–681. Springer, Heidelberg, 2010.

[139] M. Marshall. *Positive polynomials and sums of squares.* Number 146. American Mathematical Soc., 2008.

[140] M. Marshall. Representations of non-negative polynomials, degree bounds, and applications to optimization. *Can. J. Math*, 61(1):205–221, 2009.

[141] M. Marshall. Polynomials non-negative on a strip. *Proceedings of the American Mathematical Society*, 138(5):1559–1567, 2010.

[142] J. Mercer and A. R. Forsyth. Functions of positive and negative type, and their connection the theory of integral equations. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 209(441–458):415–446, 1909.

[143] H. D. Mittelmann and F. Vallentin. High-accuracy semidefinite programming bounds for kissing numbers. *Experimental Mathematics*, 19(2):175–179, 2010.

[144] B. Mohar and S. Poljak. Eigenvalues in combinatorial optimization. In *Combinatorial and Graph-Theoretical Problems in Linear Algebra*, pages 107–151, New York, NY, 1993. Springer New York.

[145] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[146] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canad. J. Math.*, 17:533–540, 1965.

[147] A. Mu'alem and M. Schapira. Setting lower bounds on truthfulness: Extended abstract. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07:1143 – 1152, 2007.

[148] K. Murty and S. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Math. Program.*, 39:117–129, 1987.

[149] O. R. Musin. The one-sided kissing number in four dimensions. *Periodica Mathematica Hungarica*, 53(1):209–225, Sep 2006.

[150] O. R. Musin. The kissing number in four dimensions. *Annals of Mathematics*, 168: 1–32, 2008.

[151] O. R. Musin. Multivariate positive definite functions on spheres. In *Proceedings of the AMS Special Session on Discrete Geometry and Algebraic Combinatorics*, pages 177–190, 2014.

[152] G. Narasimhan. The maximum $k-$colorable subgraph problem. *Ph.D. thesis, University of Wisconsin-Madison*, 1989.

[153] G. Narasimhan and R. Manber. A generalization of Lovasz's ϑ- function. *DIMACS Series in Discrete Math. and Comp. Sci.*, pages 19–27, 1990.

[154] G. Nebe and N. J. A. Sloane. Table of the highest kissing numbers presently known. http://www.math.rwth-aachen.de/~Gabriele.Nebe/LATTICES/kiss.html, 2011.

[155] R. Nelsen. *An Introduction to Copulas*. Springer-Verlag New York, 2006.

[156] Y. Nesterov and A. Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics. SIAM, 1994.

[157] T. Netzer. Stability of quadratic modules. *Manuscripta Mathematica*, 129(2):251–271, 2009.

[158] H. Nguyen and V. Powers. Polynomials non-negative on strips and half-strips. *Journal of Pure and Applied Algebra*, 216(10):2225–2232, 2012.

[159] J. Nie. An exact Jacobian SDP relaxation for polynomial optimization. *Math. Program.*, 137(1):225–255, Feb 2013.

[160] J. Nie and M. Schweighofer. On the complexity of Putinar's Positivstellensatz. *Journal of Complexity*, 23(1):135–150, 2007.

[161] J. Nie, J. Demmel, and B. Sturmfels. Minimizing polynomials via sum of squares over the gradient ideal. *Math. Program.*, 106(3, Ser. A):587–606, 2006.

[162] J. Nie, Z. Yang, and X. Zhang. A complete semidefinite algorithm for detecting copositive matrices and tensors. *SIAM Journal on Optimization*, 28(4):2902–2921, 2018.

[163] N. Nisan and A. Ronen. Algorithmic mechanism design. *Games Econom. Behav.*, 35 (1):166 – 196, 2001.

[164] A. Odlyzko and N. Sloane. New bounds on the number of unit spheres that can touch a unit sphere in n dimensions. *J. Combin. Theory Ser. A*, 26:210–214, 1979.

[165] J. Ostrowski, J. Linderoth, F. Rossi, and S. Smriglio. Orbital branching. *Mathematical Programming*, 126(1):147–178, Jan 2011.

[166] M. Overton and R. Womersley. On the sum of the largest eigenvalues of a symmetric matrix. *SIAM Journal on Matrix Analysis and Applications*, 13(1):41–45, 1992.

[167] M. Padberg. The boolean quadric polytope: Some characteristics, facets and relatives. *Mathematical Programming*, 45(1):139–172, Aug 1989.

[168] A. Papachristodoulou, S. Prajna, and P. Parrilo. Sum of squares optimization toolbox for Matlab, apr 2002. URL http://citeseer.ist.psu.edu/539939.html;http://www.cds.caltech.edu/sostools/sostools.pdf. available at http://www.cds.caltech.edu/sostools/.

[169] C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43(3):425 – 440, 1991.

[170] D. Papp and S. Yildiz. Sum-of-squares optimization without semidefinite programming. *ArXiv preprint 1712.01792*, 2018.

[171] P. Parrilo. Structured semidefinite programming and algebraic geometry methods in robustness and optimization. *Ph.D. thesis, California Institute of Technology, Pasadena, CA*, 2000.

[172] P. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Program.*, 96(2):293–320, 2003.

[173] J. Peña, J. Vera, and L. Zuluaga. Computing the stability number of a graph via linear and semidefinite programming. *SIAM J. Optim.*, 18(2):87–105, 2007.

[174] J. Peña, J. Vera, and L. Zuluaga. Exploiting equalities in polynomial programming. *Operations Research Letters*, 36:223–228, 2008.

[175] J. Peña, J. C. Vera, and L. F. Zuluaga. Completely positive reformulations for polynomial optimization. *Mathematical Programming*, 151(2):405–431, 2015.

[176] J. Peña, J. C. Vera, and L. F. Zuluaga. Positive polynomials on unbounded domains. *ArXiv preprint 1709.03435*, 2017.

[177] F. Pfender. Improved Delsarte bounds for spherical codes in small dimensions. *Journal of Combinatorial Theory*, A 114(6):1133–1147, 2007.

[178] M. E. Pfetsch and T. Rehn. A computational comparison of symmetry handling methods for mixed integer programs. *Mathematical Programming Computation*, 11 (1):37–93, Mar 2019.

[179] C. Potts. Analysis of a linear programming heuristic for scheduling unrelated parallel machines. *Discrete Appl. Math*, 10(2):155 – 164, 1985.

[180] J. Povh and F. Rendl. Copositive programming motivated bounds on the stability and the chromatic numbers. *Discrete Optimization*, 6:231–241, 2009.

[181] V. Powers. Positive polynomials and the moment problem for cylinders with compact cross-section. *Journal of Pure and Applied Algebra*, 188(1–3):217–226, 2004.

[182] V. Powers and B. Reznick. A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *Journal of Pure and Applied Algebra*, 164(1-2): 221–229, 2001.

[183] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[184] M. Putinar and F. Vasilescu. Positive polynomials on semi-algebraic sets. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 328(7):585–589, 1999.

[185] M. Putinar and F.-H. Vasilescu. Solving moment problems by dimensional extension. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 328(6):495–499, 1999.

[186] F. Rendl. Semidefinite relaxations for partitioning, assignment and ordering problems. *4OR*, 10(4):321–346, Dec 2012.

[187] F. Rendl and R. Sotirov. The min-cut and vertex separator problem. *Computational Optimization and Applications*, 69(1):159–187, Jan 2018.

[188] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*, volume 3 of *MPS/SIAM Ser. Optim.* SIAM, 2001.

[189] J. Renegar. Accelerated first-order methods for hyperbolic programming. *Math. Program.*, 173(1-2):1–35, Jan. 2019.

[190] B. Reznick. Some concrete aspects of Hilbert's 17th problem. In C. N. Delzell and J. Madden, editors, *Real Algebraic Geometry and Ordered Structures*, volume 253 of *Cont. Math.*, pages 251–272. AMS, 2000.

[191] T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1970.

[192] T. Rockafellar and R. Wets. *Variational Analysis.* Springer–Verlag, Berlin, 1998.

[193] L. Roebers. Bounds on the generalized kissing number. *Master's thesis, Tilburg University*, 2018.

[194] M. S. S. Burgdorf, C. Scheiderer. Pure states, nonnegative polynomials and sums of squares. *arXiv preprint 0905.4161*, 2009.

[195] G. Sağol and E. A. Yıldırım. Analysis of copositive optimization based linear programming bounds on standard quadratic optimization. *Journal of Global Optimization*, 63(1):37–59, 2015.

[196] M. E. Saks and L. Yu. Weak monotonicity suffices for truthfulness on convex domains. *In Proceedings 6th ACM Conference on Electronic Commerce (EC)*, pages 286 – 293, 2005.

[197] J. Saunderson. Certifying polynomial nonnegativity via hyperbolic optimization. *ArXiv preprint 1904.00491*, 2019.

[198] C. W. Scherer and C. W. J. Hol. Matrix sum-of-squares relaxations for robust semidefinite programs. *Math. Program.*, 107(1):189–211, 2006.

[199] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(1):203–206, 1991.

[200] I. J. Schoenberg. Metric spaces and positive definite functions. *Transactions of the American Mathemticl Society*, 44:522–536, 1938.

[201] I. J. Schoenberg. Positive definite functions on spheres. *Duke Math. J.*, 9(1):96–108, 1942.

[202] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Transactions on Information Theory*, 25(4):425–429, July 1979. ISSN 0018-9448. doi: 10.1109/TIT.1979.1056072.

[203] A. Schrijver. New code upper bounds from the Terwilliger algebra and semidefinite programming. *IEEE Transactions on Information Theory*, 51(8):2859–2866, 2005.

[204] C. Schulze. Schmüdgen's theorem and results of positivity. *ArXiv preprint 1411.4446*, 2014.

[205] K. Schütte and B. van der Waerden. Das problem der dreizehn kugeln. *Mathematische Annalen*, 125:325–334, 1953.

[206] M. Schweighofer. An algorithmic approach to Schmüdgen's Positivstellensatz. *Journal of Pure and Applied Algebra*, 166(3):307 –319, 2002.

[207] M. Schweighofer. On the complexity of Schmüdgen's Positivstellensatz. *Journal of Complexity*, 20(4):529 – 543, 2004.

[208] E. V. Shchepin and N. Vakhania. An optimal rounding gives a better approximation for scheduling unrelated machines. *Oper. Res. Lett.*, 33(2):127 – 133, 2005.

[209] H. Sherali and W. Adams. A hierarchy of relaxations for mixed-integer zero-one programming problems. *Discrete Applied Mathematics*, 52:83–106, 1994.

[210] N. Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23: 731–734, 1987.

[211] R. Sotirov. SDP relaxations for some combinatorial optimization problems. In M. Anjos and J. Lasserre, editors, *Handbook of Semidefinite, Conic and Polynomial Optimization: Theory, Algorithms, Software and Applications*, pages 795–820. Springer US, 2012.

[212] R. Sotirov. An efficient semidefinite programming relaxation for the graph partition problem. *INFORMS Journal on Computing*, 26(1):16–30, 2014.

[213] G. Stengle. A Nullstellensatz and a Positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207(2):87–97, 1974.

[214] A. P. Subramanian, H. Gupta, S. R. Das, and M. M. Buddhikot. Fast spectrum allocation in coordinated dynamic spectrum access based cellular networks. In *2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pages 320–330, 2007.

[215] E. van Dam and R. Sotirov. New bounds for the max-k-cut and chromatic number of a graph. *Linear Algebra and its Applications*, 488:216 – 234, 2016. ISSN 0024-3795.

[216] R. Vaze and M. Coupechoux. Online budgeted truthful matching. *SIGMETRICS Perform. Eval. Rev.*, 44(3):3–6, 2017.

[217] H. Vui and P. So'n. Representations of positive polynomials and optimization on noncompact semialgebraic sets. *SIAM Journal on Optimization*, 20(6):3082–3103, 2010.

[218] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1):218–242, 2006.

[219] J. Wang. Nonnegative polynomials and circuit polynomials. *ArXiv preprint 1804.09455*, 2018.

[220] L. Wang. *Semidefinite Relaxations Approach to Polynomial Optimization and Its Extensions*. PhD thesis, University of California, San Diego, 2014.

[221] T. Weisser, J. Lasserre, and K.-C. Toh. Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity. *Math. Program. Computation*, 10(1):1–32, Mar 2018.

[222] S. Willard. *General topology*. Addison-Wesley Publishing Company, Inc, 1970.

[223] Wolfram|Alpha, 2009. http://www.wolframalpha.com.

[224] H. Wolkowicz and Q. Zhao. Semidefinite programming relaxations for the graph partitioning problem. *Discrete Applied Mathematics*, 96–97:461–479, 1999.

[225] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*. Springer US, 2000.

[226] W. Xia and L. F. Zuluaga. Completely positive reformulations of polynomial optimization problems with linear constraints. *Optimization Letters*, 11(7):1229–1241, Oct 2017.

[227] W. Xia, J. Vera, and L. F. Zuluaga. Globally solving non-convex quadratic programs via linear integer programming techniques, 2015.

[228] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441– 466, 1991.

[229] M. Yannakakis and F. Gavril. The maximum k-colorable subgraph problem for chordal graphs. *Information Processing Letters*, 24(2):133 – 137, 1987.

[230] E. Yildirim. On the accuracy of uniform polyhedral approximations of the copositive cone. *Optim. Methods Softw.*, 27:155–173, 2012.

[231] E. A. Yildirim. On the accuracy of uniform polyhedral approximations of the copositive cone. *Optimization Methods and Software*, 27(1):155–173, 2012.

[232] G. M. Ziegler. *Lectures on Polytopes*. Springer-Verlag New York, 1995.

[233] L. F. Zuluaga, J. Vera, and J. Peña. LMI approximations for cones of positive semidefinite forms. *SIAM Journal on Optimization*, 16(4):1076–1091, 2006.