UNIVERSIDADE D COIMBRA

Mina Saee Bostanabad

A SEMIDEFINITE APPROACH TO ALGEBRAIC OPTIMIZATION

Tese no âmbito do Programa Interuniversitário de Doutoramento em Matemática, orientada pelo Professor Doutor João Eduardo da Silveira Gouveia e apresentada ao Departamento de Matemática da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Agosto de 2019

A Semidefinite Approach to Algebraic Optimization

Mina Saee Bostanabad

1 2 9 0 UNIVERSIDADE D COIMBRA

U. PORTO

UC|UP Joint PhD Program in Mathematics

Programa Interuniversitário de Doutoramento em Matemática

PhD Thesis | Tese de Doutoramento

February 2020

Acknowledgements

I would like to thank first my adviser, Prof. João Gouveia. I can surely say that I would never have written a Ph.D. thesis without his help, guidance, and patience. I am very grateful to him for teaching me the optimization course and later teaching me what I have needed for my research from the initial to the final stages. Besides, I thank him for his tremendous help whom writing my papers, listening to my practice talks and working with me to perfect the way I answer questions during talks. I appreciate him for his patience in reviewing many inferior drafts and guiding me in almost every single detail. Our 3-year of collaboration has been a great pleasure to me and I hope that it will continue past this Ph.D..

My endless thanks goes to my co-authors, Prof. Alexander Kovačec and Prof. Ting Kei Pong. It was a real pleasure to work and interact with them. A special thank to Prof. Alexander Kovačec for investing so much time in helping me, reading and reviewing over our exchanging notes until they were perfect for submission for publication. I learned too much from his observant view and his attention to details.

I am grateful to the Fundação para a Ciência e a Tecnologia (FCT) which provided me the financial support to do my Ph.D. (scholarship PD/BD/128060/2016). I also thank CMUC and MobiWise project for financing me to travel to many interesting and useful conferences. I would also like to acknowledge all committee members of the UC|UP Ph.D. program throughout my Ph.D. years which provided me this opportunity to do my research in a pleasant environment. Moreover I would also like to thank the University of Coimbra and in particular, the Department of Mathematics, for all the academic, administrative, and technical supports. I am also thankful to Mrs. Rute Andrade for having helped me in so many administrative matters.

I also acknowledge many professors, library staff, friends, and all the people who followed, advised, and supported my studies and research efforts over the years.

Words cannot describe how thankful I am to my family and especially to my parents who have always filled my life with unconditional love and support and my heart is full of gratitude to my dear husband Ali for his emotional support and constant encouragement. My deepest thanks goes to my lovely and kind friend Indira which was always so inspiring and supportive to me.

Finally, I would like to dedicate this thesis to my mother and my husband.

Abstract

Conic optimization is one of the most important and thriving research areas in the optimization field and it has a clos connection with polynomial optimization through semidefinite programming. In fact, it is known that global nonnegativity of a polynomial can be checked using sums of squares and this amounts to solving a semidefinite program. However, semidefinite programming is expensive for large-scale problems. Several attempts have been done in literature to inner approximate the positive semidefinite cone by replacing the psd condition with conditions that are cheaper but still effective in practice. In this thesis, we give some certificates for nonnegativity of polynomials using bounded factor width matrices since the cones of matrices of bounded factor width give a hierarchy of inner approximations to the PSD cone. The concept of factor width for a positive semidefinite matrix has been introduced recently and very few works have been done in this area, with the most relevant being an exploration on the cone of factor width two matrices as an inner approximation for SOS problems, by Ahmadi and Majumdar, the so called SDSOS. We will prove new results for matrices with bounded factor width and use them to derive new results on the existence of certificates of nonnegativity of polynomials.

We also propose the use of the cone of nonnegative factor width two matrices as a natural inner approximation for the completely positive cone. Using projections of this cone we derive new graph-based second-order cone approximation schemes for completely positive programming. This approach is a compromise between the expressive power of existing SDP and speed of LP based inner approximations. We also present numerical results on random problems and the stable set problem to illustrate the effectiveness of our approach.

Keywords: Bounded factor width matrices, sums of squares, conic optimization, polynomial optimization, copositive optimization

Resumo

A optimização cónica é uma das áreas mais importantes e ativas no campo da otimização e está intimamente ligada à otimização polinomial através da programação semidefinida. De facto, é sabido que a não negatividade de polinómios pode ser verificada recorrendo a somas de quadrados, e isto resume-se a um programa semidefinido. Contudo, a programação semidefinida é dispendiosa para problemas de grande escala. Vários métodos foram propostos na literatura para aproximar pelo interior o cone de matrizes semidefinidas positivas substituindo a condição de sdp por condições mais leves mas ainda eficientes na prática. Nesta tese, estudamos certificados de não negatividade de polinómios com recurso a matrizes com largura de factores limitada, já que os cones de matrizes de largura de factores limitada formam uma hierarquia de aproximações interiores ao cone SDP. O conceito de largura de factores para uma matriz semidefinida positiva foi introduzido recentemente e poucos trabalhos formas produzidos nesta área, com o mais relevante sendo uma exploração da utilização do cone de matrizes de largura de factores menor ou igual a dois como aproximação interior para problemas de soma de quadrados, por Ahmadi e Majumdar, designada de SDSOS. Provaremos novos resultados para matrizes de largura de factores limitada e usá-los-emos para derivar novos resultados sobre a existência de certificados de não negatividade de polinómios.

Propomos ainda o uso do cone das matrizes não negativas de largura de factores no máximo dois como uma aproximação interior natural para o cone de matrizes completamente positivas. Usando projeções deste cone derivamos novos esquemas de aproximação por cones de segunda ordem para programação completamente positiva. Esta abordagem oferece um compromisso entre o poder expressivo da programação semidefinida e a velocidade das aproximações interiores por programas lineares. Apresentamos ainda resultados numéricos para problemas aleatórios e problemas de independência em grafos para ilustrar a eficiência da nossa abordagem.

Palavras-chave: Matrizes com largura de factores limitada, somas de quadrados, otimização cónica, otimização polinomial, otimização copositiva

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Recently, many important developments in convex optimization have been concentrated on conic programming. A conic program or conic optimization problem is the problem of optimizing a linear function over the intersection of a hyperplane and a proper cone. The major conic optimization problems that we will cover are linear, second order, semidefinite, copositive and completely positive optimization problems. Each of these classes can be used as a tool to deal with many diverse problems. In fact, these problems permit us to manifest the rich structure that some convex programs possess and to use this structure in order to process the program in a more efficient way. In general, linear, second order and semidefinite programming problems are more tractable than copositive and completely positive programmings since the two later ones are known to be NP-hard optimization problems, but even for those there are several approximation techniques to make them more tractable. In this dissertation we will review all of these conic programs and their connections to other problems. One optimization problem which is closely connected to conic programing is polynomial optimization. This problem can be reduced to certifying global nonnegativity of a polynomial and it is known to be NP-hard if the degree of the polynomial is at least four. Hence a usual strategy is to replace the nonnegativity constraint by a condition that is more tractable. A well known sufficient condition for checking nonnegativity of a polynomial is being a sum of squares of other polynomials since then it is clearly nonnegative. The advantage of using sum of squares is that that condition can be numerically checked efficiently by way of semidefinite programming (SDP). However, semidefinite programming does not scale well with the size of the problem and this limits the use of SDP for large scale problems. There are some attempts to increase the scalability of the sums of squares technique in the literature including taking advantage of the problem structure [24], customizing solvers for sos programs [55] or replacing the psd condition with some cheaper conditions to obtain a more efficient inner approximation to the sos problem [3]. The idea in [3] is to replace the psd condition used in sums of squares by a diagonally dominant or scaled diagonally dominant condition. The authors call these two problems DSOS and SDSOS problems respectively. These optimization problems are linear and second order cone programs and are in general more scalable with the size of the problem.

Scaled diagonally dominant matrices are a special case of bounded factor width matrices. The concept of factor width was introduced recently by Boman, et al [10] and it is defined as the smallest positive integer k for which the positive semidefinite matrix A can be written as $A = VV^T$ where each column of V contains at most k non-zeros. The cones of matrices of bounded factor width provide a hierarchy of inner approximations to the PSD cone. For $k = 2$ we recover scaled diagonally dominant matrices. In this thesis we explore the use of the cone of bounded factor width matrices for both polynomial and copositive (completely positive) optimizations.

The contribution of the thesis

This thesis has been constructed based on two papers. Chapter 3 is based on a paper entitled as *On sums of k-nomial squares* [26]. This paper is about bounded factor width matrices and using them as a certificate for nonnegativity of polynomials. The factor width is a concept which has been introduced recently and very little work has been done in this area. It has been proven by Boman [10] that matrices with factor width at most two are scaled diagonally dominant and from a polynomial optimization point of view, a Gram matrix of a polynomial having factor width two corresponds to the polynomial being sum of binomial squares, a fact that motivated Ahmadi and Majumdar [3] to reduce semidefinite programming to a second order cone programming which is more scalable. Nevertheless, very few results have been explored in literature for matrices with factor width more than two. In Chapter 3, we will prove some results on the geometry of the cones of matrices with bounded factor widths and their duals, and use them to derive new results on the existence of certificates of nonnegativity of polynomials by sums of k-nomial squares.

As another contribution of this thesis, in Chapter 4 we present another paper entitled as *Inner approximating the completely positive cone via the cone of scaled diagonally dominant matrices* [27]. Copositive programming and its dual counterpart of completely positive programming are classes of convex optimization problems that have in the past decades developed as a particularly expressive tool to encode optimization problems, especially for many problems arising from combinatorial or quadratic optimization. Since copositive and completely positive programs are in general NP hard problems, there are several inner and outer approximations to this problem. Although there are several outer approximations to the completely positive cone, for inner approximation the literature is somehow sparser with only a few existing LP and SDP based inner approximations. In Chapter 4 we propose a new inner approximation for the completely positive cone based on nonnegative factor width two or scaled diagonally dominant matrices. This approach is a compromise between the speed of LP and expressive power of SDP based inner approximations. We will support our idea with some numerical results for both random and stable set problems.

Organization of the thesis

This thesis is divided in four chapters. The first of which is this introduction. Chapter 2 is mainly reviewing some important definitions and concepts. We start it by reviewing conic programming, specially linear programming, second order cone programming, semidefinite programming and copositive and completely positive programming. We will cover more carefully the two later ones corresponding to the cones of copositive and completely positive matrices. We will describe the structure of these problems and their possible applications. We then review the bases of polynomial optimization reducing it to checking global nonnegativity of a polynomial. In Chapter 3 using the concept of factor width for positive semidefinite matrices and its connection to sums of squares, we give new certificates for checking global nonnegativity of polynomials. Finally in Chapter 4, using nonnegative scaled diagonally dominant matrices, we give a new inner approximation for the completely positive cone.

Chapter 2

Conic programming and polynomial optimization

In this chapter, we first give some basic definitions and concepts regarding convex cones and conic programming, we will review several types of conic programs that we will use throughout the thesis. In addition, we bring some geometric concepts of convex cones which are crucial for proving the results of Chapter 3. Then we will look to the problem of nonnegativity of a polynomial which is known to be a NP hard problem, hence we will review the sum of squares approach as an approximation for that. However, since this approximation does not scale well, we will review some techniques to improve the performance obtained from sums of squares problems trading off accuracy.

2.1 Convex cones

Throughout the thesis, we denote the matrices by upper case letters and their entries are represented in the corresponding lower case letters, e.g., d_{ij} as the (i,j)th entry of the matrix D. We also use lower case letters to denote vectors. For vectors u, $v \in \mathbb{R}^n$, we write $u \geq 0$ if u is elementwise nonnegative, and use $[u,v]$ to denote the line segment between u and v, i.e.,

$$[u,v] := \{tu + (1-t)v : \ t \in [0,1]\}.$$

For a set $S \subseteq \mathbb{R}^n$ we define the span of S as

$$\mathrm{span}(S) = \{\lambda_1 x_1 + \ldots + \lambda_k x_k \mid x_i \in S, \ \lambda_i \in \mathbb{R}\}$$

We also define the affine hull of S as

$$\mathrm{aff}(S) = \{\lambda_1 x_1 + \ldots + \lambda_k x_k \mid x_i \in S, \ \lambda_i \in \mathbb{R}, \ \lambda_1 + \ldots + \lambda_k = 1\}$$

A set S is convex if the line segment between any two points in S lies in S, i.e., if for any $x_1, x_2 \in S$ and any λ with $0 \leq \lambda \leq 1$, we have $\lambda x_1 + (1-\lambda)x_2 \in S$.

We call the set of all convex combinations of points in S the convex hull of a set S and we define it as

$$Conv(S) = \{\lambda_1 x_1 + \ldots + \lambda_k x_k \mid x_i \in S,\ \lambda_i \geq 0,\ \lambda_1 + \ldots + \lambda_k = 1\}$$

A set C is called a cone, if for every $x \in C$ and $\lambda \geq 0$, we have $\lambda x \in C$. A set C is a convex cone if it is convex and a cone, which means that for any $x_1, x_2 \in C$ and $\lambda_1, \lambda_2 \geq 0$, we have

$$\lambda_1 x_1 + \lambda_2 x_2 \in C.$$

Given a cone K in the Hilbert space H, we can define the dual cone by

$$K^* = \{y \in H^* \mid \langle y, x \rangle \geq 0, \text{ for all } x \in K\}$$

where $\langle\ ,\ \rangle$ is the inner product on H.
For instance, letting $H = \mathcal{S}^n$ to be the set of symmetric $n \times n$ matrices, for given two matrices A and B in $\mathcal{S}^n$, their standard matrix inner product is denoted by $\langle A, B \rangle := \sum_{i,j} A_{ij} B_{ij} = tr(AB)$ where we denote the trace of X as $\mathrm{tr}(X)$ and for a cone K of matrices in $\mathcal{S}^n$, its dual cone K^* can be defined as

$$K^* = \{Y \in \mathcal{S}^n \mid \langle Y, X \rangle \geq 0, \text{ for all } X \in K\}.$$

Note that the cone K^* is a convex closed cone regardless of K being closed and convex or not. In addition, $(K^*)^*$ is the closure of K, if K is convex. Also $K_1 \subseteq K_2$ implies $(K_2)^* \subseteq (K_1)^*$, for any two cones K_1, K_2. Proof for these properties can be found in [13].

A cone is called self-dual when $K = K^*$. Examples of self-dual cones that we will explain here are the nonnegative orthant, second order cone, positive semidefinite cone and nonnegative symmetric matrices.

Example 2.1.1. *Consider the nonnegative orthant,* $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_i \geq 0\}$*, the dual to this cone is itself since* $x^T y \geq 0$ *for all* $x \geq 0$ *if and only if* $y \geq 0$.

A second-order cone is defined as

$$Q^n = \{x \in \mathbb{R}^n \mid x = (x_n, \bar{x}), x_n \geq \|\bar{x}\|\}$$

where $\bar{x}$ is a subvector of x consisting of entries 1 to $n-1$ and $\|.\|$ is the standard Euclidean norm defined as $\|\bar{x}\| = (\sum_{i=1}^{n-1} x_i^2)^{\frac{1}{2}}$. This cone is also called the Lorentz cone or the ice cream cone.

The dual cone to the cone Q_n is defined as

$$(Q^n)^* = \{y \in \mathbb{R}^n \mid y^T x \geq 0, \text{ for all } x \in Q^n\}.$$

Example 2.1.2. *The second order cone is self-dual. To see this, first we will show* $Q^n \subseteq (Q^n)^*$*. If* $x, y \in Q^n$*, then from Cauchy-Schwartz inequality we have*

$$x^T y = x_n y_n + \sum_{i=1}^{n-1} x_i y_i \geq x_n y_n - \sqrt{\sum_{i=1}^{n-1} x_i^2} \sqrt{\sum_{i=1}^{n-1} y_i^2}$$

and from $x, y \in Q^n$, we have

$$x_n y_n - \sqrt{\sum_{i=1}^{n-1} x_i^2}\sqrt{\sum_{i=1}^{n-1} y_i^2} \geq 0.$$

Hence $x^T y \geq 0$. Next we will show that $(Q^n)^* \subseteq Q^n$. Suppose $y^T x \geq 0$, for all $x \in Q^n$. There are two cases. First, if $(y_1, \ldots, y_{n-1}) = (0, \ldots, 0)$, then consider $x_1, \ldots, x_{n-1} = (0, \ldots, 0)$ and $x_n = 1$. Then we have

$$y^T x \geq 0 \Leftrightarrow y_n \geq 0 \Leftrightarrow y_n^2 \geq \sum_{i=1}^{n-1} y_i^2 \Leftrightarrow y \in Q^n.$$

Otherwise, set $x_n = \sqrt{\sum_{i=1}^{n-1} y_i^2}$ and $x_i = -y_i$ for $i = 1, \ldots, n-1$, and hence

$$y^T x \geq 0 \Leftrightarrow y_n \sqrt{\sum_{i=1}^{n-1} y_i^2} - \sum_{i=1}^{n-1} y_i^2 \geq 0 \Leftrightarrow y_n^2 \geq \sum_{i=1}^{n-1} y_i^2, y_n \geq 0 \Leftrightarrow y \in Q^n.$$

A symmetric matrix A is positive semidefinite (psd), if $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. The set of positive semidefinite $n \times n$ symmetric matrices is denoted by $\mathcal{S}_+^n$, more precisely

$$\mathcal{S}_+^n = \{A \in \mathcal{S}^n \mid x^T A x \geq 0 \text{ for all } x \in \mathbb{R}^n\}.$$

A positive semidefinite matrix is denoted by the standard notation $\succeq 0$. The following proposition gives some useful properties of positive semidefinite matrices.

Proposition 2.1.1. *Let $A \in \mathcal{S}^n$. The following conditions are equivalent*

- $A \succeq 0$
- *The eigenvalues of A are nonnegative*
- *There exists $L \in \mathbb{R}^{n \times n}$ lower triangular such that $A = LL^T$ (Cholesky factorization) and we denote by $L = chol(x)$*
- $A = \sum_{i=1}^{Rank(A)} v_i v_i^T$, *for some $v_i \in \mathbb{R}_+^n$*
- *$det(A_I) \geq 0$, for all $I \subseteq \{1, 2, \ldots, n\}$ where A_I denotes the (principal) submatrix of A composed from rows and columns of A with indices in I*

Proof. Proofs to these properties can be found in [30]. ■

Example 2.1.3. *The positive semidefinite cone is self-dual, i.e., for $Y \in \mathcal{S}^n$,*

$$tr(XY) \geq 0 \text{ for all } X \succeq 0 \Leftrightarrow Y \succeq 0.$$

Assume that $Y \notin \mathcal{S}_+^n$, then there exists $x \in \mathbb{R}^n$ such that $x^T Y x = tr(x^T x Y) < 0$. Thus the positive semidefinite matrix $X = xx^T$ satisfies $tr(XY) < 0$ which leads to $Y \notin (\mathcal{S}_+^n)^*$. Now suppose $X, Y \in \mathcal{S}_+^n$. Then X can be written in terms of its eigenvalue decomposition as $X = \sum_{i=1}^n \lambda_i x_i x_i^T$, where (the *eigenvalues)* $\lambda_i \geq 0$, $i = 1,\dots,n$. *Then we have*

$$tr(XY) = tr\left(Y\sum_{i=1}^{n}\lambda_i x_i x_i^T\right) = \sum_{i=1}^{n}\lambda_i x_i^T Y x_i \geq 0.$$

This shows that $Y \in (\mathcal{S}_+^n)^*$.

Another well known self-dual cone is the set of nonnegative symmetric matrices defined as

$$\mathcal{N}^n = \{A \in \mathcal{S}^n \mid (A)_{ij} \geq 0, \text{ for all } i,j\}.$$

It can be easily seen that this cone is also self-dual, i.e., $\mathcal{N}^n = (\mathcal{N}^n)^*$.

Naturally, not all cones are self-dual. Examples of cones that are not self-dual include copositive and completely positive cones, that will play an important role in this work.

A symmetric matrix X is defined to be copositive if $v^T X v \geq 0$ for all nonnegative vectors v, and we denote the copositive cone by

$$\mathcal{COP}^n = \{X \in \mathcal{S}^n \mid v^T X v \geq 0, \text{ for all } v \geq 0\}.$$

A symmetric matrix X is defined to be completely positive if there exists a nonnegative matrix B such that $X = B^T B$, and we denote the completely positive cone by

$$\mathcal{CP}^n = \{X \in \mathcal{S}^n \mid \exists B \geq 0, \;\; X = B^T B\}.$$

Proposition 2.1.2. *[6, Theorem 2.3] Copositive and completely positive cones are dual to each other, i.e.,* $\mathcal{CP}^n = (\mathcal{COP}^n)^*$ *and* $\mathcal{COP}^n = (\mathcal{CP}^n)^*$.

Proof. Let A be an $n \times n$ symmetric matrix. Then $A \in (\mathcal{CP}^n)^*$ if and only if for every completely positive matrix B, $tr(AB) \geq 0$. Since B is completely positive, there exists a nonnegative matrix C such that $B = C^T C$. Hence $A \in (\mathcal{CP}^n)^*$ if and only if for every nonnegative matrix C with n rows $tr(AC^T C) \geq 0$ if and only if for every $c \in \mathbb{R}_+^n$, $c^T A c \geq 0$, which means that A is copositive. The dual claim that $\mathcal{CP}^n = (\mathcal{COP}^n)^*$ follows from the fact that $((\mathcal{CP}^n)^*)^* = \mathcal{CP}^n$ since both these cones are closed. ∎

We will talk about copositive and completely positive cones in more details in Chapter 4.

A cone K is pointed if $K \cap (-K) = \{0\}$, and solid if the interior of K is not empty. A cone that is convex, closed, pointed and solid is called a proper cone. Nonnegative orthants, second order cones, semidefinite cones, nonnegative cones and copositive and completely positive cones are all proper cones [21]. Another interesting example is that of nonnegative polynomials.

Proposition 2.1.3. *The set of nonnegative polynomials of fixed degree is a proper cone.*

Proof. We identify a polynomial with its coefficients and note that the constraints $p(x) \geq 0$ are linear in the coefficients of p for every fixed x, then it follows directly that this set is a convex set. it is also clear that this cone is solid and closed. In addition, this cone is pointed since if $p(x) \geq 0$ and $p(x) \leq 0$, then we have $p(x) = 0$. ∎

The dual cone of a proper cone is also a proper cone. A good introduction about convex cones and their duals can be found in [5].

2.2 Conic programming

Conic optimization is a subfield of convex optimization that studies problems consisting of minimizing a convex function over the intersection of an affine subspace and a proper cone. A standard conic programming problem in the primal form is

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \langle c, x \rangle \\ \text{s.t.} \quad & Ax = b \\ & x \in K, \end{aligned} \tag{2.1}$$

where $c \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$, and $\langle \ , \ \rangle$ denotes the standard Euclidean inner product.

The dual problem to problem (2.1) can be formulated as

$$\begin{aligned} \max_{y \in \mathbb{R}^m} \quad & \langle b, y \rangle \\ \text{s.t.} \quad & A^T y + s = c \\ & s \in K^*. \end{aligned} \tag{2.2}$$

There is an important connection between these two problems which is called weak duality.

Lemma 2.2.1. *For every x in the feasible set of problem* (2.1) *and y in the feasible set of problem* (2.2)*, we have*

$$\langle c, x \rangle \geq \langle b, y \rangle$$

Proof. For all x in the feasible set of problem (2.1) and y in the feasible set of (2.2), we have

$$\langle c, x \rangle = \langle A^T y + s, x \rangle = \langle A^T y, x \rangle + \langle s, x \rangle = \langle y, Ax \rangle + \langle s, x \rangle = \langle y, b \rangle + \langle s, x \rangle \geq \langle b, y \rangle$$

where $\langle s, x \rangle \geq 0$ follows since $s \in K^*$. ∎

When the optimal values of these problems are equal, we say that there is strong duality. In order to have strong duality, the most common sufficient condition considered is when one of the problems (2.1) or (2.2) is strictly feasible, i.e., there is a feasible point even if we replace the cone by its interior. This condition is called Slater condition.

Among the classes of problems that can be interpreted as particular cases of the general conic programming we have linear programming (LP), which has the following form

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$.

A set defined by finitely many linear inequalities or equations is called a polyhedron. Thus, linear programming corresponds exactly to the minimization of a linear function over a polyhedron. If a polyhedron is bounded, it is called a polytope.

In addition, to every LP problem we can associate a corresponding dual problem. This is another LP problem which can be formulated as following

$$\begin{array}{rl} \max & b^T y \\ \text{s.t.} & A^T y + s = c \\ & s \in \mathbb{R}^n_+. \end{array}$$

One of the most interesting properties of linear program is that strong duality always holds. In particular, when the primal problem is feasible and has bounded optimal objective value, then the primal and the dual both attain their optima with no duality gap.

Another kind of conic programming is called semidefinite programming (SDP). A semidefinite programming is a broad generalization of linear programming, where the decision variables are symmetric matrices and has the following primal form

$$\begin{array}{rl} \min & \langle C, X \rangle \\ \text{s.t.} & \langle A_i, X \rangle = b_i,\ i = 1, \ldots, m \\ & X \in \mathcal{S}^n_+ \end{array}$$

where $\langle\ ,\ \rangle$ denotes the matrix inner product.

The positive semidefinite cone is a self-dual proper cone and we can construct the dual SDP problem as

$$\begin{array}{rl} \max & b^T y \\ \text{s.t.} & C - \sum_i^m y_i A_i \in \mathcal{S}^n_+. \end{array}$$

Semidefinite programming problem in dual form corresponds to the optimization of a linear function subject to a linear matrix inequality (LMI) constraint (an LMI has the form $A_0 + \sum_{i=1}^m A_i x_i \succeq 0$ where $A_i \in \mathcal{S}^n$ are given symmetric matrices). The set defined by an LMI is called a spectrahedron, hence the feasible set of a semidefinite program is a spectrahedron. An important difference between a linear program and a semidefinite program is that in semidefinite programming there may be a finite or infinite duality gap. The primal and/or dual may or may not attain their optima. However both programs will attain their common optimal value if both programs are strictly feasible. Semidefinite programming can be cast as a powerful technique for tackling a diverse set of problems in applied and computational mathematics. We refer the interested readers to [52] for a more thorough introduction to semidefinite programming.

Another well known type of conic program is second-order cone programming (SOCP) formulated as

$$\begin{array}{rl} \min & c_1^T x_1 + \ldots + c_r^T x_r \\ \text{s.t.} & A_1 x_1 + \ldots + A_r x_r = b \\ & x_i \in Q, \quad i = 1, \ldots, r \end{array}$$

where $c_i, x_i \in \mathbb{R}^{n_i}$, Q is the Cartesian product of several cones, which means $Q = Q^{n_1} \times \ldots \times Q^{n_r}$, where each $Q^{n_i} \subseteq \mathbb{R}^{n_i}_+$ is a second order cone, and $n = \sum_{i=1}^{n} n_i$ is the dimension of the problem, m is the number of rows in each A_i and $A_i \in \mathbb{R}^{m \times n_i}$.

Recall that again, second order cones are self-dual proper cones and therefore so is Q, the dual SOCP problem can be formulated as

$$\begin{array}{rl} \max & b^T y \\ \text{s.t.} & A_i^T y + z_i = c_i, \quad i = 1, \ldots, r \\ & z_i \in Q, \quad i = 1, \ldots, r. \end{array}$$

with $z_i \in \mathbb{R}^{n_i}$. Second-order cone programs partially enjoy the expressive modeling power that nonpolyhedral cones such as the PSD cone have, but at the same time share with LP the scalability properties necessary for solving large-scale instances (of the order of tens of thousands of variables), which are currently out of reach for SDP.

Two other well known conic programming types are copositive and completely positive programming. Completely positive problem has the following primal form

$$\begin{array}{rl} \min & \langle C, X \rangle \\ \text{s.t.} & \langle A_i, X \rangle = b_i, \ i = 1, \ldots, m, \\ & X \in \mathcal{CP}^n, \end{array}$$

where C and A_i, $i = 1, \ldots, m$ are symmetric matrices and $\langle \ , \ \rangle$ denotes the matrix inner product. The dual problem which is a copositive programming problem has the following form

$$\begin{array}{rl} \max & b^T y \\ \text{s.t.} & C - \sum_{i=1}^{m} y_i A_i \in \mathcal{COP}^n, \end{array}$$

where $\mathcal{COP}^n$ is copositive cone.

Contrary to other types of conic programming that we mentioned before, copositive and completely positive programs can not be solved efficiently. Hence we need to solve them approximately. However, copositive and completely positive program have very high expressive power and many problems in combinatorial and quadratic optimization can be exactly formulated using them. We will study these problems in Chapter 4. Also, a more thorough introduction about copositive and completely positive programming can be found in [22] and references therein.

2.3 Geometry of convex cones

In this section we study some geometric properties of convex sets taken from [47]. We will need these definitions and concepts later on throughout the thesis. We start by defining a face of a convex set.

A face of a convex set S is a convex subset F of S such that every (closed) line segment in S with a relative interior point in S has both endpoints in S. The empty set and S itself are faces of S. The zero dimensional faces of S are called the extreme points of S. Thus we say that a point $x \in S$ is an extreme point of S if and only if there is no way to express x as a convex combination $(1-\lambda)y + \lambda z$ such that $y \in S$, $z \in S$ and $0 < \lambda < 1$, except by taking $y = z = x$.

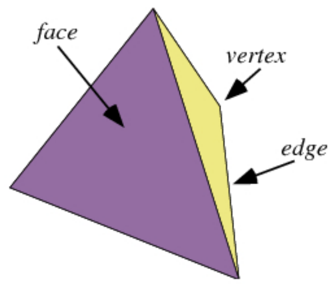

Fig. 2.1 Face, edge and vertex of a polyhedron

The concept of an extreme point for convex cones is not very useful since the origin would be the only candidate for an extreme point. Instead we study extreme rays of the cone. An extreme ray is a face which is a half-line emanating from the origin. Now, we explain when certain faces of convex sets are called exposed faces.

If we call F the set of points where a certain linear function h achieves its maximum over S, then F is a face of S. Namely, F is convex because it is the intersection of S and $\{x \mid h(x) = a\}$, where a is the maximum, and if the maximum is achieved on the relative interior of a line segment $L \subset S$, then h must be constant on L, so that $L \subset F$. A face of this type is called an exposed face. The exposed faces of S (aside from S itself and possibly the empty set) are thus precisely the sets of the form $S \cap H$, where H is a non-trivial supporting hyperplane to S.

We derive immediately from the definition of exposed face the following definition for exposed point of the convex set S. An exposed point of S is an exposed face which is a point, i.e. a point through which there is a supporting hyperplane which contains no other point of S. An exposed ray of a convex cone is an exposed face which is a half-line emanating from the origin. Notice that an exposed point is an extreme point and an exposed ray is an extreme ray. Figure 2.1 shows a two dimensional face, an edge and a vertex of a polyhedron. We also have the following lemma regarding the faces of a polyhedron.

Lemma 2.3.1. *[5, p55]. Every face of a polyhedron is exposed.*

It has been also proven in [44] that every face of a spectrahedron is exposed. We will talk about this later in Chapter 3.

2.4 Polynomial optimization

We start this section with some basic definitions, then we state the polynomial optimization problem.

For a vector variable $x \in \mathbb{R}^n$ and a vector $\alpha \in \mathbb{Z}_+^n$, we denote by x^α the monomial $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$ which by definition has degree $\sum_{i=1}^n \alpha_i$.

Definition 1. *A polynomial p in $x_1, \ldots, x_n$ is a finite linear combination of monomials:*

$$p = \sum_\alpha c_\alpha x^\alpha = \sum_\alpha c_\alpha x_1^{\alpha_1} \ldots x_n^{\alpha_n}$$

where the sum is over a finite number of n-tuples $\alpha = (\alpha_1, ..., \alpha_n)$, $\alpha_i \in \mathbb{Z}_{\geq 0}$ and c_α belong to some field K.

The set of all polynomials in x is denoted by $K[x]$. In this work we will only consider $K = \mathbb{R}$, and let P_n be the ring of real polynomials in n variables $x = (x_1, \ldots, x_n)$ and $P_{n,d}$ be the vector space of real polynomials in n variables and of degree less than or equal to d.

When all the monomials of a polynomial have the same degree, we call it a homogeneous polynomial or a form. It is well-known that there is a correspondence between forms and polynomials. A form in n variables and degree d can be dehomogenized to a polynomial in $n-1$ variables, of degree less than or equal to d, by fixing any variable to the constant value 1. Conversely, given a polynomial, it can be converted into a form by multiplying each monomial by powers of a new variable, in such a way that the degree of all monomials is the same.

A multivariate polynomial $p(x) = p(x_1, \ldots, x_n) \in P_{n,d}$ is nonnegative if it takes only nonnegative values, i.e.,

$$p(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^n.$$

We denote by $POS_{n,2d}$ the cone of nonnegative polynomials in $P_{n,2d}$,

$$POS_{n,2d} = \{p \in P_{n,2d} : p(x) \geq 0,\ \forall x \in \mathbb{R}^n\}.$$

The problem of minimizing a polynomial globally over $\mathbb{R}^n$ is a standard NP-hard problem in optimization [9]. This problem can be expressed as

$$\begin{aligned} \lambda^* = \quad & \min\ p(x) \\ & s.t.\ x \in \mathbb{R}^n. \end{aligned} \tag{2.3}$$

We can reformulate this in terms of checking non-negativity of a polynomial $p(x)$ in the following way

$$\begin{aligned} \lambda^* = \quad & \max\ \lambda \\ & s.t.\ p(x) - \lambda \geq 0,\ \forall x \in \mathbb{R}^n. \end{aligned} \tag{2.4}$$

Checking non-negativity of a polynomial is still an NP-hard problem. A usual approach to this problem is to replace the nonnegativity constrain on $p(x) - \lambda$ by something easier to check. A popular such condition is being able to write it as a sum of squares of real polynomials.

2.5 Sums of squares

A multivariate polynomial $p := p(x_1, \ldots, x_n)$ is a sum of squares (sos) if it can be written as the sum of squares of some other polynomials. Formally, we have the following.

Definition 2. *A polynomial $p(x) \in P_{n,2d}$ is a sum of squares (sos) if there exist $q_1, \ldots, q_m \in P_{n,d}$ such that*

$$p(x) = \sum_{k=1}^{m} q_k^2(x).$$

Example 2.5.1. *As an example the polynomial $p = 3x^4 + 4x^3y + 4x^2y^2 + 2xy^3 + y^4$ is sum of squares and it can be decomposed as $p = (x^2)^2 + (x^2 + xy)^2 + (x^2 + xy + y^2)^2$*

We denote the cone of sum of squares polynomials of n variables and degree $2d$ by $SOS_{n,2d}$.

$$SOS_{n,2d} = \{p(x) = \sum_i q_i^2(x) \text{ for some } q_i \in P_{n,d}\}$$

A sufficient condition for nonnegativity of a polynomial is being a sum of squares of polynomials, hence one can formulate a relaxation of problem (2.4) as following.

$$\begin{aligned} \lambda_{\text{SOS}} = \quad & \max \lambda \\ & s.t.\ p(x) - \lambda \ is\ sos \end{aligned} \tag{2.5}$$

Since sum of squares implies nonnegativity (i.e., $SOS_{n,2d} \subseteq POS_{n,2d}$), we have $\lambda_{\text{SOS}} \leq \lambda^*$.

Now, a natural question is to understand when a nonnegative polynomial can be written as a sum of squares. More than a century ago, David Hilbert showed that equality between the set of nonnegative polynomials $POS_{n,2d}$ and sum of squares polynomials $SOS_{n,2d}$ holds only in the following three cases [29]

- Univariate polynomials (i.e., $n = 1$).
- Quadratic polynomials ($2d = 2$).
- Bivariate quartics ($n = 2, 2d = 4$).

For all other cases, there always exist nonnegative polynomials that are not sum of squares. Perhaps the most celebrated example is the ternary sextic ($n = 3, 2d = 6$) due to Motzkin [38], given by

$$M(x,y,z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2.$$

This polynomial is nonnegative but is not a sum of squares. Nonnegativity of $M(x,y,z)$ follows from the arithmetic-geometric inequality as following

$$\frac{x^4y^2 + x^2y^4 + z^6}{3} \geq x^2y^2z^2.$$

Non-existence of a sos decomposition can be shown by assuming a decomposition $M = \sum_i q_i^2$ (with each q_i being a ternary form of degree 3), in other words, if we assume that $M(x,y,z)$ has a sos decomposition, then we can write it as

$$M(x,y,z) = \sum_k (A_kx^3 + B_kx^2y + C_kx^2z + D_kxy^2 + E_kxyz + F_kxz^2 + G_ky^3 + H_ky^2z + I_kyz^2 + J_kz^3)^2$$

Since $M(x,y,z)$ does not have x^6, y^6, x^4z^2, y^4z^2, x^2z^4 ,y^2z^4, then we have $A_k = G_k = C_k = H_k = F_k = I_k = 0$. So

$$M(x,y,z) = \sum_k (B_kx^2y + D_kxy^2 + E_kxyz + J_kz^3)^2$$

Hence the coefficient of $x^2y^2z^2$ in $M(x,y,z)$ is

$$\sum_k E_k^2 \geq 0$$

which is a contradiction. So we conclude that not all of nonnegative polynomials are sums of squares.

The next question that comes to our mind is how many nonnegative polynomials are sum of squares? Or, more precisely, is there any quantitative relationship between sum of squares polynomials and nonnegative polynomials?

Blekherman answered this question in [8] by comparing the relative sizes of theses two cones. He used the Urysohn's inequality as a main tool for computing the volumes of the sections of these two cones.

Lemma 2.5.1. *Let $K \subset \mathbb{R}^n$ be a convex body with 0 in its interior and let K° be the dual convex body. The following is known as Urysohn's inequality [49].*

$$\left(\frac{VolK}{VolB^n}\right)^{\frac{1}{n}} \leq \int_{S^{n-1}} G_{K^\circ}(x)d\sigma_x = \int_{S^{n-1}} \max_{y\in K}\langle x,y\rangle$$

where G_{K° is the Gauge function of K° and B^n and S^{n-1} denote the unit ball and unit sphere respectively.

Since the cones $SOS_{n,2d}$ and $POS_{n,2d}$ are unbounded objects, in order to compare the relative sizes of them, a section of each cone is taken with the same hyperplane so that both sections are compact. Let $L_{n,2d}$ to be an affine hyperplane in $P_{n,2d}$ consisting of all forms which integrate to 1 on the unit sphere S^{n-1} in $\mathbb{R}^n$

$$L_{n,2d} = \left\{p \in P_{n,2d} | \int_{S^{n-1}} pd\sigma = 1\right\},$$

where σ is the rotation invariant probability measure on S^{n-1}.

First, compact sections of both cones are taken by intersecting them with the hyperplane $L_{n,2d}$.

$$\overline{SOS}_{n,2d} = SOS_{n,2d} \cap L_{n,2d},$$

$$\overline{POS}_{n,2d} = POS_{n,2d} \cap L_{n,2d}.$$

Then, since cones should contain the origin, they are translated using a form $v^{2d} = (x_1^2 + \ldots + x_n^2)^d$ which is constantly 1 on the unit sphere and the compact convex bodies

$$\widetilde{SOS}_{n,2d} = \overline{SOS}_{n,2d} - v^{2d} = \{p \in P_{n,2d} \mid p + v^{2d} \in \overline{SOS}_{n,2d}\}$$

and

$$\widetilde{POS}_{n,2d} = \overline{POS}_{n,2d} - v^{2d} = \{p \in P_{n,2d} \mid p + v^{2d} \in \overline{POS}_{n,2d}\}$$

are obtained. Then the volumes of both convex bodies are bounded using Urysohn's inequality. The results are as following

Theorem 2.5.2. *[8, Theorem 1.1] There are the following bounds on the volume of $\widetilde{POS}_{n,2d}$*

$$\frac{1}{2\sqrt{4d+2}}n^{\frac{-1}{2}} \leq \left(\frac{Vol\widetilde{POS}_{n,2d}}{VolB_N}\right)^{\frac{1}{N}} \leq 4\left(\frac{2d^2}{4d^2+n-2}\right)^{\frac{1}{2}}.$$

Theorem 2.5.3. *[8, Theorem 1.1] There are the following bounds on the volume of $\widetilde{SOS}_{n,2d}$*

$$\frac{d!^2}{4^{2d}2d!\sqrt{24}}\frac{n^{\frac{d}{2}}}{(\frac{n}{2}+2d)^d} \leq \left(\frac{Vol\widetilde{SOS}_{n,2d}}{VolB_N}\right)^{\frac{1}{N}} \leq \frac{4^{2d}2d!\sqrt{24}}{d!}n^{-\frac{d}{2}}.$$

In both cases N is the ambient dimension. These bounds are asymptotically exact if the degree is fixed and number of variables tends to infinity. By comparing these two volumes the following relations between the volumes of these two convex bodies can be found [8].

Theorem 2.5.4.

$$\left(\frac{Vol\widetilde{POS}_{n,2d}}{Vol\widetilde{SOS}_{n,2d}}\right)^{\frac{1}{N}} \geq \frac{d!^2}{2(2d!)4^{2d}\sqrt{24(4d+2)}}n^{\frac{d-1}{2}}$$

When the degree is larger than two, if we fix the degree of the polynomials and let the number of variables increase, then there are significantly more nonnegative polynomials than sum of squares polynomials.

Although there are significantly more nonnegative polynomials than sum of squares polynomials, sum of squares still can be used as a main tool for checking nonnegativity of a polynomial.

2.6 Sums of squares as a semidefinite program

Now continuing our problem (2.5), one may ask how we can solve an sos problem? The problem (2.5) can be solved easily using semidefinite programming. We write the monomials in $x_1, x_2, \ldots, x_n$ up to degree d as a column vector $z(x)_d := [1, x_1, ..., x_n, x_1^2, x_1x_2, ..., x_n^d]^T$. By standard combinatorial reasoning, this vector has $\binom{n+d}{d}$ entries.

Proposition 2.6.1. *A polynomial $p(x)$ of degree $2d$ is sum of squares if and only if $p(x)$ can be written as $p(x) = z(x)_d^T Q z(x)_d$ where $z(x)_d$ is the vector of monomials of degree at most d and the matrix Q where we will call a Gram matrix for p, is positive semidefinite.*

Proof. Assume that the matrix Q is psd, then it can be factorized as $Q = L^T L$ (where L is an upper triangular matrix) and we have

$$p(x) = z(x)_d^T Q z(x)_d = z(x)_d^T L^T L z(x)_d = \|Lz(x)_d\|^2 = \sum_i (Lz(x)_d)_i^2$$

Here $Lz(x)_d$ is a vector of polynomials and when it is multiplied by its transpose, it gives a sum of squares polynomial, hence $p(x)$ has a sum of squares representation. Conversely, when $p(x)$ is a sum of squares, for some vectors of coefficients a_i, we must have

$$p(x) = \sum_i (a_i^T z(x)_d)^2 = \sum_i ((z(x)_d^T a_i)(a_i^T z(x)_d)) = z(x)_d^T (\sum_i a_i a_i^T) z(x)_d$$

so the positive semidefinite matrix $Q := \sum_i a_i a_i^T$ can be extracted. ■

Note that when $p(x)$ is homogeneous, we can pick the vector of monomials $z(x)$ to be also homogeneous and of degree exactly d. But when the polynomial $p(x)$ is not homogeneous, the vector of monomials $z(x)_d$ is of degree at most d. In addition, the size of the Gram matrix Q is $\binom{n+d-1}{d}$ when polynomial $p(x)$ is a form, otherwise it has size $\binom{n+d}{d}$.

The following example uses the certificate given in Proposition (2.6.1) to check whether a given polynomial is a sum of squares or not and, if yes, how we can decompose it as a sum of squares.

Example 2.6.1. *Consider the polynomial $p(x,y) = x^4 + 2x^3y + 4x^2y^2 + 4xy^3 + 3y^4$, this polynomial is a sum of squares if and only if $p(x,y) = z(x,y)^T Qz(x,y)$ and Q is positive semidefinite, where $z(x,y) = [x^2, xy, y^2]$.*
In other words, $p(x,y)$ is a sum of squares if and only if

$$p(x,y) = \begin{bmatrix} x^2 & xy & y^2 \end{bmatrix} \underbrace{\begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix}}_{Q} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$

and Q is positive semidefinite. Note that the equality can be rewritten as

$$x^4 + 2x^3y + 4x^2y^2 + 4xy^3 + 3y^4 = q_{11}x^4 + 2q_{12}x^3y + 2q_{13}x^2y^2 + 2q_{23}xy^3 + q_{33}y^4 + q_{22}x^2y^2,$$

so by comparison of coefficients, we have the following system of equalities

$$\begin{cases} q_{11} = 1, \\ 2q_{12} = 2, \\ 2q_{13} + q_{22} = 4, \\ 2q_{23} = 4, \\ q_{33} = 3. \end{cases}$$

So $p(x,y)$ is sum of squares if and only if there exists x such that

$$Q = \begin{bmatrix} 1 & 1 & x \\ 1 & 4-2x & 2 \\ x & 2 & 3 \end{bmatrix} \succeq 0.$$

For example, for $x = 1$, the matrix

$$Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

is positive semidefinite. Note that this x is not unique and different choices of x lead to different matrix Q and accordingly different sum of squares decompositions.
In order to recover the sum of squares decomposition, we factorize the matrix Q as $Q = U^TU$ using Cholesky factorization

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Also, note that this factorization is not unique and other factorizations can be used like the ones obtained by SVD or LU factorizations. Now by the factorization obtained above we have,

$$z(x,y)U^T Uz(x,y)^T = \begin{bmatrix} x^2 & xy & y^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$

$$= \left\| \begin{bmatrix} x^2+xy+y^2 \\ y^2+xy \\ y^2 \end{bmatrix} \right\|^2 = (y^2)^2 + (y^2+xy)^2 + (x^2+xy+y^2)^2$$

and this gives us a decomposition of polynomial $P(x,y)$ as a sum of squares. Note that here the matrix Q has rank three, therefore $p(x,y)$ is a sum of three squares.

Now, based on the certificate of sum of squares, problem (2.5) can be formulated as an SDP problem

$$\begin{aligned} \lambda_{\text{SDP}} = \quad & \max \lambda \\ s.t. \quad & p(x) - \lambda = z(x)_d^T Q z(x)_d \\ & Q \succeq 0 \end{aligned} \tag{2.6}$$

This problem can be solved efficiently using semidefinite programming solvers [37]. Although semidefinite programming is very useful in practice, it does not scale well when the size of the problem increases. In fact, in the absence of problem structure, sum of squares problems involving degree 4 or 6 polynomials are currently limited to around a dozen variables.

2.7 DSOS and SDSOS

In order to increase the scalability of sum of squares techniques, some contributions can be found in the literature. One approach is to take advantage of problem structure which means using sparsity or symmetry of the underlying polynomials to reduce the size of the SDP problem. This approach has been explored in [24], [51], [19].

Another approach is to customize solvers for SOS programs. Some works in this direction include [7], [40], [55]. There has also been recent work by Lasserre et al [35], that increases scalability of sum of squares optimization problems at the cost of accuracy of the solutions obtained. This can be done by bounding the size of the largest SDP constraint which appears in the SOS formulation, and this leads to what the authors call the BSOS (bounded SOS) hierarchy.

Recently Ahmadi and Majumdar [3] introduced more scalable alternatives to SOS optimization that they refer to as diagonally dominant sum of squares (dsos) and scaled diagonally dominant sum of squares (sdsos) programs.

The idea in [3] is to replace the conditions that the Gram matrix Q should be positive semidefinite with conditions which are stronger but cheaper, to obtain more efficient inner approximations to the SOS cone. Two such conditions come from the concepts of diagonally dominant and scaled diagonally dominant matrices in linear algebra.

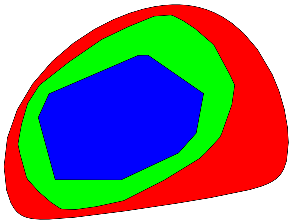

Fig. 2.2 PSD cone versus SDD and DD cones.

Definition 3. *A symmetric matrix A is diagonally dominant (dd) if* $a_{ii} \geq \sum_{j \neq i} |a_{ij}|$ *for all i. We say that A is scaled diagonally dominant (sdd) if there exists a diagonal matrix D, with positive diagonal entries, which makes DAD diagonally dominant.*

Note that this differs from the standard definition of diagonal dominance, since we are requiring the entries of the diagonal to be nonnegative.

Example 2.7.1. *Consider the matrices*

$$A = \begin{bmatrix} 1 & 0.6 & 0.5 \\ 0.6 & 2 & 0.2 \\ 0.5 & 0.2 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}, \quad DAD = \begin{bmatrix} 1 & 0.3 & 0.25 \\ 0.3 & 0.5 & 0.05 \\ 0.25 & 0.05 & 0.5 \end{bmatrix}.$$

Since DAD is diagonally dominant, A is scaled diagonally dominant.

We denote by DD^n and SDD^n the cones of all $n \times n$ diagonally dominant and scaled diagonally dominant matrices, respectively, and recall that $\mathcal{S}^n_+$ denotes the cone of positive semidefinite matrices. Then, from Gershgorin's circle theorem [25], it is clear that $DD^n \subseteq \mathcal{S}^n_+$ and since sdd matrices are diagonally dominant matrices scaled by a diagonal matrix, and scaling rows and columns does not effect psd-ness of a matrix, we have

$$DD^n \subseteq SDD^n \subseteq \mathcal{S}^n_+.$$

In order to illustrate this inclusion, in Figure 2.2 we generated two random symmetric matrices A and B of size 10×10. The outermost set is the feasible set of an SDP with the constraint $I + xA + yB \succeq 0$. where I denotes the identity matrix. The green and blue areas inside correspond to the points (x, y) for which $I + xA + yB$ are respectively sdd and dd.

Definition 4. *Recall that* $z(x)_d$ *denotes the vector of monomials of degree at most* d *. A polynomial* $p(x)$ *of degree* $2d$ *is said to be diagonally dominant sum of squares (resp. scaled diagonally dominant sum of squares) dsos (resp. sdsos) if it admits a representation as* $p(x) = z(x)_d^T Q z(x)_d$ *where the Gram matrix Q is a diagonally dominant (resp. scaled diagonally dominant) matrix.*

Now, the problem (2.6) can be approximated by the following DSOS and SDSOS problems

$$\begin{aligned} \lambda_{\text{DSOS}} = \quad & \max \lambda \\ & s.t.\ p(x) - \lambda = z(x)_d^T Q z(x)_d \\ & Q \quad \text{is dd} \end{aligned} \tag{2.7}$$

$$\begin{aligned} \lambda_{\text{SDSOS}} = \quad & \max \lambda \\ & s.t.\ p(x) - \lambda = z(x)_d^T Q z(x)_d \\ & Q \quad \text{is sdd} \end{aligned} \tag{2.8}$$

The problems (2.7) and (2.8) are linear and second order cone programs respectively and hence are more scalable with the size of the problem compared to the semidefinite problem (2.6). It is clear that DSOS and SDSOS programming are inner approximations to the SOS problem and hence are in general weaker than SOS, but solutions to these problems can be strengthened. Two strengthening techniques are basis pursuit and column generation which have been studied in [2] and [1] respectively. We will briefly explain these two techniques in the next sections.

Barker and Carlson in [4] gave another characterization of sdd matrices which will be very useful throughout this thesis. They proved that a matrix is sdd if and only if it can be written as the sum of positive semidefinite matrices whose supports are contained in some 2 by 2 submatrices. Here, support of a matrix X is defined as $\text{supp}(X) = \{(i,j) \in \{1,2,...,n\}^2 : x_{ij} \neq 0\}$. In other words,

Proposition 2.7.1. *A matrix A is scaled diagonally dominant if and only if it can be expressed as*

$$A = \sum_{i<j} M^{ij}_{2\times 2}$$

where each M^{ij} is an $n \times n$ matrix with zeros everywhere except on the 2×2 submatrix $\begin{bmatrix} M^{ij}_{ii} & M^{ij}_{ij} \\ M^{ij}_{ji} & M^{ij}_{jj} \end{bmatrix}$ *which is symmetric and positive semidefinite.*

As an example, the following matrix is scaled diagonally dominant since it can be written as sum of psd matrices with 2×2 support.

Example 2.7.2.

$$\begin{bmatrix} 1 & 0.6 & 0.5 \\ 0.6 & 2 & 0.2 \\ 0.5 & 0.2 & 2 \end{bmatrix} = \begin{bmatrix} 0.3 & 0.6 & 0 \\ 0.6 & 1.8 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0.7 & 0 & 0.5 \\ 0 & 0 & 0 \\ 0.5 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.2 & 0.2 \\ 0 & 0.2 & 1 \end{bmatrix}$$

Definition 5. *A polynomial $p := p(x)$ is sum of binomial squares (sobs) if it can be written as*

$$p = \sum_{i,j} (\alpha_{ij} m_i + \beta_{ij} m_j)^2$$

where m_i and m_j are monomials and $\alpha_{ij}, \beta_{ij} \geq 0$.

The characterization of SDD given by Barker and Carlson immediately implies that sdsos and sobs coincide, as illustrated in the example below.

Example 2.7.3. *Consider the polynomial* $p(x,y) = x^4 + 1.2x^3y + 3x^2y^2 + 0.4xy^3 + 2y^4$*. This polynomial is a sum of binomial squares since there exists a symmetric matrix* $Q \in \mathbb{R}^{3\times 3}$

$$Q = \begin{bmatrix} 1 & 0.6 & 0.5 \\ 0.6 & 2 & 0.2 \\ 0.5 & 0.2 & 2 \end{bmatrix}$$

such that $p(x,y) = z(x)^T Qz(x)$ *with* $z(x,y) = [x^2, xy, y^2]$ *and* Q *is sdd. This is true because as we can see from Example* (2.7.2)*, the matrix* Q *can be written as a sum of psd matrices with* 2×2 *support, hence* $p(x,y)$ *is sum of binomial squares. One such sobs decomposition is the following*

$$p(x,y) = 0.3(x^2+2xy)^2 + 0.6(xy)^2 + (0.5x^2+y^2)^2 + 0.45(x^2)^2 + 0.2(xy+y^2)^2 + 0.8(y^2)^2.$$

Now, based on the Proposition 2.7.1, the SDSOS problem can be reformulated as

$$\begin{aligned} \lambda_{\text{SDSOS}} = \quad & \max \lambda \\ & s.t.\ p(x) - \lambda = z(x)_d^T Q z(x)_d \\ & Q = \textstyle\sum_{i<j} M_{2\times 2}^{ij} \\ & M_{2\times 2}^{ij} \succeq 0 \end{aligned} \tag{2.9}$$

The constraints ($M_{2\times 2}^{ij} \succeq 0$) are rotated quadratic cone constraints and can be imposed using second order cone programming (SOCP), in other words

$$\begin{aligned} \lambda_{\text{SDSOS}} = \quad & \max \lambda \\ & s.t.\ p(x) - \lambda = z(x)_d^T Q z(x)_d \\ & Q = \textstyle\sum_{i<j} M_{2\times 2}^{ij} \\ & \left\| \begin{pmatrix} 2M_{ij}^{ij} \\ M_{ii}^{ij} - M_{jj}^{ij} \end{pmatrix} \right\| \leq M_{ii}^{ij} + M_{jj}^{ij} \\ & M_{ii}^{ij} + M_{jj}^{ij} \geq 0. \end{aligned} \tag{2.10}$$

As we mentioned before, second order cone programs can be solved easily for larger scale problems. The problem (2.10) is weaker than the problem (2.5), since we strengthened the restrictions, but it can be altered to better approximate the SOS problem which is the topic of the next section. We will mainly focus on strengthening techniques for the SDSOS problem.

2.8 Basis pursuit

As we mentioned before SDSOS programming can be strengthened to better approximate the SOS problem. One such strengthening is proposed by Ahmadi and Hall in [2].

The idea of basis pursuit can be illustrated by the following example which is taken from [2].

Example 2.8.1. *Suppose we would like to show that the degree-4 polynomial*

$$p(x) = x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 - 14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4$$

has a sum of squares decomposition. One way to do this is to try to write p as $p = z(x)^T Qz(x)$ *where*

$$z(x) = (x_1^2, x_1x_2, x_2^2, x_1x_3, x_2x_3, x_3^2)^T, \tag{2.11}$$

and the matrix Q is symmetric and positive semidefinite. As we explained before the search for such a Q can be done with semidefinite programming, and one feasible solution is as follows.

$$\begin{bmatrix} 1 & -3 & 0 & 1 & 0 & 2 \\ -3 & 9 & 0 & -3 & 0 & -6 \\ 0 & 0 & 16 & 0 & 0 & -4 \\ 1 & -3 & 0 & 2 & -1 & 2 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 2 & -6 & 4 & 2 & 0 & 5 \end{bmatrix}$$

Suppose now that instead of the basis z in (2.11)*, we pick a different basis*

$$\tilde{z}(x) = (2x_1^2 - 6x_1x_2 + 2x_1x_3 + 2x_3^2,\ x_1x_3 - x_2x_3,\ x_2^2 - \frac{1}{4}x_3^2)^T \tag{2.12}$$

With this new basis, we can get a sum of squares decomposition of p by writing it as

$$p = \tilde{z}(x)^T \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \tilde{z}(x).$$

In fact, by using a better basis, the Gram matrix has been simplified to be diagonal. Now positive semidefiniteness can be imposed as a linear constraint (diagonals should be nonnegative).

Hence, the goal of basis pursuit is to "pursue" the basis which simplifies the problem by starting with an arbitrary basis (typically the standard monomial basis), and then iteratively improving it by solving a sequence of SOCPs and performing some efficient matrix decomposition tasks in the process. In what follows we explain this method briefly.

Assume that we want to approximate the following SDP problem

$$\begin{aligned} &\min\ \langle C, X\rangle \\ &s.t.\ \langle A_i, X\rangle = b_i,\ i = 1, \dots, m \\ &X \succeq 0 \end{aligned} \tag{2.13}$$

In other to do this, first a family of cones is defined as following

$$SDD(U) := \{M \in S^n | M = U^T QU \text{ for some sdd matrix } Q\}$$

parametrized by an $n \times n$ matrix U. Optimizing over the set $SDD(U)$ is an SOCP and we have $SDD(U) \subseteq PSD$. This leads to the following iterative SOCP sequence

$$\begin{aligned} SDSOS_k = \quad & \min \langle C, X \rangle \\ & s.t.\ \langle A_i, X \rangle = b_i,\ i = 1, \ldots, m \\ & X \in SDD(U_k) \end{aligned} \tag{2.14}$$

Assuming existence of an optimal solution X_k at each iteration, we can define the sequence $\{U_k\}$ iteratively as

$$U_0 = I$$

$$U_{k+1} = chol(X_k).$$

Note that the first SOCP problem of this iterative approach is problem (2.8) where the Gram matrix is scaled diagonally dominant. Then by defining $U_{k+1} = chol(X_k)$, the optimal solution improves at each iteration. In fact, since $X_k = U_{k+1}^T I U_{k+1}$, and the identity matrix is scaled diagonally dominant, $X_k \in SDD(U_{k+1})$ and hence it is a feasible solution for iteration $k+1$ and therefore the optimal value improves at each iteration which means $SDSOS_{k+1} \leq SDSOS_k$. Now, since the sequence $\{SDSOS_k\}$ is lower bounded by SOS^* and monotonic, it must converge to a limit $SDSOS^* \geq SOS^*$.

Ahmadi and Hall [2] proved that if X_k is positive definite, then the solution improves strictly from step k to $k+1$. They proved this result for DSOS case, but the same is true for SDSOS.

Theorem 2.8.1. *[2, Theorem 3.1], Let X_k (resp. X_{k+1}) be an optimal solution of iterate k (resp.$k+1$) of problem* (2.14) *and assume that X_k is positive definite and $SOS^* < SDSOS_k$. Then,*

$$SDSOS_{k+1} < SDSOS_k$$

Since our original problem was to check nonnegativity of polynomial using sum of squares program, we iteratively improve SDSOS approximation to the SOS problem. In order to do that, yet another family of cones of degree $2d$ polynomials are defined as following

$$SDSOS(U) = \{p \mid p(x) = z(x)_d^T U^T Q U z(x)_d \quad \text{for some sdd matrix } Q\}$$

This set can also be viewed as the set of polynomials that are sdsos in the basis $Uz(x)_d$. In order to construct a sequence of SOCPs that generate improving bounds on the SOS optimal value, the constraint p is sos is replaced by $p \in SDSOS(U_k)$

$$\begin{aligned} \lambda_{SDSOS_k} = \quad & \max \lambda \\ & s.t.\ p(x) - \lambda = z(x)_d^T U^T Q U z(x)_d \\ & Q \in SDD(U_k) \end{aligned} \tag{2.15}$$

The sequence of matrices $\{U_k\}$ is defined as

$$U_0 = I$$

$$U_{k+1} = chol(U_k^T Q_k U_k).$$

where Q_k is an optimal Gram matrix of iteration k.

2.9 Column generation

Another strengthening technique is the one proposed by Ahmadi and Hall in [1] which is called column generation. In this iterative approach, the solutions to the SOCP problem improve iteratively by adding some atoms (columns) to the problem at each iteration. Geometrically, this amounts to optimizing over structured subsets of sum of squares polynomials that are larger than the set of sum of binomial squares polynomials.

In fact, SOS or SDSOS approaches can be considered as ways of proving that a polynomial is nonnegative by writing it as a nonnegative linear combination of certain atom polynomials that are already known to be nonnegative. For SOS, these atoms are all the squares and for SDSOS, these atoms are all the binomial squares. The idea of column generation is to start with a certain cheap subset of atoms (columns) and only add new ones, one or a limited number in each iteration if they improve the desired objective function.

We start column generation for a general SDP problem in the following form

$$\begin{array}{rl} \max & b^T y \\ \text{s.t.} & C - \sum_i^m y_i A_i \succeq 0 \end{array} \tag{2.16}$$

with $b \in \mathbb{R}^m$, $C, A_i \in \mathcal{S}^n$ and its dual

$$\begin{array}{l} \min \langle C, X \rangle \\ s.t.\ \langle A_i, X \rangle = b_i,\ i = 1, \dots, m \\ X \succeq 0 \end{array} \tag{2.17}$$

SOCP-based column generation optimizes over structured subsets of the positive semidefinite cone that are SOCP representable and are larger than the set of scaled diagonally dominant matrices. This will be achieved by working with the following SOCP

$$\begin{array}{rl} \max_{y \in \mathbb{R}^m, \Lambda_i \in \mathcal{S}^2} & b^T y \\ \text{s.t.} & C - \sum_i^m y_i A_i = \sum_{i=1}^t V_i \Lambda_i V_i^T \\ & \Lambda_i \succeq 0, i = 1, \dots, t \end{array} \tag{2.18}$$

Here, the decision matrices Λ_i are 2×2 and the positive semidefiniteness constraints on them can be imposed via rotated quadratic cone constraints. The $n \times 2$ matrices V_i are fixed for all $i = 1, \dots, t$.

To generate a new SOCP atom, we work with the dual of (2.18)

$$\begin{array}{l} \min \langle C, X \rangle \\ s.t.\ \langle A_i, X \rangle = b_i,\ i = 1, \dots, m \\ V_i^T X V_i \succeq 0,\ i = 1, \dots, t \end{array} \tag{2.19}$$

Now, if the optimal solution X^* to the problem (2.19) is psd, we are done, if not, one way to use X^* for producing new SOCP-based cuts is to put the two eigenvectors of X^* corresponding to its two most negative eigenvalues as the columns of an $n \times 2$ matrix V_{t+1}, in this way a new useful atom is produced.

We will always be initializing our SOCP iterations with the SDSOS bound. It is not hard to see that this corresponds to the case where we have $\binom{n}{2}$ initial $n \times 2$ atoms V_i, which have zeros everywhere, except for a 1 in the first column in position j and a 1 in the second column in position $k > j$. If only one exists, we can complete it with a row of zeros.

Chapter 3

Bounded factor width matrices and sums of squares polynomials

In 2005, Boman et al [10] introduced the concept of factor width for a positive semidefinite matrix A. This is the smallest positive integer k for which one can write the matrix as $A = VV^T$ with each column of V containing at most k non-zeros. The cones of matrices of bounded factor width give a hierarchy of inner approximations to the PSD cone. In the polynomial optimization context, the Gram matrix of a polynomial having factor width k corresponds to the polynomial being a sum of squares where each polynomial being squared has support of size at most k.

Recently, Ahmadi and Majumdar [3] explored this connection and proposed to replace the reliance on sum of squares polynomials in semidefinite programming to sum of binomial squares polynomials (sobs) which they call SDSOS, for which semidefinite programming can be reduced to a second order programming to gain scalability at the cost of some tolerable loss of precision. In fact, the study of sobs goes back to Reznick [46] and Hurwitz [31]. In this chapter we will prove some results on the geometry of the cones of matrices with bounded factor widths and their duals, and use them to derive new results on the existence of certificates of nonnegativity of polynomials by sums of k-nomial squares.

3.1 On the factor width of a matrix

The concept of factor width for a matrix was first introduced by Boman et al in [10]. They gave the following definition for the factor width of a matrix

Definition 6. *The factor width of a positive semidefinite matrix A is the smallest integer k such that there exists a real (rectangular) matrix V where $A = VV^T$ and each column of V contains at most k non-zeros.*

Example 3.1.1. *As an example the following matrix has factor width* 2

$$A = \begin{bmatrix} 4 & 0 & 4 \\ 0 & 1 & 1 \\ 4 & 1 & 5 \end{bmatrix}$$

since it can be factorized as

$$\begin{bmatrix} 4 & 0 & 4 \\ 0 & 1 & 1 \\ 4 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 2 & 1 \end{bmatrix}^T$$

and each column of $V = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 2 & 1 \end{bmatrix}$ *has two non zeros.*

It has been proved in [10, Proposition 2] that diagonally dominant (dd) and scaled diagonally dominant (sdd) matrices have factor width at most two.

Theorem 3.1.1. *A matrix has factor width at most two if and only if it is diagonally dominant or scaled diagonally dominant.*

We let

$$FW_k^n = \{\text{symmetric positive semidefinite } n \times n \text{ matrices of factorwidth} \leq k.\}.$$

We have of course

$$FW_1^n \subset FW_2^n \subset FW_3^n \subset \cdots \subset FW_n^n = \mathcal{S}_+^n.$$

Next assume $A = VV^T$ is a symmetric positive semidefinite matrix where each column of V has at most k nonzero entries. By the rules of matrix multiplication, for any $i, j \in \{1, ..., n\}$, and writing $V_{*\nu}$ and $V_{\nu *}$ for the ν-th column or row of a matrix V, respectively, we have

$$(VV^T)_{ij} = \sum_{\nu=1}^{m} V_{i\nu}(V^T)_{\nu j} = \sum_{\nu=1}^{m} (V_{*\nu}V^T{}_{\nu *})_{ij} = \sum_{\nu=1}^{m} (V_{*\nu}V_{*\nu}{}^T)_{ij}.$$

Write $A = \sum_{\nu \in V}(V_{*\nu}V_{*\nu}{}^T)$. Note that each $V_{*\nu}V_{*\nu}{}^T$ is a symmetric $n \times n$ rank 1 matrix, whose support lies within a cartesian product $K^2 = K \times K$ for some $K \subseteq \{1, 2, ..., n\}$ of cardinality k. Since at the other hand every $n \times n$ matrix with the latter properties can be written as vv^T for some v with number of nonzero entries$\leq k$, we have the following

Proposition 3.1.2. *Let A be an $n \times n$ symmetric positive semidefinite matrix, and assume $k \in \mathbb{Z}_{\geq 1}$. Then $A \in FW_k^n$ if and only if A is the sum of a finite family of symmetric positive semidefinite $n \times n$ matrices whose supports are all contained in sets $K \times K$ with $|K| = k$.*

Example 3.1.2. *As an example, the following* 4×4 *matrix in the left hand side has factor width* 3 *since it can be written as sum of psd matrices on the right hand side which these matrices are zeros everywhere except on some* 3×3 *sub matrices.*

$$\begin{bmatrix} 3 & 2 & 2 & 2 \\ 2 & 3 & 2 & 2 \\ 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix},$$

From Proposition 3.1.2, it follows immediately that each set FW_k^n is a convex closed subcone of $\mathcal{S}_+^n$.

Let us determine the dual cone of FW_k^n.

Proposition 3.1.3. *a. The dual of FW_k^n is given by*

$$(FW_k^n)^* = \{X \in \mathcal{S}^n \mid X_K \in \mathcal{S}_+^k \textit{ for all } K \subseteq \{1,2,...,n\} \textit{ with } |K| = k\}.$$

b. There hold the following inclusions between FW_k^n and its dual and double dual:
$FW_k^n \subseteq (FW_k^n)^*$ and $FW_k^n = (FW_k^n)^{**}$.

Proof. a. We have the following computation

$$\begin{aligned}(FW_k^n)^* &= \{X \in \mathcal{S}^n \mid \langle X, A\rangle \geq 0 \text{ for all } A \in FW_k^n\}\\ &= \{X \in \mathcal{S}^n \mid \langle X, B\rangle \geq 0 \text{ for all } B \in \mathcal{S}_+^n \text{with } supp(B) \subseteq K \text{ for some } k-\text{set } K \subseteq \{1,...,n\}\}\\ &= \{X \in \mathcal{S}^n \mid \langle X_K, B\rangle \geq 0 \text{ for all } B \in \mathcal{S}_+^k \text{ and all } K \subseteq \{1,\dots,n\}, \text{with } |K| = k\}\\ &= \{X \in \mathcal{S}^n \mid X_K \in (\mathcal{S}_+^k)^* \text{ for all } K \subseteq \{1,\dots,n\}, \text{ with } |K| = k\}\\ &= \{X \in \mathcal{S}^n \mid X_K \in \mathcal{S}_+^k \text{ for all } K \subseteq \{1,\dots,n\}, \text{ with } |K| = k\},\end{aligned}$$

where the first equality follows from the definition of a dual cone, the second from the characterization of the cone FW_k^n given in the Proposition 3.1.2, the third by one of the possibilities to write the inner product $\langle , \rangle$ on $\mathcal{S}_n$, the last two equalities are then consequences of the characterization and the selfduality of the cone $\mathcal{S}_+^n$ mentioned in Chapter 2.

b. If $X \in FW_k^n$ then as seen, X is a symmetric positive semidefinite matrix and so $X_K \in \mathcal{S}_+^k$ for any $K \subseteq \{1,2,...,n\}$ with $|K| = k$, by [30]. So by part a, $X \in (FW_k^n)^*$. The final equality follows from the fact that FW_k^n is a closed convex cone. ■

3.2 On the geometry of bounded factor width matrices

In this section, we give some geometric properties of the cone of bounded factor width matrices. In particular, we characterize some of the extreme rays of their duals.

We start with the following lemma about exposedness of the extreme rays of $(FW_k^n)^*$.

Lemma 3.2.1. *The cone $(FW_k^n)^*$ is (linearly equivalent to) a spectrahedron. Therefore a matrix in $(FW_k^n)^*$ which spans an extreme ray is an exposed ray.*

Proof. Let $E_{\{i,j\}}$ be the symmetric $n \times n$ matrix which has zeros everywhere except at the entries (i, j) and (j, i) where it has 1s and define for $l = 1, 2, \dots, \binom{n}{k}$ the matrix

$$E_{\{i,j\}}^l = \begin{cases} E_{\{i,j\}} & \text{if } i, j \text{ are both contained in the } l\text{th of the sets } I_1, I_2, \dots, I_{\binom{n}{k}} \text{ which} \\ & \text{they are } \binom{n}{k} \text{ distinct } k \text{ element subsets of } \{1,2,\dots,n\}; \\ 0 & \text{otherwise.} \end{cases}$$

Consider now the condition

$$\sum_{1 \leq i \leq j \leq \binom{n}{k}} b_{ij} (E_{\{i,j\}}^1 \oplus E_{\{i,j\}}^2 \dots \oplus E_{\{i,j\}}^n) \succeq 0.$$

Since a direct sum of matrices is positive semidefinite if and only if each of its summands is positive semidefinite, the attentive reader finds that this condition expresses precisely that the submatrices B_{I_r}, $r = 1,\dots,\binom{n}{k}$ with $|I_r| = k, I_r \subseteq \{1,\dots,n\}$ of $B = (b_{ij}) \in \mathcal{S}^n$ should be positive semidefinite. Since this is the defining property of B to be in $(FW_k^n)^*$ we find that $(FW_k^n)^*$ is a spectrahedron. The second part is a consequence of the theorem that every face of a spectrahedron is exposed. This is proved in [44, Corollary 1]. ■

Our first result about the extreme rays of the cone $(FW_k^n)^*$ is as following

Lemma 3.2.2. *The matrix $A \in \mathcal{S}_+^n$ spans an extreme ray of $(FW_k^n)^*$ if and only if it has rank* 1.

Proof. Let $A \in \mathcal{S}_+^n$ span an extreme ray of $(FW_k^n)^*$ and assume $\operatorname{rank}(A) = r \geq 2$. Then, as $A \in \mathcal{S}_+^n$, one can write $A = x_1x_1^T + \cdots + x_rx_r^T$ with real pairwise orthogonal x_i. Since the $x_ix_i^T \in \mathcal{S}_+^n$, $i = 1,...,r$, these $x_ix_i^T$ are elements of $(FW_k^n)^*$ since $FW_k^n \subset \mathcal{S}_+^n \subset (FW_k^n)^*$ and are not multiples of each other, thus A is not an extreme ray. So for extremality of A rank equal to 1 is necessary.

Now we prove that if the matrix A has rank 1, then it spans an extreme ray of $(FW_k^n)^*$. Assume not. i.e., $A = xx^T = X + Y$ with some $X, Y \in (FW_k^n)^*$ and some $x \in \mathbb{R}^n$. Then for any k element subset $I \subseteq \{1,2,\dots,n\}$, $x_Ix_I^T = X_I + Y_I$. By the characterization of $(FW_k^n)^*$, X_I, Y_I are positive semidefinite; that is we have found in $\mathcal{S}_+^n$ a representation of a rank 1 matrix as a sum of two other matrices. Since the null space of a sum of two psd matrices is contained in the nullspace of each, we infer that X_I, Y_I are multiples of $x_Ix_I^T$: for some real λ_I, $X_I = \lambda_I x_Ix_I^T$, $Y_I = (1-\lambda_I)x_Ix_I^T$. Now, considering any two $k \times k$ submatrices of X indexed by I and J, we have if $i \in I \cap J$, then $x_{ii} = \lambda_I x_i^2 = \lambda_J x_i^2$ so if $x_{ii} \neq 0$ then $\lambda_I = \lambda_J$. Note that if $x_{ii} = 0$, the entire i-th row and column of X must be zero. For any I and J such that $i \in I$ and $j \in J$ with $x_{ii} \neq 0$ and $x_{jj} \neq 0$, we can pick K such that $\{i, j\} \in K$ and the above argument gives $\lambda_I = \lambda_J = \lambda_K$. So all are equal to some λ and $X = \lambda xx^T$. ■

Next, we present a simple fact which will help us in the next theorem to characterize the extreme rays of $(FW_{n-1}^n)^*$.

Lemma 3.2.3. *Assume that $A \in (FW_{n-1}^n)^*$ and let A_I to be an $n-1 \times n-1$ principal submatrix of A for some I with $I \subseteq \{1,2,\dots,n\}$, if $\operatorname{rank}(A_I) \leq n-3$, then A is psd.*

Proof. Since $A \in (FW_{n-1}^n)^*$, all its proper principal minors are nonnegative. So A is psd if and only if $det(A) \geq 0$. But by Cauchy's interlacing theorem [30] if $\beta_1,\dots,\beta_{n-1}$ are the eigenvalues of A_I and $\gamma_1,\dots,\gamma_n$ are the eigenvalues of A,

$$\gamma_1 \leq \beta_1 \leq \gamma_2 \leq \beta_2 \leq \ldots \leq \beta_{n-1} \leq \gamma_n$$

Now, since $rank(A_I) \leq n-3$, β_1 and β_2 should be zero which leads to $\gamma_2 = 0$ and so $det(A) = 0$, hence A is psd. ■

Theorem 3.2.4. *If the matrix $A \in (FW_{n-1}^n)^*$ is not psd, the matrix A spans an extreme ray of $(FW_{n-1}^n)^*$, if and only if all of its $(n-1)\times(n-1)$ principal submatrices have rank $n-2$.*

Proof. We first prove that if the matrix A spans an extreme ray of $(FW_{n-1}^n)^*$, then all of its $(n-1)\times(n-1)$ principal submatrices have rank $n-2$. Assume that this does not happen, which means there is one $(n-1)\times(n-1)$ principal submatrix which is full rank, otherwise by Lemma 3.2.3 A will be psd and suppose $A_{\{1,2,\dots,n-1\}}$ is such principal submatrix. Since the cone $(FW^n_{n-1})^*$ is a spectrahedron, by Lemma 3.2.1 every face of it is exposed. Hence A is an exposed extreme ray of $(FW^n_{n-1})^*$. So, there exists a $B \in (FW^n_{n-1})^{**} = FW^n_{n-1}$ such that $\langle B,A\rangle = 0$ and $\langle B,X\rangle > 0$ for all $X \in (FW^n_{n-1})^* \setminus \{\lambda A \mid \lambda \geq 0\}$.

This $B \in FW^n_{n-1}$, and so it can be written as

$$B = \sum_{I\subseteq\{1,2,\dots,n\},|I|=n-1} \iota_I(B_I), \quad \text{for } B_I \in \mathcal{S}^{n-1}_+.$$

We thus get

$$0 = \langle B,A\rangle = \sum_{I\subseteq\{1,2,\dots,n\},|I|=n-1} \langle \iota_I(B_I),A\rangle = \sum_{I\subseteq\{1,2,\dots,n\},|I|=n-1} \langle B_I,A_I\rangle.$$

Since the $(n-1)\times(n-1)$ principal submatrices of A are all positive semidefinite, we get that all the inner products are nonnegative and hence must be 0. Which means $\langle B_I,A_I\rangle = 0$ for all I.

Under the current supposition that $A_{\{1,2,\dots,n-1\}}$ is not singular, we conclude that $B_{\{1,2,\dots,n-1\}} = 0$.

Let now a be the n-th column of A and let $\tilde{A} = aa^T$. Of course $\tilde{A} \in \mathcal{S}^n_+$ and so $\tilde{A} \in (FW^n_{n-1})^*$. We have

$$\langle \iota_I(B_I),\tilde{A}\rangle = \langle \iota_I(B_I),aa^T\rangle = \langle B_I,a_Ia_I^T\rangle.$$

But note that a_I is a column of A_I for $I \neq \{1,2,\dots,n-1\}$, so $A_I = a_Ia_I^T + A'_I$ for some $A'_I \succeq 0$ and $\langle B_I,A_I\rangle = 0$ implies $\langle B_I,a_Ia_I^T\rangle = 0$. Since we know already $B_{\{1,2,\dots,n-1\}} = 0$ we get $\langle B,\tilde{A}\rangle = 0$. Now evidently $\tilde{A}$ is not a multiple of A so it does not span the same ray and we have a contradiction.

For the reverse direction, assume that A does not span an extreme ray of $(FW^n_{n-1})^*$ which means that we can write it as

$$A = \gamma X + (1-\gamma)Y \text{ for some } X,Y \in (\widetilde{FW}^n_{n-1})^* \text{ and } \gamma \in [0,1]$$

where $(\widetilde{FW}^n_{n-1})^*$ is the compact section of the cone $(FW^n_{n-1})^*$ where every matrix in this section has the same trace as matrix A.

Let $X_\lambda = \lambda X + (1-\lambda)Y$, $\lambda \in \mathbb{R}_+$. Given some I, we know that $(X_\lambda)_I$ has rank at most $n-2$, in fact, there is a 2 dimensional space, $ker(A_I)$, which is always contained in $ker(X_\lambda)_I$. Then the set $L = \{\lambda | X_\lambda \in (FW^n_{n-1})^*\} = [\lambda_{min},\lambda_{max}]$ since $L \cap (FW^n_{n-1})^* \subseteq (\widetilde{FW}^n_{n-1})^*$ which is compact. The eigenvalues and eigenvectors of $(X_\lambda)_I$ change continuously with λ. Since two zero eigenvalues correspond to fixed eigenvectors, the only way for $(X_\lambda)_I$ to stop being psd is if a third eigenvalue switches from positive to negative, which implies that for some I, $rank((X_{\lambda_{max}})_I) \leq n-3$ and the same for $(X_{\lambda_{min}})_I$ and this means by Lemma 3.2.3 that both are psd. Hence A is psd since it is a convex combination of both. ■

In the following observation, we observe that conjugate permutation and scaling of a matrix does not affect extreme rays.

Observation 1. *Let D be a positive definite $n \times n$ diagonal matrix and P a $n \times n$ permutation matrix. Then*

a. The operation $\bullet \mapsto D \bullet D$ defines a bijection from $\mathbb{R}^{n\times n}$ onto itself which also induces bijections from $\mathcal{S}^n$ onto itself and from $\mathcal{S}^n_+$ onto itself and similarly bijections of the families of extreme rays of these cones onto themselves.

b. The cones FW^n_k and $(FW^n_k)^*$ are by $\bullet \mapsto D \bullet D$ also bijectively mapped onto themselves and analogous claims are true for the families of respective extreme rays.

c. The claims of parts a and b remain literally true if we replace in them the corresponding operation by $\bullet \mapsto P^T \bullet P$.

Proof. a. The operation $\bullet \mapsto D \bullet D$ is evidently a map of $\mathbb{R}^{n\times n}$ to itself. Since with D also D^{-1} is a well defined and positive definite diagonal matrix, the map $\bullet \mapsto D^{-1} \bullet D^{-1}$ is evidently the inverse of the former map on space $\mathbb{R}^{n\times n}$. It is clear that these maps preserve symmetry. Also if $x \in \mathbb{R}^n$, and $A \in \mathcal{S}^n_+$, then $x^T DADx = x^T D^T ADx = (Dx)^T A(Dx) \geq 0$. So the cone $\mathcal{S}^n_+$ is also preserved. Take next a matrix, E, say that spans an extremal ray of $\mathcal{S}^n_+$. If we had $DED = E_1 + E_2$ for two matrices E_1, E_2 that are not multiples of DED then they are not multiples of each other, and we have $E = D^{-1}E_1D^{-1} + D^{-1}E_2D^{-1}$ where the matrices on the right are in S^n_+ and are not multiples of each other and hence are not multiples of E. Hence E would not be extreme. A similar argument goes for the extreme rays of $\mathcal{S}^n$.

b. If $A \in FW^n_k$ then A is a sum of k matrices of the form $\iota_I(B)$ with $I \subset [n], |I| = k$ and B is a positive semidefinite $k \times k$ matrix. It is clear that $D\iota_I(B)D = \iota_I(D_IBD_I)$ and that D_IBD_I is positive semidefinite again. Consequently DAD is again sum of k matrices $\iota_I(B')$ with B' all positive semidefinite. So $DAD \in FW^n_k$. Similarly as above we see that our operation will also yield a bijection from the extreme rays of FW^n_k to itself. Finally if $A \in (FW^n_k)^*$ then all its principal $k \times k$ matrices are positive semidefinite. So the arguments above can also serve to show the invariance of $(FW^n_k)^*$ and the invariance of the set of extreme rays.

c. The proofs we gave in parts a and b can with minimal changes be transferred to valid proofs for the map $\bullet \mapsto P^T \bullet P$. Observe that $P^T P = I$. ∎

3.2.1 Characterizing extreme rays of $(FW^4_3)^*$

We start this section with the following lemma

Lemma 3.2.5. *If Q is positive semidefinite and $Q \notin FW^4_3$ then there exists a symmetric 4×4 matrix B and a positive definite diagonal matrix D such that*

i. B spans an extreme ray in $(FW^4_3)^*$

ii. B has the diagonal only entries all equal to 1

iii. $\langle DQD, B\rangle < 0$.

Proof. Suppose first that for all $B \in (FW_3^4)^*$ we had $\langle Q,B\rangle \geq 0$. This would show by definition of dual cones, that $Q \in (FW_3^4)^{**}$. But we know by Proposition 3.1.3 that $(FW_3^4)^{**} = FW_3^4$. So we get a contradiction. So there exists a matrix $B \in (FW_3^4)^*$ such that $\langle Q,B\rangle < 0$. Now every matrix in $(FW_3^4)^*$ is a finite positive linear combination of some matrices that span extreme rays of $(FW_3^4)^*$. Hence for at least one of these extreme-ray- defining matrices whose combination is B, taken at the place of B, we again must have the inequality. We call this extremal matrix now B.

By hypothesis $Q \in \mathcal{S}_+^4$; so $\langle Q,B\rangle < 0$, implies $B \notin \mathcal{S}_+^4$. Since every diagonal entry of B is diagonal entry of some principal 3×3 submatrix of B, and these submatrices are positive semidefinite, the diagonal entries of B are all nonnegative. Assume now that some diagonal entry, say $b_{11} = 0$. Then by a standard argument, see e.g. [30, p400c3], all the entries of column 1 and row 1 would be 0. The nonzero entries of B are thus found in B_{234}, which is positive semidefinite. Hence B is psd, a contradiction.

Thus we have $b_{11}, b_{22}, b_{33}, b_{44} > 0$ and the diagonal matrix $D = \mathrm{Diag}(b_{11}^{-1/2}, b_{22}^{-1/2}, b_{33}^{-1/2}, b_{44}^{-1/2})$ is well defined. By Observation 1, the matrix $B' = DBD$ will be again an extreme ray of $(FW_3^4)^*$ and it is clear that $B' = (b_{ii}^{-1/2} b_{ij} b_{jj}^{-1/2})_{i,j=1}^4$ is a matrix which has only ones on the diagonal. Finally $\langle D^{-1}QD^{-1}, B'\rangle = \langle Q,B\rangle < 0$. Thus renaming D^{-1}, B' to D, B, respectively, we get the claim. ∎

Based on the results that we have proven so far, we can fully characterize the extreme rays of $(FW_3^4)^*$.

Proposition 3.2.6. *Let B be a symmetric 4×4 not positive semidefinite matrix which spans an extreme ray of $(FW_3^4)^*$, then for some $a, c \in]-\pi,\pi[\setminus\{0\}$, some permutation P and some nonsingular diagonal matrix D, the matrix B has the following form*

$$DPBP^TD^T = \begin{bmatrix} 1 & \cos(a) & \cos(a-c) & \cos(c) \\ \cos(a) & 1 & \cos(c) & \cos(a-c) \\ \cos(a-c) & \cos(c) & 1 & \cos(a) \\ \cos(c) & \cos(a-c) & \cos(a) & 1 \end{bmatrix}.$$

Proof. First note that by the considerations of the previous lemma, we can always assume a scaling that takes all diagonal entries of B to 1. Furthermore, by assumption, $B \in (FW_3^4)^*$ which means all of its 3×3 and accordingly its 2×2 principal submatrices are psd, hence for all $i,j \in \{1,2,3,4\}$, $0 \leq b_{ii}b_{jj} - b_{ij}^2 = 1 - b_{ij}^2$ and hence $b_{ij}^2 \leq 1$ for all pairs (i,j). Therefore, using that the image of the cossine function is $[-1,1]$, we can write B as

$$B = \begin{bmatrix} 1 & \cos(a) & \cos(b) & \cos(c) \\ \cos(a) & 1 & b_{23} & b_{24} \\ \cos(b) & b_{23} & 1 & b_{34} \\ \cos(c) & b_{24} & b_{34} & 1 \end{bmatrix},$$

for some $a, b, c \in]-\pi,\pi[$. Now since B spans an extreme ray of $(FW_3^4)^*$, by Theorem 3.2.4 all of its 3×3 principal submatrices have rank 2 and hence have zero determinant. Hence by starting with principal submatrix B_{123}, we have

$$0 = \det\left(\begin{bmatrix} 1 & \cos(a) & \cos(b) \\ \cos(a) & 1 & b_{23} \\ \cos(b) & b_{23} & 1 \end{bmatrix}\right) = 1 - b_{23}^2 - \cos(a)^2 + 2b_{23}\cos(a)\cos(b) - \cos(b)^2.$$

By solving this quadratic equation for b_{23} one finds

$$\begin{aligned} b_{23} &\in \{\cos(a)\cos(b) \pm \sqrt{1-\cos(a)^2-\cos(b)^2+\cos(a)^2\cos(b)^2}\} \\ &= \{\cos(a)\cos(b) \pm \sqrt{(1-\cos^2(a))(1-\cos^2(b))}\} \\ &= \{\cos(a)\cos(b) \pm \sin(a)\sin(b)\} \\ &= \{\cos(a \mp b)\}. \end{aligned}$$

We do completely analogous calculations for principal submatrices B_{134} and B_{124} and obtain $b_{34} \in \{\cos(b \pm c)\}$ and $b_{24} \in \{\cos(a \pm c)\}$, respectively. Now we have eight matrices that emerge from choosing one of the symbols $+$ or $-$ in each of the patterns $a \pm b, a \pm c, b \pm c$ existent in the matrix below by taking care that the symmetry of the matrix is preserved.

$$\begin{bmatrix} 1 & \cos(a) & \cos(b) & \cos(c) \\ \cos(a) & 1 & \cos(a \pm b) & \cos(a \pm c) \\ \cos(b) & \cos(a \pm b) & 1 & \cos(b \pm c) \\ \cos(c) & \cos(a \pm c) & \cos(b \pm c) & 1 \end{bmatrix}.$$

The following table indicates in the first column the possible selections of signs in $a \pm b, a \pm c, b \pm c$, respectively; and in the second column and the third column the determinants of the respective matrices B_{234} and B.

$x \pm y$	$\det(B_{234})$	$\det(B)$
$+,+,+$	$4\sin(a)\sin(b)\sin(c)\sin(a+b+c)$	$-4\sin(a)^2\sin(b)^2\sin(c)^2$
$+,+,-$	0	0
$+,-,+$	0	0
$+,-,-$	$-4\sin(a)\sin(b)\sin(a+b-c)\sin(c)$	$-4\sin(a)^2\sin(b)^2\sin(c)^2$
$-,+,+$	0	0
$-,+,-$	$-4\sin(a)\sin(b)\sin(c)\sin(a-b+c)$	$-4\sin(a)^2\sin(b)^2\sin(c)^2$
$-,-,+$	$4\sin(a)\sin(b)\sin(a-b-c)\sin(c)$	$-4\sin(a)^2\sin(b)^2\sin(c)^2$
$-,-,-$	0	0

Now assume one of the reals a,b,c is 0 or $\pm\pi$. Then the table shows that all entries in columns two and three vanish. Hence the matrix B in this case is positive semidefinite. Thus in order that B, as required, is not positive semidefinite it is necessary that $a,b,c \neq \{-\pi, 0, \pi\}$. In this case column 3 guarantees we get a not positive semidefinite matrix B in exactly the cases of the sign choices $+++, +--, -+-, --+$ for $a \pm b, a \pm c, b \pm c$, respectively. The matrices corresponding to rows, 2,3,5,8 of the table are positive semidefinite independent of choices a,b,c.. Explicitly this means that

B must be one of the following four matrices

$$\begin{bmatrix} 1 & \cos(a) & \cos(b) & \cos(c) \\ \cos(a) & 1 & \cos(a+b) & \cos(a+c) \\ \cos(b) & \cos(a+b) & 1 & \cos(b+c) \\ \cos(c) & \cos(a+c) & \cos(b+c) & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & \cos(a) & \cos(b) & \cos(c) \\ \cos(a) & 1 & \cos(a+b) & \cos(a-c) \\ \cos(b) & \cos(a+b) & 1 & \cos(b-c) \\ \cos(c) & \cos(a-c) & \cos(b-c) & 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & \cos(a) & \cos(b) & \cos(c) \\ \cos(a) & 1 & \cos(a-b) & \cos(a+c) \\ \cos(b) & \cos(a-b) & 1 & \cos(b-c) \\ \cos(c) & \cos(a+c) & \cos(b-c) & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & \cos(a) & \cos(b) & \cos(c) \\ \cos(a) & 1 & \cos(a-b) & \cos(a-c) \\ \cos(b) & \cos(a-b) & 1 & \cos(b+c) \\ \cos(c) & \cos(a-c) & \cos(b+c) & 1 \end{bmatrix}.$$

Note by substituting the letter c by $-c$ in the left upper matrix we get the right upper matrix because $\cos(-c) = \cos(c)$. Exactly the same remark leads from the left lower matrix to the right lower matrix. Finally note that after doing the transpositions of rows and columns $3,4$, the upper left matrix shown takes the form

$$\begin{bmatrix} 1 & \cos(a) & \cos(c) & \cos(b) \\ \cos(a) & 1 & \cos(a+c) & \cos(a+b) \\ \cos(c) & \cos(a+c) & 1 & \cos(b+c) \\ \cos(b) & \cos(a+b) & \cos(b+c) & 1 \end{bmatrix}$$

and after changing the name of variable c to $-b$ and of variable b to c and noting that $\cos(b-c) = \cos(c-b)$ we see we have obtained the following matrix

$$\begin{bmatrix} 1 & \cos(a) & \cos(b) & \cos(c) \\ \cos(a) & 1 & \cos(a-b) & \cos(a+c) \\ \cos(b) & \cos(a-b) & 1 & \cos(c-b) \\ \cos(c) & \cos(a+c) & \cos(c-b) & 1 \end{bmatrix}.$$

Hence we have one form and its possible permutations. Now, we know that the determinant of the submatrix B_{234} is $4\sin(a)\sin(b)\sin(a-b-c)\sin(c)$. We know by Theorem 3.2.4 that all 3×3 principal minors must vanish, so $det(B_{234}) = 0$ which happens if and only if $b = a - c + k\pi$, $k \in \mathbb{Z}$. Substituting this in the start matrix B we get the following two forms

$$\begin{bmatrix} 1 & \cos(a) & \delta\cos(a-c) & \cos(c) \\ \cos(a) & 1 & \delta\cos(c) & \cos(a-c) \\ \delta\cos(a-c) & \delta\cos(c) & 1 & \delta\cos(a) \\ \cos(c) & \cos(a-c) & \delta\cos(a) & 1 \end{bmatrix},$$

with $\delta = \pm 1$. But note that these are the same up to scaling by a diagonal matrix, so we may assume $\delta = 1$, finishing the proof. ■

3.3 Factor width k matrices and sums of k-nomial squares polynomials

Ahmadi and Majumdar in [3] considered the following example

$$p_n^a = (\sum_{i=1}^n x_i)^2 + (a-1)\sum_{i=1}^n x_i^2$$

when $n = 3$ and proved that if $a < 2$ then no nonnegative integer r can be chosen so that $(x_1^2 + x_2^2 + x_3^2)^r p_3^a$ is a sum of squares of binomials, although it is clearly nonnegative for $a \geq 1$.

In this section, we give negative results along the same direction. We first characterize when p_n^a is a sum of k-nomial squares, then we show that $p_{n,r}^a$, that is, the multiplication of p_n^a with $(\sum_{i=1}^n x_i^2)^r$, is a sum of k-nomial squares if and only if this is the case for $r = 0$. Before presenting our proof, we make the connection between factor width k matrices and sums of k-nomial squares polynomials which will be used along the proof. The following proposition states this connection

Proposition 3.3.1. *A multivariate polynomial $p(x)$ of degree $2d$ is a sum of k-nomial squares (soks) if and only if it can be written in the form $p(x) = z(x)_d^T Q z(x)_d$ with matrix $Q \in FW_k^{\binom{n+d}{d}}$.*

Proof. Consider an expression $a_1m_1 + \cdots + a_km_k$ with reals $a_1, \ldots, a_k$ and monomials $m_1, \ldots, m_k$. Note that monomials $m_1, \ldots, m_k$ occur necessarily in the column $z(x)_d$ at positions $i_1, \ldots, i_k$, say. Construct a column q of size $\binom{n+d}{d}$ by putting into positions $i_1, \ldots, i_k$ respectively the reals $a_1, \ldots, a_k$, and into all other positions 0s. Then evidently $z(x)_d^T q = a_1m_1 + \cdots + a_km_k$, and consequently $z(x)_d^T qq^T z(x)_d = (a_1m_1 + \cdots + a_km_k)^2$. Consequently, a polynomial which is a sum of, say, t squares of k-nomial can be written as $z(x)_d^T Q z(x)_d$, where $Q = \sum_{v=1}^t q_v q_v^T$, with suitable columns $q_1, \ldots, q_t$ of size $\binom{n+d}{d}$ each of which has at most k nonzero entries. It follows that Q is a matrix of factor width k. Conversely if Q is of factor width k, then we already know we can write $Q = \sum_{q \in Q} qq^T$ where each column has at most k nonzero real entries. Clearly from the arguments above follows now that $z(x)_d^T Q z(x)_d$ yields a polynomial which is a finite sum of k-nomial squares. ∎

We shall also need the following proposition.

Lemma 3.3.2. *Consider a quadratic form $q(x) = x^T Q x$ and a polynomial p related to q by $p = (\sum_{i=1}^n (\lambda_i x_i)^2)^r q$. Then every monomial of p has at most two odd degree variables and we have $p_{(i,j)} = 2(\sum_{i=1}^n \lambda_i^2)^r q_{ij}$ and $p_0 = (\sum_{i=1}^n \lambda_i^2)^r\, tr(Q)$ where $p_{(i,j)}$ is the sum of coefficients of the monomials in which x_i and x_j have odd degree and p_0 is the sum of coefficients of even monomials of p and q_{ij} is the coefficient of x_ix_j in q, $i < j$.*

Proof. The quadratic form is

$$q(x) = \sum_{1 \leq i,j \leq n} x_i q_{ij} x_j = \sum_{i=1}^n q_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2q_{ij} x_i x_j,$$

while by the multinomial theorem [14] we have

$$((\lambda_1 x_1)^2 + \cdots + (\lambda_n x_n)^2)^r = \sum_{i_1 + \cdots + i_n = r} \binom{r}{i_1, \ldots, i_n} (\lambda_1 x_1)^{2i_1} (\lambda_2 x_2)^{2i_2} \ldots (\lambda_n x_n)^{2i_n}.$$

Thus, putting the λs into evidence, by definition of p, we get

$$
\begin{aligned}
p = & \sum_{(i,\underline{i}) \in J_1} q_{ii} \binom{r}{\underline{i}} \lambda_1^{2i_1} \cdots \lambda_i^{2i_i} \cdots \lambda_n^{2i_n} \cdot x_1^{2i_1} \cdots x_i^{2i_i+2} \cdots x_n^{2i_n} \\
& + \sum_{((i,j),\underline{i}) \in J_2} 2q_{ij} \binom{r}{\underline{i}} \lambda_1^{2i_1} \cdots \lambda_n^{2i_n} \cdot x_1^{2i_1} \ldots x_i^{2i_i+1} \ldots x_j^{2i_j+1} \ldots x_n^{2i_n}
\end{aligned}
$$

where $\underline{i} = (i_1, \ldots, i_n)$ and, with $|\underline{i}| = i_1 + \cdots + i_n$,

$$
\begin{aligned}
J_1 &= \{(i,\underline{i}) : i \in \{1, \ldots, n\}, \underline{i} \in \mathbb{Z}^n_{\geq 0}, |\underline{i}| = r\}, \\
J_2 &= \{((i,j),\underline{i}) : 1 \leq i < j \leq n, \underline{i} \in \mathbb{Z}^n_{\geq 0}, |\underline{i}| = r\}.
\end{aligned}
$$

From the above equation for p , we recognize that

$$
p_{(i,j)} = 2q_{ij} \sum_{i_1+\cdots+i_n=r} \binom{r}{i_1, \ldots, i_n} \lambda_1^{2i_1} \cdots \lambda_i^{2i_i} = 2q_{ij}(\lambda_1^2 + \cdots + \lambda_n^2)^r,
$$

again by the multinomial theorem; and similarly we have

$$
\begin{aligned}
p_0 = \sum_{i=1}^{n} \sum_{i_1+\cdots+i_n=r} q_{ii} \binom{r}{i_1, \ldots, i_n} \lambda_i^{2i_i} \cdots \lambda_n^{2i_n} &= \sum_{i=1}^{n} q_{ii} \sum_{i_1+\cdots+i_n=r} \binom{r}{i_1, \ldots, i_n} \lambda_i^{2i_i} \cdots \lambda_n^{2i_n} \\
&= (\lambda_1^2 + \cdots + \lambda_n^2)^r trace(Q).
\end{aligned}
$$

■

In addition, we find the following useful fact proved in Muir's treatise [39, p 61].

Lemma 3.3.3. *For the determinant at the left hand side below which has only a's except on the diagonal, we have*

$$
\begin{vmatrix}
b_1 & a & \ldots & a \\
a & b_2 & \ldots & a \\
\vdots & \vdots & \ddots & \vdots \\
a & a & \ldots & b_n
\end{vmatrix}
= \prod_{i=1}^{n} (b_i - a) + a \sum_{j=1}^{n} \prod_{i:i \neq j}^{n} (b_i - a)
$$

Now we are ready to prove our results regarding Ahmadi and Majumdar's example.

Proposition 3.3.4. *If $a \geq \frac{n-1}{k-1}$, then p_n^a is a sum of k-nomial squares.*

Proof. The quadratic form p_n^a, can be written as $a\sum_{i=1}^n x_i^2 + 2\sum_{i<j} x_i x_j$, so $p = z(x)^T Q z(x)$ by means of the $n \times n$ matrix Q shown.

$$
Q = \begin{bmatrix}
a & 1 & \cdots & 1 & 1 \\
1 & a & \cdots & 1 & 1 \\
\vdots & & \ddots & & \vdots \\
1 & & & & 1 \\
1 & 1 & \cdots & 1 & a
\end{bmatrix}.
$$

Now there exist $\binom{n}{k}$ subsets K of cardinality k of the set $\{1,2,\ldots,n\}$. Let $i,j \in \{1,2,\ldots,n\}$. A pair (i,i) lies in exactly $\binom{n-1}{k-1}$ of the sets $K\times K$ while a pair (i,j) with $i\neq j$ lies in $K\times K$ if and only if $\{i,j\}\subseteq K$. It hence lies in exactly $\binom{n-2}{k-2}$ sets $K\times K$. Consider the $k\times k$ matrix B as following

$$B = \binom{n-2}{k-2}^{-1} \begin{bmatrix} \frac{(k-1)a}{n-1} & 1 & \cdots & 1 & 1 \\ 1 & \frac{(k-1)a}{n-1} & \cdots & 1 & 1 \\ \vdots & & \ddots & & \vdots \\ 1 & & & & 1 \\ 1 & 1 & \cdots & 1 & \frac{(k-1)a}{n-1} \end{bmatrix},$$

and define $\iota_K(B)$ to be the $n\times n$ matrix of support $K\times K$ which carries on it the matrix B. Then our arguments yield that $\sum_{K:|K|=k}\iota_K(B) = Q$.

Take an arbitrary $l\times l$ submatrix of the matrix factor of B. By the previous lemma , this submatrix has determinant $(\frac{(k-1)a}{n-1}-1)^{-1+l}(\frac{(k-1)a}{n-1}-1+l)$. It follows from the hypothesis for a that this determinant is nonnegative. So B, and thus $\iota_K(B)$, is a positive semidefinite matrix and Q hence a matrix of factor width $\leq k$ by Proposition 3.1.2. This means by Proposition 3.3.1 that p_n^a is a sum of k-nomial squares. ∎

Theorem 3.3.5. *For integers $n\geq 0$ and $r\geq 0$, define*

$$p_{n,r}^a = (\sum_{i=1}^n x_i^2)^r\, p_n^a = (\sum_{i=1}^n x_i^2)^r \cdot \left((\sum_{i=1}^n x_i)^2 + (a-1)\sum_{i=1}^n x_i^2 \right).$$

Then $p_{n,r}^a$ is a sum of k-nomial squares if and only if $p_n^a = p_{n,0}^a$ is a sum of k-nomial squares.

Proof. Clearly, if p_n^a is a so$k$s then $p_{n,r}^a$ is a so$k$s. So we need to show the inverse. Assume that the degree $2(r+1)$ polynomial $p_{n,r}^a$ is a so$k$s. Let $I_{n,r+1} = \{(i_1,\ldots,i_n)\ s.t.\ i_k\in\mathbb{N}_0, \sum_{k=1}^n i_k = r+1\}$ be the set of vectors of exponents in $\mathbb{Z}_{\geq 0}^n$ that occurs in the family of monomials of a homogeneous polynomial of degree $r+1$ in variables $x_1,\ldots,x_n$. Let this family of monomials be also the one that occurs in $z(x)_{r+1}$.

By Proposition 3.3.1, we can write

$$p_{n,r}^a = z(x)_{r+1}^T H_{n,r} z(x)_{r+1} \text{ for some } H_{n,r}\in FW_k^{\binom{n+r}{r+1}}.$$

Call an $i\in\mathbb{Z}_{\geq 0}^n$ *even* if it has only even entries and consider now the matrix $B_{n,r}\in\mathbb{R}^{I_{n,r+1}\times I_{n,r+1}}$ given by

$$(B_{n,r})_{ij} = \begin{cases} k-1 & \text{if } i+j \text{ is even,} \\ -1 & \text{otherwise.} \end{cases}$$

We will show now that $B_{n,r}\in (FW_k^{\binom{n+r}{r+1}})^*$; that is we shall prove that every $k\times k$ principal submatrix of $B_{n,r}$ is positive semidefinite, see Proposition 3.1.3. Since n,r are fixed, we write B and H for matrices $B_{n,r}, H_{n,r}$ respectively.

Note that a sum $i+j$ of such n-uples is even if and only if the sets of positions in i where odd entries occur equals the corresponding set in j. (Example: The 5-uple $i=(1,0,0,3,2)$ has $\{1,4\}$ as the set of positions of odd entries.)

So take a $k\times k$ submatrix M of B with rows and columns indexed by the n-uples $i_1,\dots,i_k$, say. Determine for each n-uple its set of positions of odd entries. Let $S_1,\dots,S_l$ $(l\le k)$ be the distinct non empty sets of such positions. Now rearrange the n-uples so that the first few n-uples each have S_1 as set of positions of odd entries, the next few have S_2 as such set of positions, etc. Let $s_1,\dots,s_l$ be the sizes of these sets. To the rearrangement of the n-uples corresponds a $k\times k$ permutation matrix P such that PMP^T is 'a direct sum of blocks of sizes $s_1\times s_1,\dots,s_l\times s_l$ with entries $k-1$ over a background of -1s'. Formally, for suitable P we can express this as

$$PMP^T=(-1)J_k+k(J_{s_1}\oplus J_{s_2}\oplus\cdots\oplus J_{s_l})$$

This same matrix can be produced as follows. Define $l\times l$ matrix N and $l\times k$ matrix C by

$$N=(-1)J_l+kI_l=\begin{bmatrix} k-1 & -1 & \dots & -1\\ -1 & k-1 & \dots & -1\\ \vdots & & \ddots & \\ -1 & & \dots & k-1\end{bmatrix},$$

$$C=\begin{pmatrix} 1 & 1 & \cdots & 1 & & & & & & & & \\ & & & & 1 & 1 & \cdots & 1 & & & & \\ & & & & & & & & \dots & & & \\ & & & & & & & & & 1 & 1 & \cdots & 1\end{pmatrix}$$

where rows, $1,2,...,l$ of C have, respectively, $s_1,s_2,...,s_l$ entries equal to 1. Check that then $PMP^T=C^TNC$. Now, again by Lemma 3.3.3, N is positive semidefinite, Hence M will be psd. Since the $k\times k$ submatrix M of B was arbitrary, we are done with proving that $B\in(FW_k^{\binom{n+r}{r+1}})^*$. By definition of the concept of a dual cone, we have $\langle B,H\rangle\ge 0$, and by the definitions of $\langle,\rangle$, H and B, hence

$$\langle B,H\rangle=(k-1)\sum_{i,j:i+j\text{ even}}h_{ij}+(-1)\sum_{i,j:i+j\text{ non-even}}h_{ij}\ge 0.$$

Since the quadratic form underlying our construction of $p^a_{n,r}$ is $p^a_n=a\sum_{i=1}^n x_i^2+2\sum_{i<j}x_ix_j$, and it has defining matrix Q mentioned in the previous proposition, we get by Lemma 3.3.2 that

$$\sum_{i,j:i+j\text{ even}}h_{ij}=n^r trQ=n^{r+1}a,\quad\text{and}\sum_{i,j:i+j\text{ non-even}}h_{ij}=2n^r\times\sum_{1\le i<j\le n}q_{ij}=2n^r\frac{1}{2}n(n-1)=n^{r+1}(n-1).$$

Hence the inequality above reads $(k-1)n^{r+1}a\ge n^{r+1}(n-1)$ or $a\ge\frac{n-1}{k-1}$, which means by the previous proposition that p^a_n is a sum of k-nomial squares. ■

3.4 Factor width 2 matrices and sum of binomial squares

For the case of $k = 2$, sums of squares of k-nomials are also known as sums of binomial squares [23] or scaled diagonally dominant sums of squares (SDSOS) [3]. In this section we will try to generalize Ahmadi And Majumdar's [3] counterexample in this setting. We prove that a quadratic form is a r-sobs, if and only if it is a sobs. But before we proceed further, we shall need the following proposition.

Proposition 3.4.1. *[23, Corollary 2.9] Given a quadratic form $P(x) = \sum_{i=1}^{n} q_i x_i^2 + \sum_{i<j} q_{ij} x_i x_j$, then if $F(x) = \sum_{i=1}^{n} q_i x_i^2 - \sum_{i<j} |q_{ij}| x_i x_j$ is nonnegative, $P(x)$ is sum of binomial squares.*

Theorem 3.4.2. *Let $q(x) = q(x_1, \ldots, x_n)$ be a real quadratic form and let $r \in \mathbb{Z}_{\geq 0}$. Then if $q(x)(x_1^2 + \cdots + x_n^2)^r$ is sum of squares of binomials, so is $q(x)$ itself.*

Proof. Assume that $q(x)(x_1^2 + \cdots + x_n^2)^r$ is a sum of binomial squares. We will prove that $q(x)$ is a sum of binomial squares. Write $q(x) = \sum_{i=1}^{n} a_i x_i^2 + \sum_{1 \leq i < j \leq n} d_{ij} x_i x_j$, say. Then considerations as in the proof of Lemma 3.3.2 yield

$$\begin{aligned} q(x).(x_1^2 + \ldots, + x_n^2)^r = &\sum_{(i,\underline{i}) \in J_1} a_i \binom{r}{\underline{i}} x_1^{2i_1} \cdots x_i^{2i_i+2} \cdots x_n^{2i_n} \\ &+ \sum_{((i,j),\underline{i}) \in J_2} d_{ij} \binom{r}{\underline{i}} x_1^{2i_1} \cdots x_i^{2i_i+1} \cdots x_j^{2i_j+1} \cdots x_n^{2i_n}, \end{aligned}$$

where again,

$$J_1 = \{(i,\underline{i}) : i \in \{1,\ldots,n\}, \underline{i} \in \mathbb{Z}_{\geq 0}^n, |\underline{i}| = r\}, \quad J_2 = \{((i,j),\underline{i}) : 1 \leq i < j \leq n, \underline{i} \in \mathbb{Z}_{\geq 0}^n, |\underline{i}| = r\}.$$

Now the monomials of degree r are of the form $x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ with $i_1 + \cdots + i_n = r$. There are as we know $L = \binom{r+n-1}{r}$ such monomials. We order these and denote them by $m_1, \ldots, m_L$. Every binomial is of the form $(\alpha_{ij} m_i + \beta_{ij} m_j)$ with some selection of i, j with $1 \leq i < j \leq L$. By possibly redefining α_{ii}, we can thus assume the binomials are of the form $\alpha_{ii} m_i$, $1 \leq i \leq L$ or $(\alpha_{ij} m_i + \beta_{ij} m_j)$ with $1 \leq i < j \leq L$. A sum of binomial squares is thus given as

$$\begin{aligned} \sum_{i=1}^{L} \alpha_{ii}^2 m_i^2 + \sum_{1 \leq i < j \leq L} (\alpha_{ij} m_i + \beta_{ij} m_j)^2 &= \sum_{i=1}^{L} \alpha_{ii}^2 m_i^2 + \sum_{i<j} \alpha_{ij}^2 m_i^2 + \sum_{i<j} \beta_{ij}^2 m_j^2 + \sum_{1 \leq i < j \leq L} 2\alpha_{ij}\beta_{ij} m_i m_j \\ &= \sum_{i=1}^{L} (\alpha_{ii}^2 + \alpha_{i,i+1}^2 + \ldots + \alpha_{iL}^2 + \beta_{1i}^2 + \ldots \beta_{i-1,i}^2) m_i^2 + \sum_{1 \leq i < j \leq L} 2\alpha_{ij}\beta_{ij} m_i m_j. \end{aligned}$$

Now assuming, as we do, that $q(x)(x_1^2 + \cdots + x_n^2)^r$ is a sobs, by means of comparison of coefficients, we get a system of $|J_1| + |J_2|$ equations between reals. It is easily seen that these can be obtained as follows: For each $(i,\underline{i}) \in J_1$ define

$$T(i,\underline{i}) = \{\text{indices } t \in \{1,\ldots,L\} \text{ for which } m_t^2 = x_1^{2i_1} \cdots x_i^{2i_i+2} \cdots x_n^{2i_n}\},$$

$$S(i,\underline{i}) = \{\text{index pairs } s_1 < s_2 \text{ for which } m_{s_1} m_{s_2} = x_1^{2i_1} \cdots x_i^{2i_i+2} \cdots x_n^{2i_n}\}$$

and write the equation

$$a_i \binom{r}{\underline{i}} r_{i_1,\ldots,i_n} = \sum_{t\in T(i,\underline{i})} (\alpha_{tt}^2 + \ldots + \alpha_{tL}^2 + \beta_{1t}^2 + \ldots + \beta_{t-1,t}^2) + \sum_{(s_1,s_2)\in S(i,\underline{i})} 2\alpha_{s_1s_2}\beta_{s_1s_2}$$

for each $((i,j),\underline{i}) \in J_2$, let

$$S'((i,j),\underline{i}) = \{\text{index pairs } s_1' < s_2' \text{ for which } m_{s_1'} m_{s_2'} = x_1^{2i_1} \ldots x_i^{2i_i+1} \ldots x_j^{2i_j+1} \ldots x_n^{2i_n}\},$$

and write the equation

$$d_{ij} \binom{r}{\underline{i}} = \sum_{s_1',s_2' \in S'((i,j),\underline{i})} 2\alpha_{s_1's_2'}\beta_{s_1's_2'}.$$

Every system of reals $(\{a_i\}_{i=1}^n, \{d_{ij}\}_{i,j=1}^n, \{\alpha_{ij}\}_{1\le i\le j\le L}, \{\beta_{ij}\}_{1\le i<j\le L})$ which satisfies the system of equations gives rise to a quadratic form q and binomials so that $q(x)(x_1^2 + \cdots + x_n^2)^r$ is sum of squares of these binomials and if we have a system of reals satisfying the system, then we can find a particular new solution by replacing the d_{ij} which are positive by $-d_{ij}$ and simultaneously replacing the $\beta_{s_1's_2'}$ for which $s_1', s_2' \in S'((i,j),\underline{i})$ by $-\beta_{s_1's_2'}$. Indeed note that the sets $S'((i,j),\underline{i})$ are disjoint from the sets $S(i,\underline{i})$ and $(-\beta_{s_1's_2'})^2 = (\beta_{s_1's_2'})^2$, hence the first set of $|J_1|$ equations will again be satisfied. In what concerns the second set of equations we note that the sets $S'((i,j),\underline{i})$ are also mutually disjoint, because a choice (s_1', s_2') defines via forming $m_{s_1'} m_{s_2'}$ a unique power product $x_1^{2i_1} \ldots x_i^{2i_i+1} \ldots x_j^{2i_j+1} \ldots x_n^{2i_n}$ with exactly two odd exponents determining i, j and then $\underline{i}$. In the other words (s_1', s_2') lives in only one of the sets $S'((i,j),\underline{i})$ hence carrying through the replacements indicated we change the sign at the left hand side of equation if and only if we change the sign of the corresponding right hand side. We therefore satisfy also the second group of equations.

The new solution tells us that $\hat{q}(x)(x_1^2 + \cdots + x_n^2)^r$ is sum of squares of binomials where $\hat{q}(x) = \sum_{i=1}^n a_i x_i^2 - \sum_{1\le i<j\le n} |d_{ij}| x_i x_j$. Now since the multiplier is evidently positive definite, $\hat{q}$ is nonnegative. Hence by Proposition 3.4.1, q is a sum of squares of binomials. ∎

The following example illustrates the theorem for a quadratic form in 3 variables and $r = 1$.

Example 3.4.1. *Let $q(x,y,z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz$ be a ternary quadratic form. Assume that $q.(x^2 + y^2 + z^2)$ is a sum of binomial squares, then q itself is a sobs. By direct computation we find that*

$$\begin{aligned}
&(x^2 + y^2 + z^2).q(x,y,z) \\
&= ax^4 + by^4 + cz^4 \\
&+ (a+b)x^2y^2 + (a+c)x^2z^2 + (b+c)y^2z^2 \\
&+ dx^3y + dxy^3 + ex^3z + fy^3z + exz^3 + fyz^3 \\
&+ dxyz^2 + exy^2z + fx^2yz
\end{aligned}$$

In such forms all possible monomials in three variables of degree 4 occur. In order to show these polynomials as sobs, think of 6 degree 2 monomials as laid out lexicographically xx, xy, xz,yy yz,zz. These monomials have ordinal 1,2,3,4,5,6 respectively. Any binomial is a linear combination of two of these monomials, the i-th and the j-th say, with $i \le j$. We index the coefficients of the monomials in such a binomial by α_{ij} and β_{ij}.

Thus, examples of our binomials are $\alpha_{12}x^2+\beta_{12}xy$ *and* $\alpha_{23}xy+\beta_{23}xz$*. Also the binomials* $\alpha_{ii}x^2+\beta_{ii}x^2$ *can be simplified to* $\alpha_{ii}x^2$ *via redefining* α_{ii} *so we will suppress* β_{ii} *throughout. There remain thus* $\binom{6}{2}+6=21$ *index pairs. Assume now a quartic form is a sum of the* 21 *squares of binomials. Then the monomials of this form have the following coefficients:*

$$
\begin{array}{lll}
x^4 & \alpha_{11}^2+\alpha_{12}^2+\alpha_{13}^2+\alpha_{14}^2+\alpha_{15}^2+\alpha_{16}^2 & =a \\
y^4 & \beta_{14}^2+\beta_{24}^2+\beta_{34}^2+\alpha_{44}^2+\alpha_{45}^2+\alpha_{46}^2 & =b \\
z^4 & \beta_{16}^2+\beta_{26}^2+\beta_{36}^2+\beta_{46}^2+\beta_{56}^2+\alpha_{46}^2 & =c \\
x^2y^2 & \beta_{12}^2+\alpha_{22}^2+\alpha_{23}^2+\alpha_{24}^2+\alpha_{25}^2+\alpha_{26}^2+ & 2\alpha_{14}\beta_{14}=a+b \\
x^2z^2 & \beta_{13}^2+\beta_{23}^2+\alpha_{33}^2+\alpha_{34}^2+\alpha_{35}^2+\alpha_{36}^2+ & 2\alpha_{16}\beta_{16}=c+a \\
z^2y^2 & \beta_{15}^2+\beta_{25}^2+\beta_{35}^2+\beta_{45}^2+\alpha_{55}^2+\alpha_{56}^2+ & 2\alpha_{46}\beta_{46}=c+b \\
x^3y & & 2\alpha_{12}\beta_{12}=d \\
xy^3 & & 2\alpha_{24}\beta_{24}=d \\
xyz^2 & & 2\alpha_{26}\beta_{26}+2\alpha_{35}\beta_{35}=d \\
x^3z & & 2\alpha_{13}\beta_{13}=e \\
xy^2z & & 2\alpha_{25}\beta_{25}+2\alpha_{34}\beta_{34}=e \\
xz^3 & & 2\alpha_{36}\beta_{36}=e \\
x^2yz & & 2\alpha_{23}\beta_{23}+2\alpha_{15}\beta_{15}=f \\
yz^3 & & 2\alpha_{56}\beta_{56}=f \\
y^3z & & 2\alpha_{45}\beta_{45}=f
\end{array}
$$

In this scheme at the right hand side you find also the respective coefficients of $q.(x^2+y^2+z^2)$*. So the following is clear:*

Every family of reals α*s,* β*s and* $a,b,\ldots,f$ *that satisfies the system of equations at the right hand side gives rise to a quadratic form* $q(x,y,z)$ *such that* $q.(x^2+y^2+z^2)$ *is a sobs, and conversely if* q *is a ternary quadratic form such that* $q.(x^2+y^2+z^2)^2$ *is a sobs it produces such a system. We also denote* $\hat{q}$ *as following*

$$\hat{q}=ax^2+by^2+cz^2-|d|xy-|e|xz-|f|yz$$

A quadratic form q *is a sobs if and only if* $\hat{q}$ *is positive semidefinite. So all we need to show is that* $\hat{q}$ *is positive semidefinite.*

In order to do this in our system of equations, whenever d,e,f *is positive, replace in the system the* $\beta_\star$*s that occur in the LHS of those equations that have* d,e,f *at RHS by* $-\beta_\star$*s and* d,e,f *by* $-d,-e,-f$*. This yield a new set of reals which satisfies the old system. The old* a,b,c *together with partially new* d,e,f *give a system that corresponds to saying that for* $\hat{q}$*, we have that* $\hat{q}.(x^2+y^2+z^2)$ *is a sobs. Hence it is a nonegative form. Now since* $x^2+y^2+z^2$ *is positive definite it follows that* $\hat{q}$ *is positive semidefinite and then* q *is a sobs.*

The Theorem 3.4.2 is true only for quadratics which means that in general it is not true for other polynomials with degree greater than two. The Motzkin polynomial is an example which shows this fact. In what follows we will show that Motzkin polynomial is neither a sobs nor a 1-sobs , but it is a

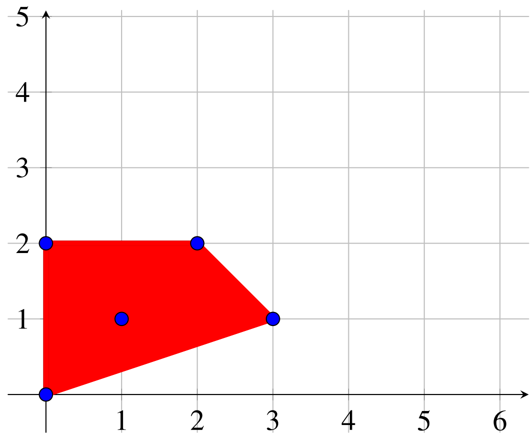

Fig. 3.1 Newton polytope of polynomial $p(x,y) = 5 - xy + x^2y^2 - 5y^2 + x^3y$

2-sobs. In our proof, we shall need the concept of Newton polytope for polynomials which is defined as following.

Definition 7. *Consider a multivariate polynomial $p(x_1,\dots,x_n) = \sum_\alpha c_\alpha x^\alpha$. The Newton polytope of p, denoted by $New(p)$, is the convex hull of the set of exponents α, considered as vectors in $\mathbb{R}^n$.*

Example 3.4.2. *As an example consider the polynomial $p(x,y) = 5 - xy - x^2y^2 + 3y^2 + x^3y$. Its Newton polytope $N(p)$, displayed in Figure 3.1, is the convex hull of the points $(0,0),(1,1),(2,2),(0,2),(3,1)$.*

Theorem 3.4.3. *([45, Theorem 1]) If polynomial $p(x)$ is a sums of squares, then the Newton polytope of p has only even vertices corresponding to positive coefficients of p. Moreover, the Newton polytope of any square on the decomposition is contained in $New(p)$.*

Example 3.4.3. *Polynomial $M = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2$ which is the so called Motzkin polynomial, is not a sobs neither a 1-sobs.*

Proof. We prove that Motzkin Polynomial is neither a sobs nor a 1-sobs. First we will show that the Motzkin polynomial is not a sobs. We have $M(1,1,1) = 0$, therefore, assuming M is a sobs every binomial square entering in the representation of M has to vanish in $(1,1,1)$. So, if $\alpha x^{i_1}y^{i_2}z^{i_3} - \beta x^{j_1}y^{j_2}z^{j_3}$ enters in the representation, then we get $(\alpha-\beta)^2 = 0$ and so $\alpha = \beta$. Hence, every binomial square has to be of the form $(m_1 - m_2)^2$ where m_1,m_2 are monomials of degree 2 which we can assume to be distinct. For monomial $-3x^2y^2z^2$ with negative coefficient, there must exist a binomial square $\alpha^2(m_1 - m_2)^2$ so that $m_1m_2 = x^2y^2z^2$.

We have the following sum representations for exponents of monomial $-3x^2y^2z^2$, where we always assume the first monomial smaller than the second (lexicographically) and where we order the first monomial increasingly

$$\begin{aligned} 222 =& 012 + 210 \\ & 021 + 201 \\ & 102 + 120 \end{aligned}$$

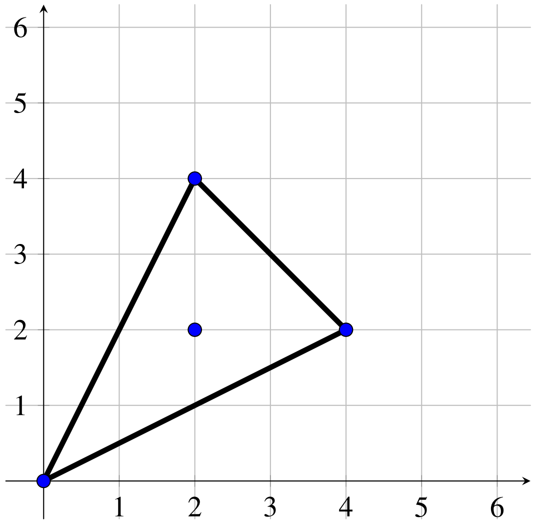

Fig. 3.2 Newton polytope of Motzkin polynomial

in any case the resultant square has Newton polytope which is not contained in the Newton polytope of M (shown in Figure 3.2). As a result $-3x^2y^2z^2$ can not be obtained by any sum of binomial squares. In order to show that Motzkin polynomial is not a 1-sobs, we do a similar argument, we first multiply $(x^2+y^2+z^2)$ with Motzkin polynomial as following

$$\begin{aligned} p(x,y,z) &= (x^2+y^2+z^2)(x^4y^2+x^2y^4+z^6-3x^2y^2z^2) \\ &= 2x^4y^4+x^6y^2+x^2y^6+y^2z^6+x^2z^6+z^8-2x^4y^2z^2-2x^2y^4z^2-3x^2y^2z^4 \end{aligned} \tag{3.1}$$

By the same reasoning as before, for every degree 8 monomial m occuring in p with negative coefficient, there must exist a binomial square $\alpha^2(m_1-m_2)^2$ so that $m_1m_2=m$. We have the following sum representations for exponents of monomials with negative coefficients

$$\begin{aligned} 224 &= 004+220 & 242 &= 022+220 & 422 &= 022+400 \\ &= 013+211 & &= 211+031 & &= 112+310 \\ &= 022+202 & &= 040+202 & &= 121+301 \\ &= 103+121 & &= 112+130 & &= 202+220 \end{aligned}$$

The representations $211+031$, $121+301$, $022+400$, $040+202$, $022+220$, $202+220$ and $022+202$ are eliminated since in any case the resultant squares are not in representation of polynomial p shown in (3.1). Figure 3.3 shows the Newton polytope of p which is the convex hull of points $\{(0,0,8),(4,2,2),(2,4,2),(2,2,4),(2,0,6),(0,2,6),(2,6,0),(6,2,0),(4,4,0)\}$. Now we are left with the following possibilities for monomials in p with negative coefficient

$$\begin{aligned} 224 &= 004+220 & 242 &= 112+130 & 422 &= 112+310 \\ &= 013+211 & & & & \\ &= 103+121 & & & & \end{aligned}$$

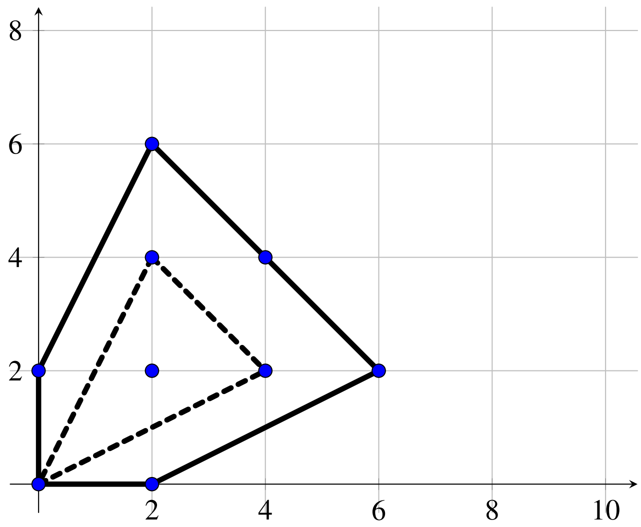

Fig. 3.3 Newton polytope of p

and with these insights, if p has a sobs representation, then we can write

$$p = p_1 + \alpha_1(z^4 - x^2y^2)^2 + \alpha_2(yz^3 - x^2yz)^2 + \alpha_3(xz^3 - xy^2z)^2 + \beta(xyz^2 - xy^3)^2 + \gamma(xyz^2 - x^3y)^2 \quad (3.2)$$

Where p_1 is a sobs which does not contain any of the shown binomials. Also p_1 by previous arguments cannot produce any of the terms with negative coefficients of p. From the representation (3.2) it follows

$$\begin{aligned} p = p_1 &+ \alpha_1 z^8 + \alpha_1 x^4y^4 + \alpha_2 x^4y^2z^2 + \alpha_3 x^2z^6 + \alpha_3 x^2z^6 + \alpha_3 x^2y^4z^2 \\ &- (2\alpha_1 + 2\alpha_2 + 2\alpha_3)x^2y^2z^4 + \beta x^2y^2z^4 + \beta x^2y^6 + \gamma x^2y^2z^4 + \gamma x^6y^2 - 2\beta x^2y^4z^2 - 2\gamma x^4y^2z^2 \end{aligned}$$

by comparison with p, we find $\beta = \gamma = 1$ and $\alpha_1 + \alpha_2 + \alpha_3 = \frac{3}{2}$ after cancellations we now have

$$\begin{aligned} &2x^4y^4 + x^6y^2 + x^2y^6 + y^2z^6 + x^2z^6 + z^8 \\ &= p_1 + \alpha_1 z^8 + \alpha_1 x^4y^4 + \alpha_2 x^4y^2z^2 + \alpha_3 x^2z^6 + \alpha_3 x^2z^6 \\ &+ \alpha_3 x^2y^4z^2 + x^2y^2z^4 + x^2y^6 + x^2y^2z^4 + x^6y^2 \end{aligned}$$

or

$$\begin{aligned} &(2-\alpha_1)x^4y^4 + (1-\alpha_1)y^2z^6 + (1-\alpha_3)x^2z^6 + (1-\alpha_1)z^8 \\ &- \alpha_2 x^4y^2z^2 - \alpha_3 x^2y^4z^2 - 2x^2y^2z^4 = p_1 \end{aligned}$$

Now we know $-2x^2y^2z^4$ cannot be generated by p_1 since it does not have the right binomials for this, as we observed already and this is a contradiction which proves Motzkin polynomial is not a 1-sobs. ■

The following proposition determines when we can say that a general multivariate homogeneous polynomial which vanishes in one, can be written as sum of binomial squares.

Proposition 3.4.4. *If $p = p(x_1,\dots,x_n)$ is a form for which $p(1,1,\dots,1) = 0$, then the question whether p is a sobs can be reduced to the question whether a certain system of linear equations has a nonnegative solution.*

Proof. We know already why p can be written as a nonnegative linear combination of expression of the form $(m-m')^2$ where m amd m' are monomials of degree d, where $2d$ is degree of p. If the monomials are $\{m_i\}_{i=1}^{L}$, we have in case (p is a sobs) a representation $p = \sum_{1\le i<j\le L} \alpha_{ij}(m_i - m_j)^2$ with $\alpha_{ij} \ge 0$. It is now clear that a linear system of equations will emerge from coefficient comparisons. ■

In the following example we show that Motzkin polynomial is 2-sobs.

Example 3.4.4. *Consider the polynomial*

$$\begin{aligned} p(x) &= (x^2+y^2+z^2)^2(x^4y^2+x^2y^4+z^6-3x^2y^2z^2) \\ &= x^2y^8+3x^6y^4+3x^4y^6-2x^4y^4z^2-x^2y^6z^2-5x^2y^4z^4+x^8y^2 \\ &\quad -x^6y^2z^2-5x^4y^2z^4+y^4z^6+x^4z^6-x^2y^2z^6+2x^2z^8+2y^2z^8+z^{10} \end{aligned}$$

By choosing monomial vector $z(x)$ as

$$z(x) = [x^4y, x^3y^2, x^2y^3, xy^4, x^2yz^2, xy^2z^2, x^2z^3, y^2z^3, xz^4, yz^4, z^5]$$

an sdd matrix M can be computed as

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & -0.5 & 0 & 0 & 0 & 0 & -0.5 & 0 \\ 0 & 3 & 0 & 0 & 0 & -0.5 & 0 & 0 & -2.5 & 0 & 0 \\ 0 & 0 & 3 & 0 & -0.5 & 0 & 0 & 0 & 0 & -2.5 & 0 \\ 0 & 0 & 0 & 1 & 0 & -0.5 & 0 & 0 & -0.5 & 0 & 0 \\ -0.5 & 0 & -0.5 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.5 & 0 & -0.5 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -0.5 & 0 & 0 & -0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.5 & 1 & 0 & 0 & -0.5 \\ 0 & -2.5 & 0 & -0.5 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ -0.5 & 0 & -2.5 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.5 & -0.5 & 0 & 0 & 1 \end{bmatrix}$$

and we have

$$\begin{aligned} p(x) &= z(x)Mz(x)^T \\ &= \frac{1}{2}(x^4y-x^2yz^2)^2+\frac{1}{2}(x^4y-yz^4)^2+\frac{1}{2}(x^3y^2-xy^2z^2)^2+\frac{5}{2}(x^3y^2-xz^4)^2 \\ &+\frac{1}{2}(x^2y^3-x^2yz^2)^2+\frac{5}{2}(x^2y^3-yz^4)^2+\frac{1}{2}(xy^4-xy^2z^2)^2+\frac{1}{2}(xy^4-xz^4)^2 \\ &+\frac{1}{2}(x^2z^3-y^2z^3)^2+\frac{1}{2}(x^2z^3-z^5)^2+\frac{1}{2}(y^2z^3-z^5)^2. \end{aligned}$$

3.5 Factor width 3 matrices and sum of trinomial squares

The purpose of this section is to show that if a quarternary quadratic form $q(w,x,y,z)$ is not a sum of squares of trinomials then, given any positive integer r, the form $(w^2+x^2+y^2+z^2)^r \cdot q$ is not a sum of squares of trinomials. In fact it will be necessary to show more generally that for nonzero reals $\lambda_1,\lambda_2,\lambda_3,\lambda_4$, the form $(\lambda_1^2 w^2+\lambda_2^2 x^2+\lambda_3^2 y^2+\lambda_4^2 z^2)^r \cdot q$ is not a sum of squares of trinomials.

We will focus first on the case $r=1$ and then indicate later the modifications necessary for the case of larger r. We start with the following proposition.

Proposition 3.5.1. *Let $\underline{x}=[w,x,y,z]$ and let $q=\underline{x}^T Q\underline{x}$ be a psd quadratic form and B a matrix such that B spans an extreme ray in $(FW_3^4)^*$ and $\langle Q,B\rangle<0$. Then the degree $(2r+2)$ form $p=(\lambda_1^2 w^2+\lambda_2^2 x^2+\lambda_3^2 y^2+\lambda_4^2 z^2)^r q$ is not a sum of trinomial squares, for any, not all zero, reals $\lambda_1,\lambda_2,\lambda_3,\lambda_4$.*

Proof. The inequality $\langle Q,B\rangle<0$ implies that B is not psd. In addition, it spans an extreme ray, hence by Proposition 3.2.6, for some permutation P and non singular matrix D and some $a,c\in]-\pi,\pi[\setminus\{0\}$ it has the following form

$$B_2 = DPBP^T D^T = \begin{bmatrix} 1 & \cos(a) & \cos(a-c) & \cos(c) \\ \cos(a) & 1 & \cos(c) & \cos(a-c) \\ \cos(a-c) & \cos(c) & 1 & \cos(a) \\ \cos(c) & \cos(a-c) & \cos(a) & 1 \end{bmatrix}.$$

We now have the inequality $0>\langle Q,B\rangle=\langle P^T D^T QDP,B_2\rangle$. We work with the quadratic form q_n defined by $q_n=x^T P^T D^T QDPx$ and show that given any $\lambda\in(\mathbb{R}^*)^4\setminus\{0\}$, we have that the associated quartic form $p_n=(\lambda_1^2 w^2+\lambda_2^2 x^2+\lambda_3^2 y^2+\lambda_4^2 z^2)^r q_n$, for any $\lambda_1,\lambda_2,\lambda_3,\lambda_4$ is not a sum of trinomial squares. Since the property of 'not being a sum of trinomial squares for any λ' is invariant under permutations and scalings of the variables in q_n, we shall get the claim concerning the original p,q. For simplicity of notation be aware that *we redefine* $(Q,B):=(P^T D^T QDP,B_2)$ and $(p,q):=(p_n,q_n)$. The original Q,B,p,q will not play any further role in this proof.

The polynomial p is of degree $2r+2$. From Theorem 3.3.1 we know that p has a -usually nonunique- representation $p=z(x)_{r+1}^T Q' z(x)_{r+1}$, where $z(x)_{r+1}$ collects all monomials of degree $r+1$ and hence Q' is an $\binom{r+4}{3}\times\binom{r+4}{3}$ matrix. We define the matrix $B'=(b'_{ij})$ as follows (where we use for the moment as the most natural indexation, the one given by the vectors of exponents of the monomials), where $i,j\in\mathbb{Z}^4_{\geq 0}$ are uples with $|i|=|j|=r+1$ so that B' is also an $\binom{r+4}{3}\times\binom{r+4}{3}$ matrix:

$$b'_{ij} = \begin{cases} b_{kl} & \text{iff } i+j \text{ has two odd entries exactly in positions } k\neq l \\ 1 & \text{iff } i+j \text{ has only even entries} \\ 0 & \text{iff } i+j \text{ has 1 or 3 odd entries} \\ \omega & \text{iff } i+j \text{ has only odd entries} \end{cases}$$

(The case that $i+j$ has exactly 1 or 3 odd entries can actually not happen in case $|i|=|j|$, but we will need the given rules below also in cases where $|i|\neq|j|$.) We will show that $B'\in(FW_3^{\binom{r+4}{3}})^*$, and then that $\langle B',Q'\rangle<0$, thus showing $Q'\notin FW_3^{\binom{r+4}{3}}$, and hence showing by Propositions 3.1.3 and 3.3.1 that p is not a sum of squares of trinomials. We will then see from the fact that 'being a sum of squares of trinomials is invariant under permutations' that the original p is also not a sum of squares of trinomials.

We split the proof that $B' \in (FW_3^{\binom{r+4}{3}})^*$, into two parts, treating the cases $r=1$ and $r \geq 2$ separately.

Case $r=1$. Then we have to show $B' \in (FW_3^{10})^*$

In the current case the vector to represent p via Q' above is given by, say,

$$z(x)_2 = [w^2, wx, wy, wz, x^2, xy, xz, y^2, yz, z^2]$$

and the matrix B' is the inner part of the table given below

$$B' = \begin{array}{c|cccccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
 & w^2 & wx & wy & wz & x^2 & xy & xz & y^2 & yz & z^2 \\
\hline
1 & 1 & b_{12} & b_{13} & b_{14} & 1 & b_{23} & b_{24} & 1 & b_{34} & 1 \\
2 & b_{12} & 1 & b_{23} & b_{24} & b_{12} & b_{13} & b_{14} & b_{12} & \omega & b_{12} \\
3 & b_{13} & b_{23} & 1 & b_{34} & b_{13} & b_{12} & \omega & b_{13} & b_{14} & b_{13} \\
4 & b_{14} & b_{24} & b_{34} & 1 & b_{14} & \omega & b_{12} & b_{14} & b_{13} & b_{14} \\
5 & 1 & b_{12} & b_{13} & b_{14} & 1 & b_{23} & b_{24} & 1 & b_{34} & 1 \\
6 & b_{23} & b_{13} & b_{12} & \omega & b_{23} & 1 & b_{34} & b_{23} & b_{24} & b_{23} \\
7 & b_{24} & b_{14} & \omega & b_{12} & b_{24} & b_{34} & 1 & b_{24} & b_{23} & b_{24} \\
8 & 1 & b_{12} & b_{13} & b_{14} & 1 & b_{23} & b_{24} & 1 & b_{34} & 1 \\
9 & b_{34} & \omega & b_{14} & b_{13} & b_{34} & b_{24} & b_{23} & b_{34} & 1 & b_{34} \\
10 & 1 & b_{12} & b_{13} & b_{14} & 1 & b_{23} & b_{24} & 1 & b_{34} & 1
\end{array}$$

The bordering is for the convenience of the reader. For example, the entry in row 3, column 8 of matrix B' below corresponds to the pair of monomials (wy, y^2) and thus to the pair of 4-uples of exponents $(1010, 0020)$. The sum of these 4-uples is (1030) which has exactly two odd entries at positions 1 and 3 and therefore the entry of B' is b_{13}. One notices that B' is a symmetric real 10×10 matrix in which rows and hence also columns $1, 5, 8, 10$ are equal. Thus to show that all 3×3 principal submatrices are positive semidefinite, it is sufficient to examine those of the 7×7 submatrix occurring after striking out the rows and columns $5, 8, 10$ of B'. Doing so one gets the reduced matrix B'_{red}.

$$B'_{\text{red}} = \begin{array}{c|ccccccc}
 & 1 & 2 & 3 & 4 & 6 & 7 & 9 \\
 & w^2 & wx & wy & wz & xy & xz & yz \\
\hline
1 & 1 & b_{12} & b_{13} & b_{14} & b_{23} & b_{24} & b_{34} \\
2 & b_{12} & 1 & b_{23} & b_{24} & b_{13} & b_{14} & \omega \\
3 & b_{13} & b_{23} & 1 & b_{34} & b_{12} & \omega & b_{14} \\
4 & b_{14} & b_{24} & b_{34} & 1 & \omega & b_{12} & b_{13} \\
6 & b_{23} & b_{13} & b_{12} & \omega & 1 & b_{34} & b_{24} \\
7 & b_{24} & b_{14} & \omega & b_{12} & b_{34} & 1 & b_{23} \\
9 & b_{34} & \omega & b_{14} & b_{13} & b_{24} & b_{23} & 1
\end{array}$$

To see that all principal 3×3 submatrices of B'_{red} are positive semidefinite note first that the left upper 4×4 matrix of B'_{red} coincides with B and more generally all principal 3×3 submatrices of B'_{red}

which do not contain an ω are permutation equivalent to 3×3 principal submatrices of B and hence are automatically positive semidefinite. The 3×3 principal submatrices containing ω stem from selecting sets of three line indices which contain one of the sets $\{2,9\},\{3,7\},\{4,6\}$. These matrices are permutation equivalent to one of the following matrices:

$$\begin{bmatrix} 1 & \omega & b_{12} \\ \omega & 1 & b_{34} \\ b_{12} & b_{34} & 1 \end{bmatrix}, \begin{bmatrix} 1 & \omega & b_{14} \\ \omega & 1 & b_{23} \\ b_{14} & b_{23} & 1 \end{bmatrix}, \begin{bmatrix} 1 & \omega & b_{13} \\ \omega & 1 & b_{24} \\ b_{13} & b_{24} & 1 \end{bmatrix}.$$

So it is sufficient to find an $\omega \in \mathbb{R}$ such that these matrices are positive semidefinite. To see this, the easiest choices possible are $\omega = \pm 1$; (These are universal choices valid for all $0 < a,b,c < \pi$ that result in determinants equal to 0. If one has given explicit real numbers for a,b,c, then putting $\omega = -1+\varepsilon$ or $\omega = 1-\varepsilon$ for sufficiently small $\varepsilon > 0$, one will obtain strictly positive definite (sub)determinants.) With these checks the case is done.

Case $r \geq 2$. To show that B' is a matrix in $(FW_3^{\binom{r+4}{3}})^*$, we consider first the inner part of the following 16×16 matrix which is a direct sum $\tilde{B} = \hat{B}\oplus\hat{B}$ where $\hat{B}$ is the left upper 8×8 matrix and where blanks are zeros.

$$\tilde{B} = \begin{array}{|cccccccccccccccc}
1 & wx & wy & wz & xy & xz & yz & wxyz & w & x & y & z & wxy & wxz & wyz & xyz \\
\hline
1 & b_{12} & b_{13} & b_{14} & b_{23} & b_{24} & b_{34} & \omega & & & & & & & & \\
b_{12} & 1 & b_{23} & b_{24} & b_{13} & b_{14} & \omega & b_{34} & & & & & & & & \\
b_{13} & b_{23} & 1 & b_{34} & b_{12} & \omega & b_{14} & b_{24} & & & & & & & & \\
b_{14} & b_{24} & b_{34} & 1 & \omega & b_{12} & b_{13} & b_{23} & & & & & & & & \\
b_{23} & b_{13} & b_{12} & \omega & 1 & b_{34} & b_{24} & b_{14} & & & & & & & & \\
b_{24} & b_{14} & \omega & b_{12} & b_{34} & 1 & b_{23} & b_{13} & & & & & & & & \\
b_{34} & \omega & b_{14} & b_{13} & b_{24} & b_{23} & 1 & b_{12} & & & & & & & & \\
\omega & b_{34} & b_{24} & b_{23} & b_{14} & b_{13} & b_{12} & 1 & & & & & & & & \\
 & & & & & & & & 1 & b_{12} & b_{13} & b_{14} & b_{23} & b_{24} & b_{34} & \omega \\
 & & & & & & & & b_{12} & 1 & b_{23} & b_{24} & b_{13} & b_{14} & \omega & b_{34} \\
 & & & & & & & & b_{13} & b_{23} & 1 & b_{34} & b_{12} & \omega & b_{14} & b_{24} \\
 & & & & & & & & b_{14} & b_{24} & b_{34} & 1 & \omega & b_{12} & b_{13} & b_{23} \\
 & & & & & & & & b_{23} & b_{13} & b_{12} & \omega & 1 & b_{34} & b_{24} & b_{14} \\
 & & & & & & & & b_{24} & b_{14} & \omega & b_{12} & b_{34} & 1 & b_{23} & b_{13} \\
 & & & & & & & & b_{34} & \omega & b_{14} & b_{13} & b_{24} & b_{23} & 1 & b_{12} \\
 & & & & & & & & \omega & b_{34} & b_{24} & b_{23} & b_{14} & b_{13} & b_{12} & 1
\end{array}$$

For the convenience of the reader the top line contains a version of the column indexation of the matrix; the same is used for the rows. The entries of the matrix are obtained as given by this example: take the row indexed by wxy and the column indexed by wy. The product is $w^2x^1y^2z^0$. The 4-uple of exponents, 2120 has precisely one odd entry. Hence by the rules for forming a matrix B' - we shall see why we use these rules also here - we have to put 0 in the entry of address (wxy, wy) or $(1110, 1010)$. Similarly the entry of address (xy, xz) yields the product $x^2yz = 0211$ which has exactly two odd entries namely at positions $3,4$ therefore by the above rules we have to put the entry with this address equal to b_{34}.

Again we observe as before that the matrix $\tilde{B}$ has the matrix B as a 4×4 submatrix in the left upper corner and due to this many of the principal 3×3 submatrices of $\tilde{B}$ are positive semidefinite. Furthermore it is clear that a 3×3 submatrix which 'intersects' both summands $\hat{B}$ is positive semidefinite. The positive semidefiniteness of the other 3×3 submatrices of $\tilde{B}$ was also already established in the first part of the proof.

We now show that B' is a submatrix of a certain Kronecker product with one factor equal to $\hat{B}$. To any string of exponents $i=(i_1,i_2,i_3,i_4)\in\mathbb{Z}^4_{\geq 0}$ we can associate a unique 4-uple $\varepsilon=\varepsilon(i)=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)\in\{0,1\}^4$ defined by $i_\nu\equiv\varepsilon_\nu \mod 2$.

Consider the $2l\times 2l$ matrix L obtained by stacking its first two rows l times one over the other.

$$L=\begin{pmatrix} 1 & 0 & 1 & 0 & \dots & 0 & 1 & 0\\ 0 & 1 & 0 & 1 & \dots & 1 & 0 & 1\\ & & & & \vdots & & & \\ 0 & 1 & 0 & 1 & \dots & 1 & 0 & 1\end{pmatrix}$$

Since by the definition of B', the value of b'_{ij} depends only on $\varepsilon(i)$ and $\varepsilon(j)$ it is evident that B' and hence any submatrix of B' is a submatrix of $L\otimes\hat{B}$ if l is sufficiently large. So it is sufficient to show that every 3×3 submatrix of $L\otimes\hat{B}$ is positive semidefinite.

To see this, for ease of explanation, assume the lines of $\hat{B}$ indexed as $0,1,2,\dots,7$ and those of $L\otimes\hat{B}$ as indexed by $0^0,1^0,\dots,7^0,0^1,1^1,\dots,7^1,\dots\dots,0^{2l-1},1^{2l-1},\dots,7^{2l-1}$. For line indices $i^{i'},j^{j'}$ of $L\otimes\hat{B}$ one sees that

$$(L\otimes\hat{B})_{i^{i'}j^{j'}}=\begin{cases}\hat{b}_{ij} & \text{if either } i',j'\in 0,2,\dots,2l-2 \text{ or } i',j'\in 1,3,\dots,2l-1;\\ 0 & \text{in all other cases.}\end{cases}$$

Let now $i^{i'},j^{j'},k^{k'}$ be three distinct line indices. The principal 3×3 submatrix of $L\otimes\hat{B}$ selected by these line indices is in the case that i,j,k are distinct equal to the principal 3×3 submatrix of $\hat{B}$ defined by line indices i,j,k. This is the following matrix.

$$\begin{pmatrix}\hat{b}_{ii} & \hat{b}_{ij} & \hat{b}_{ik}\\ \hat{b}_{ij} & \hat{b}_{jj} & \hat{b}_{jk}\\ \hat{b}_{ik} & \hat{b}_{jk} & \hat{b}_{kk}\end{pmatrix}$$

This matrix is by hypothesis positive semidefinite.

But if i,j,k are not distinct, say that $i=j\neq k$, then the line indices are $i^{i'},i^{j'},k^{k'}$. Then the submatrix obtained is the following matrix.

$$\begin{pmatrix}\hat{b}_{ii} & \hat{b}_{ii} & \hat{b}_{ik}\\ \hat{b}_{ii} & \hat{b}_{ii} & \hat{b}_{ik}\\ \hat{b}_{ik} & \hat{b}_{ik} & \hat{b}_{kk}\end{pmatrix}$$

It is easy to see that this matrix inherits positive semidefiniteness from its left neighbour. The positive semidefiniteness in the cases $i = k, j = k$ and $i = j = k$ can be similarly inferred. In fact the whole purpose for the construction of $L \otimes \hat{B}$ was to put the matrix $\tilde{B}$ to good use in case $\varepsilon(i), \varepsilon(j), \varepsilon(k)$ in B' are not all distinct as might happen in case of large r.

This concludes the case $r \geq 2$. We now show the other claim we made for B'.

Claim: There holds $\langle B', Q' \rangle = (\sum_{i=1}^4 \lambda_i^2)^r \langle B, Q \rangle$; and consequently $\langle Q', B' \rangle < 0$.

By the definition of the inner product in matrix space, we have to show

$$\sum \{ b'_{ij} q'_{ij} : \, i, j \in \mathbb{Z}^4_{\geq 0}, |i| = |j| = 1 + r \} = (\sum_{i=1}^4 \lambda_i^2)^r \sum_{i,j=1}^4 b_{ij} q_{ij}.$$

Now, given $i, j \in \mathbb{Z}^4_{\geq 0}, |i| = |j| = 1 + r$, we have of course $|i + j| = 2r + 2$. Since for an $s \in \mathbb{Z}^4_{\geq 0}$ for which $|s|$ is even it is impossible that s has exactly one or three odd entries, we can write the left side above as follows:

$$\sum_{\substack{|s| = 2r+2 \\ s \text{ has four} \\ \text{even entries}}} \sum_{\substack{|i| = |j| = r+1 \\ i + j = s}} b'_{ij} q'_{ij} + \sum_{\substack{|s| = 2r+2 \\ s \text{ has two} \\ \text{odd entries}}} \sum_{\substack{|i| = |j| = r+1 \\ i + j = s}} b'_{ij} q'_{ij} + \sum_{\substack{|s| = 2r+2 \\ s \text{ has four} \\ \text{odd entries}}} \sum_{\substack{|i| = |j| = r+1 \\ i + j = s}} b'_{ij} q'_{ij}.$$

By the definition of B' given, this is equal to

$$\sum_{\substack{|s| = 2r+2 \\ s \text{ has four} \\ \text{even entries}}} \sum_{\substack{|i| = |j| = r+1 \\ i + j = s}} q'_{ij} + \sum_{1 \leq k < l \leq 4} \sum_{\substack{|s| = 2r+2 \\ s \text{ has odd} \\ \text{entries at } k,l}} \sum_{\substack{|i| = |j| = r+1 \\ i + j = s}} b_{kl} q'_{ij} + \sum_{\substack{|s| = 2r+2 \\ s \text{ has four} \\ \text{odd entries}}} \sum_{\substack{|i| = |j| = r+1 \\ i + j = s}} \omega q'_{ij}.$$

Now we remember that by its construction, polynomial p cannot have a monomial with only odd exponents so the third sum is 0. The sum of the coefficients of monomials whose variables have only even powers in p is given by Proposition 3.3.2 by

$$(\sum_{i=1}^4 \lambda_i^2)^r (q_{11} + q_{22} + q_{33} + q_{44});$$

while the second sum is

$$\sum_{1 \leq k < l \leq 4} b_{kl} \sum_{\substack{|s| = 2r+2 \\ s \text{ has odd} \\ \text{entries at } k,l}} \sum_{\substack{|i| = |j| = r+1 \\ i + j = s}} q'_{ij}$$

The inner double sum here can be described exactly as the sum of the coefficients of the monomials of p which have two odd entries at distinct k, l. Hence again by Proposition 3.3.2 the whole inner sum is equal to $2(\sum \lambda_i^2)^r q_{kl}$ and so the sum is

$$2(\sum_{i=1}^4 \lambda_i^2)^r \sum_{1 \leq k < l \leq 4} b_{kl} q_{kl} = (\sum_{i=1}^4 \lambda_i^2)^r \sum_{\substack{1 \leq k, l \leq 4, \\ k \neq l}} b_{kl} q_{kl}.$$

The claim now follows because $\sum_{i=1}^4 \lambda_i^2 > 0$.

To conclude the proof we detail an idea we mentioned at the beginning. We have until now shown that whatever the reals $\lambda_1 \lambda_2, \lambda_3, \lambda_4$, not all zeros are, for a polynomial $q_n = x^T P'^T Q P' x$, with Q satisfying the hypotheses, the polynomial $p_n = (\lambda_1^2 w^2 + \lambda_2^2 x^2 + \lambda_3^2 y^2 + \lambda_4^2 z^2)^r q_n$ is not a sum of trinomial squares. Now by its definition $q_n(w,x,y,z) = q(\pi(w), \pi(x), \pi(y), \pi(z))$ where π embodies the permutation matrix P'. Since the property 'to be a sum of squares of trinomials' is evidently invariant under permutations, it follows that $(\lambda_1^2 \pi^{-1}(w)^2 + \lambda_2^2 \pi^{-1}(x)^2 + \lambda_3^2 \pi^{-1}(y)^2 + \lambda_4^2 \pi^{-1}(z)^2)^r q(w,x,y,z)$ for whatever $\lambda_1, ..., \lambda_4$, is not sum of trinomial squares. Since $\{\pi^{-1}(w), \pi^{-1}(x), \pi^{-1}(y), \pi^{-1}(z)\} = \{w,x,y,z\}$ it follows that $(\lambda_1^2 w^2 + \lambda_2^2 x^2 + \lambda_3^2 y^2 + \lambda_4^2 z^2)^r q(w,x,y,z)$ is not sum of trinomial squares. ■

Theorem 3.5.2. *Assume $\lambda_1, ..., \lambda_4$ are reals, not all zero. If the quadratic form $q(\underline{x}) = q(w,x,y,z)$ is not a sum of squares of trinomials, the quarternary form $(w^2 + x^2 + y^2 + z^2)^r q(\underline{x})$ is not a sum of squares of trinomials.*

Proof. If the quadratic form is not positive semidefinite then the claim is trivial. So assume now q is positive semidefinite and let it be written as $q = x^T Q x$. Then Q is positive semidefinite and by Proposition 3.3.1, $Q \notin FW_3^4$. So there exists $B \in (FW_3^4)^*$ spanning in $(FW_3^4)^*$ an extreme ray such that $\langle B, Q \rangle < 0$. By Proposition 3.5.1 it follows that $\lambda_1^2 w^2 + \lambda_2^2 x^2 + \lambda_3^2 y^2 + \lambda_4^2 z^2)^r q(\underline{x})$ is not a sum of squares of trinomials for any $\lambda_1 \lambda_2, \lambda_3, \lambda_4$. In particular, $(w^2 + x^2 + y^2 + z^2) q(x)$ is not a sum of squares of trinomials. ■

Chapter 4

On completely positive programming and its approximations

In this chapter, we first review the copositive and completely positive cones mentioned in Chapter 2 and then introduce copositive and completely positive programming. Motivated by the expressive power of copositive and completely positive programming to encode hard optimization problems, we will present some formulations of combinatorial and quadratic optimization problems and due to the NP-hardness of these problems, we will review some approximations for them. Mainly our focus will be on completely positive programming and its approximations. Hence, we will review both outer approximations (approximations by larger cones) and some existing LP and SDP inner approximations. Then, we propose the use of the cone of nonnegative scaled diagonally dominant matrices as a natural inner approximation to the completely positive cone. Using projections of this cone we derive new graph-based second-order cone approximation schemes for completely positive programming, leading to both uniform and problem-dependent hierarchies. This offers a compromise between the expressive power of semidefinite programming and the speed of linear programming based approaches. We also present numerical results on random problems and the stable set problem to illustrate the effectiveness of our approach.

4.1 Copositive and completely positive cones

Recall from Chapter 2 that we defined the cone of copositive matrices as

$$\mathcal{COP}^n := \{X \in \mathcal{S}^n \mid v^T X v \geq 0, \text{ for all } v \geq 0\}, \tag{4.1}$$

and its dual, the cone of completely positive matrices as

$$\mathcal{CP}^n := \{X \in \mathcal{S}^n \mid \exists B \geq 0, \ X = B^T B\}. \tag{4.2}$$

Example 4.1.1. *As an example the matrix*

$$X = \begin{bmatrix} 6 & 2 & 4 \\ 2 & 4 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

is completely positive since it can be written as $X = B^T B$ *with* $B = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix}$

Recall from Chapter 2 that $\mathcal{N}^n$ is the cone of nonnegative $n \times n$ symmetric matrices, and is self-dual. By definition of $\mathcal{N}^n$ and $\mathcal{S}^n_+$ (the set of positive semidefinite matrices), it is obvious that $\mathcal{CP}^n \subset \mathcal{S}^n_+ \cap \mathcal{N}^n$ and that $\mathcal{S}^n_+ + \mathcal{N}^n \subset \mathcal{COP}^n$. We refer to $\mathcal{S}^n_+ \cap \mathcal{N}^n$ as the doubly nonnegative cone (DNN^n) and we have $(DNN^n)^* = \mathcal{S}^n_+ + \mathcal{N}^n$. In fact, for $n \times n$ matrices of order $n \leq 4$, we have equality in the two inclusions above.

Example 4.1.2. *As an example the matrix*

$$X = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 3 \\ -1 & 1 & 1 \end{bmatrix} \in \mathcal{COP}^3$$

is copositive since

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 3 \\ -1 & 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix}}_{\in \mathcal{N}^3} + \underbrace{\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}}_{\in \mathcal{S}^3_+}.$$

However, for $n \geq 5$ both inclusions are strict. An example that illustrates $DNN^n \neq \mathcal{CP}^n$ for $n \geq 5$ is the following matrix

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 3 \end{bmatrix} \in DNN^5 \setminus \mathcal{CP}^5.$$

We refer the readers for a good introduction to copositive and completely positive cones to [21], [22]. Another important cone which we will use later to give an inner approximation for the cone of completely positive matrices is the cone of nonnegative scaled diagonally dominant matrices defined as

$$SDD^n \cap \mathcal{N}^n = \{A \in \mathcal{S}^n \mid A \text{ is scaled diagonally dominant}, (A)_{ij} >= 0.\} \tag{4.3}$$

This cone lives naturally inside of the completely positive cone, thus we have the following chain of inclusions among different cones for any positive integer n,

$$SDD^n \cap \mathcal{N}^n \subset \mathcal{CP}^n \subset DNN^n \subset \mathcal{S}^n_+ \subset (DNN^n)^* \subset \mathcal{COP}^n \subset SDD^n + \mathcal{N}^n.$$

In order to illustrate theses inclusions, in Figure 4.1 we took a random 2 dimensional slice of different cones using a 5×5 matrix and plotted them in one figure. The outermost set colored in violet is $SDD^* + \mathcal{N}$, the green area inside is $PSD + N$ and the areas colored in yellow, red and blue correspond to PSD, $PSD \cap \mathcal{N}$ and $SDD \cap \mathcal{N}$ respectively. The cone $\mathcal{CP}$ is sandwiched between $SDD \cap \mathcal{N}$ and $PSD \cap \mathcal{N}$ and also the cone $\mathcal{COP}$ is sandwiched between $PSD + \mathcal{N}$ and $SDD + \mathcal{N}$.

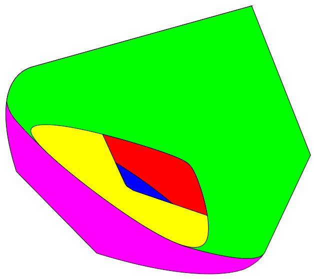

Fig. 4.1 Two dimensional sections of different cones

4.2 Copositive and completely positive programming

In Chapter 2, we briefly reviewed several types of conic programming problems, namely linear programming, semidefinite programming, second order cone programming and copositive and completely positive programming. Among this list of conic programs, in this section we will look into completely positive programming. Recall that this problem has the following form

$$\begin{aligned} v_p := \min \quad & \operatorname{tr}(CX) \\ \text{s.t.} \quad & \operatorname{tr}(A_i X) = b_i, \; i = 1, \ldots, m, \\ & X \in \mathcal{CP}^n, \end{aligned} \tag{4.4}$$

where C and A_i, $i = 1, \ldots, m$ are symmetric matrices.

We also consider the dual problem of (4.4), which is the following copositive programming problem

$$\begin{aligned} v_d := \max \quad & b^T y \\ \text{s.t.} \quad & C - \textstyle\sum_{i=1}^m y_i A_i \in \mathcal{COP}^n. \end{aligned} \tag{4.5}$$

As for any conic programming, strong duality holds if Slater condition [13] holds and in that case the optimal values of (4.4) and (4.5) are equal.

Completely positive programming and its dual counterpart of copositive programming are classes of convex optimization problems that have in the past decades developed as a particularly expressive tool to encode optimization problems, especially for many problems arising from combinatorial or quadratic optimization. As a first example consider the standard quadratic problem of optimizing a quadratic objective over the standard simplex.

$$\begin{aligned} \min \quad & x^T Q x \\ \text{s.t.} \quad & e^T x = 1, \\ & x \geq 0 \end{aligned} \tag{4.6}$$

where e denotes the all-ones vector. Easy manipulations show that the objective function can be written as $x^T Qx = \langle Q, xx^T \rangle$. Analogously the constraint $e^T x = 1$ transforms to $\langle E, xx^T \rangle = 1$, with $E = ee^T$. By considering $X = xx^T$, we can rewrite problem (4.6) as the following problem

$$\begin{aligned} \min \quad & \langle Q, X \rangle \\ \text{s.t.} \quad & \langle E, X \rangle = 1, \\ & X \in \mathcal{CP}^n \end{aligned} \tag{4.7}$$

which is obviously a relaxation of the problem (4.6). Now, since the objective is linear, an optimal solution must be attained in an extremal point of the convex feasible set. It has been shown in [20, Theorem 4.2] that these extremal points are exactly the rank-one matrices xx^T with $x \geq 0$ and $e^T x = 1$. Together, these results imply that (4.7) is in fact an exact reformulation of (4.6).

As a second example, this time combinatorial, consider the Stable Set Problem (SSP) of finding in a graph G the largest set of vertices such that no two are connected by an edge. The cardinality of such set is known as stability number of G, denoted by $\alpha(G)$. It was shown in [18] that this can be solved by the completely positive program

$$\begin{aligned} \alpha(G) = \max \quad & \langle E, X \rangle \\ \text{s.t.} \quad & \langle A_G + I, X \rangle = 1, \\ & X \in \mathcal{CP}^n \end{aligned} \tag{4.8}$$

where A_G is the adjacency matrix of G and E is the matrix of ones.

Another classical example can be found in [16], which shows that general quadratic programs with a mix of binary and continuous variables can be expressed as copositive programs. A large body of work has been developed in the area and there is a series of survey papers that can be consulted for further information. We refer the readers to [12, 17, 22] and references therein for more details.

In general, copositive and completely positive optimization problems are NP-hard problems and thus we would not expect to be able to solve them exactly in an efficient way. Instead we consider approximating the copositive and completely positive cones with other cones over which we can optimize efficiently. In what follows we first define the concept of inner and outer approximations for a closed convex set and then in the next sections we review some inner and outer approximations for the completely positive cone.

4.3 Inner and outer approximations to a convex set

In general, for a sequence of convex cones, $\{\mathcal{I}_r | r \in \mathbb{Z}_+\}$, we say that $\mathcal{I}_r$ is an inner approximation hierarchy for a closed convex set K if

$$\mathcal{I}_0 \subseteq \mathcal{I}_1 \subseteq \ldots \subseteq \bigcup_{r\in\mathbb{Z}_+} \mathcal{I}_r \subseteq K$$

and we say that an inner approximation hierarchy converges if

$$cl(\bigcup_{r\in\mathbb{Z}_+} \mathcal{I}_r) = K.$$

In other words, for all $X \in int(K)$, there exists an $r \in \mathbb{Z}_+$ such that $X \in \mathcal{I}_r$. Similarly, for a sequence of convex cones $\{\mathcal{O}_r | r \in \mathbb{Z}_+\}$, we say that $\mathcal{O}_r$ is an outer approximation hierarchy for a closed convex set K if

$$\mathcal{O}_0 \supseteq \mathcal{O}_1 \supseteq \ldots \supseteq \bigcap_{r\in\mathbb{Z}_+} \mathcal{O}_r \supseteq K$$

and we say that an outer approximation hierarchy converges if

$$\bigcap_{r\in\mathbb{Z}_+} \mathcal{O}_r = K.$$

In other words, for all $X \notin K$, there exists an $r \in \mathbb{Z}_+$ such that $X \notin \mathcal{O}_r$.

Approximation hierarchies play an important role in conic programming. For example, consider the following conic optimization problem

$$\begin{aligned} X^* = \min \quad & \langle C, X\rangle \\ \text{s.t.} \quad & \langle A_i, X\rangle = b_i, \quad i = 1,\ldots,m \\ & X \in K. \end{aligned} \tag{4.9}$$

Using inner and outer approximation hierarchies $\mathcal{I}_r$ and $\mathcal{O}_r$, we can approximate problem (4.9) and obtain lower and upper bounds for the optimal solutions of this problem.

$$\begin{aligned} X_{\mathcal{I}_r} = \min \quad & \langle C, X\rangle \\ \text{s.t.} \quad & \langle A_i, X\rangle = b_i, \quad i = 1,\ldots,m \\ & X \in \mathcal{I}_r, \end{aligned} \tag{4.10}$$

$$\begin{aligned} X_{\mathcal{O}_r} = \min \quad & \langle C, X\rangle \\ \text{s.t.} \quad & \langle A_i, X\rangle = b_i, \quad i = 1,\ldots,m \\ & X \in \mathcal{O}_r, \end{aligned} \tag{4.11}$$

then for all $r \in \mathbb{Z}_+$, we have

$$X_{\mathcal{I}_r} \geq X_{\mathcal{I}_r+1} \geq X^* \geq X_{\mathcal{O}_r+1} \geq X_{\mathcal{O}_r}.$$

Moreover, if problem (4.9) is strictly feasible and the hierarchies are convergent, then it can be seen that $\lim_{r\to\infty} X_{\mathcal{I}_r} = X^*$. Similarly, if the dual problem to problem (4.9) is strictly feasible, then it can be seen that $\lim_{r\to\infty} X_{\mathcal{O}_r} = X^*$.

4.4 Outer approximations to the completely positive cone

Several approximation schemes have been proposed and successfully used in the literature, based on outer approximations to $\mathcal{CP}^n$. The simplest one is to replace $\mathcal{CP}^n$ by the cone of nonnegative positive semidefinite matrices (doubly nonnegative) which is strictly larger than $\mathcal{CP}^n$ when $n \geq 5$, hence leading to a lower bound to v_p. This approximation is exact for $n < 5$, but gets very weak as n grows.

Another outer approximation is the one proposed by Parrilo in [41]. In fact, this approximation is an inner approximation to the copositive cone, but it can be considered as an outer approximation to the completely positive cone by defining corresponding dual cones. We explain this approach here briefly.

Given a matrix $A \in \mathcal{S}^n$, consider the polynomial $P_A(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i^2 x_j^2$. Clearly, $A \in \mathcal{COP}$ if and only if $P_A(x) \geq 0$ for all $x \in \mathbb{R}^n$ [22]. Then according to what we saw in Chapter 2, a sufficient condition for a polynomial to be nonnegative is that it can be written as sums of squares, and it is easy to see that $P_A(x)$ has a sum of squares decomposition if and only if $A \in (\mathcal{S}_+^n + \mathcal{N}^n)$, which yields the relation $\mathcal{S}_+^n + \mathcal{N}^n \subseteq \mathcal{COP}$. This idea can be extended using the following theorem from Reznick [46].

Theorem 4.4.1. *If $f(x_1,\ldots,x_n)$ is a homogeneous polynomial which is positive on $\mathbb{R}_+^n \setminus \{0\}$, then for sufficiently large $r \in \mathbb{N}$, the polynomial*

$$f(x_1,\ldots,x_n)\left(\sum_{i=1}^n x_i^2\right)^r$$

has positive coefficients.

The following hierarchy of cones can be defined to approximate the $\mathcal{COP}$ cone from the interior.

$$\mathcal{K}^r = \left\{ A \in \mathcal{S}_+^n \mid P_A(x)\left(\sum_{i=1}^n x_i^2\right)^r \text{ has an sos decomposition} \right\}$$

It can be shown that $\mathcal{S}_+^n + \mathcal{N}^n = \mathcal{K}_0 \subset \mathcal{K}_1 \subset \ldots$, and $int(\mathcal{COP}) \subseteq \bigcup_{r\in\mathbb{N}} \mathcal{K}^r$, so when r increases, the cones $\mathcal{K}^r$ converges to the $\mathcal{COP}$. Similarly, a family of dual cones $(\mathcal{K}^r)^*$ can be defined to approximate the $\mathcal{CP}$ cone from the exterior. Since the sos condition can be written as a system of linear matrix inequalities (LMIs), optimizing over $\mathcal{K}^r$ amounts to solving a semidefinite program. Figure 4.2 shows sections of the inner approximations to $\mathcal{COP}$ cone. The innermost set is a section of $\mathcal{S}_+^5 + \mathcal{N}^5$ which is equal to the initial cone of Parrilo hierarchy and the red area behind shows the approximation given by hierarchy when $r = 1$. As r gets bigger and bigger, the approximation gets finer.

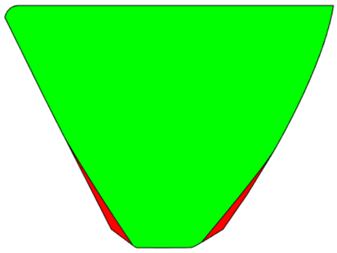

Fig. 4.2 Comparison of Parrilo approximation with $\mathcal{S}_+^5 + \mathcal{N}^5$

Also, Peña et al [42] used another certificate for nonegativity of the same polynomial to define other outer approximation for the $\mathcal{CP}$ cone. In their approach, optimizing over each of the cones is again a semidefinite program. In the Parrilo and Peña approaches the system of LMIs becomes very large quickly when r increases. Hence, dimension of the semidefinite programs increases very fast and current SDP-solvers can only solve problems over those cones for small values of r, i.e., $r \leq 3$ at most. Moreover, De Klerk and Pasechnik [18], and Bomze and De Klerk [11] used different sufficient conditions for nonnegativity of the same polynomial to define other outer approximations for the $\mathcal{CP}$ cone. In their approaches, each of the approximating cones is polyhedral, so optimizing over one of them is solving an LP. All these approximation hierarchies approximate $\mathcal{COP}$ (resp. $\mathcal{CP}$) uniformly and thus do not take into account any information provided by the objective function of an optimization problem.

4.5 Inner approximations to the completely positive cone

For upper bounds based on inner approximations to $\mathcal{CP}^n$, the literature is somewhat sparser. One way of constructing inner approximations to $\mathcal{CP}^n$ is to make use of the fact that the extreme rays of $\mathcal{CP}^n$ are matrices of the form vv^T with $v \in \mathbb{R}_+^n \backslash \{0\}$; see [20, Theorem 4.2]. Thus, one can pick uniformly spaced $v \in \Delta_n = \left\{x \in \mathbb{R}_+^n : \sum x_i = 1\right\}$, and approximate $\mathcal{CP}^n$ by the cone the matrices vv^T generate (see [15, 53]). This leads to linear programming (LP) approximations to (4.4). Another inner approximation to $\mathcal{CP}^n$ is that proposed in [34], based on the theory of moments, leading to semidefinite programming (SDP) approximations to (4.4). In both cases we have hierarchies that give upper bounds to (4.4), and dually lower bounds to (4.5), and converge to the optimal value/solutions of (4.4). These inner approximations are uniform (i.e., problem-independent) approximations, giving rise to either LP or SDP problems. See also [54] for a more thorough treatment of inner approximations. An extra step taken as an adaptive linear approximation algorithm was proposed in [15]. This uses information obtained from an upper bound approximation to selectively refine the hierarchy, leading to problem-dependent LP approximations.

Here we will take a look to two approximations given by [15] and [34].

4.5.1 Bundfuss and Dür approximation

One inner approximation to $\mathcal{CP}$ cone is the one proposed by Bundfuss and Dür in [15]. This approximation relies on the observation that A is copositive if and only if the quadratic form $x^T Ax \geq 0$

on the standard simplex. In fact, letting $v_1,\dots,v_n$ to denote the vertices of a simplex, a point x in the simplex can be written in barycentric coordinates as $x=\sum_{i=1}^n \lambda_i v_i$ with $\lambda_i \geq 0$ and $\sum_{i=1}^n \lambda_i = 1$. Using this fact we have

$$x^T Ax = \sum_{i,j=1}^{n} v_i^T A v_j \lambda_i \lambda_j$$

Thus, a sufficient condition for $x^T Ax$ to be nonnegative on the simplex is that

$$v_i^T A v_j \geq 0,\ \forall i,j.$$

A simplicial partition $\mathcal{P}$ of the standard simplex Δ is a family of smaller simplices $\Delta^1,\dots,\Delta^m$ such that $\Delta = \bigcup_{i=1}^m \Delta^i$ and $int(\Delta^i)\cap int(\Delta^j)=\emptyset$ for $i\neq j$. In [15], the authors use such partition to define an inner approximation.

Let $V_{\mathcal{P}}$ denote the set of all vertices of simplices in $\mathcal{P}$ and $E_{\mathcal{P}}$ the set of all edges of simplices in $\mathcal{P}$. Define the diameter of a partition $\mathcal{P}$ to be

$$\delta(\mathcal{P}) = \max_{\{u,v\}\in E_{\mathcal{P}}} \|u-v\|$$

Define also the following set corresponding to a given partition $\mathcal{P}$

$$\mathcal{I}_{\mathcal{P}} = \{A \in \mathcal{S}^n \mid v^T A v \geq 0,\ \text{for all } v\in V_{\mathcal{P}}, u^T A v \geq 0,\ \text{for all } (u,v)\in E_{\mathcal{P}}\}$$

It is not difficult to see that $\mathcal{I}_{\mathcal{P}}$ is a closed, convex, polyhedral cone which approximates $\mathcal{COP}$ from the interior. Similarly define the sets

$$\mathcal{O}_{\mathcal{P}} = \{A \in \mathcal{S}^n \mid v^T A v \geq 0,\ \text{for all } v\in V_{\mathcal{P}}\}$$

and this set is also a closed, convex, polyhedral cone which approximates $\mathcal{COP}$ from the exterior. In fact, it is not also hard to see that for any partition $\mathcal{P}$ of Δ,

$$\mathcal{I}_{\mathcal{P}}^* = \{ \sum_{\{u,v\}\in E_{\mathcal{P}}} \lambda_{uv}(uv^T + vu^T) + \sum_{v\in V_{\mathcal{P}}} \lambda_v vv^T : \lambda_{uv}, \lambda_v \in \mathbb{R}_+\}$$

is an outer approximation of $\mathcal{CP}$, and

$$\mathcal{O}_{\mathcal{P}}^* = \{ \sum_{v\in V_{\mathcal{P}}} \lambda_v vv^T : \lambda_v \in \mathbb{R}_+\}$$

is an inner approximation of $\mathcal{CP}$. Both inner and outer approximations of $\mathcal{COP}$ (resp. $\mathcal{CP}$) converge to $\mathcal{COP}$ (resp. $\mathcal{CP}$) if the diameter of the partitions goes to zero. Note that since inner and outer approximations are polyhedra, optimizing over them is solving an LP problem.

They then create a refining strategy for the partitions, starting at an initial partition and subdividing simplices where the constraints are active. This partitioning strategy is guided adaptively by the objective function, which yields to a good approximation of the completely positive cone in those parts that are relevant for the optimization and only a coarse approximation in those parts that are not and more complete description about the adaptive approach can be found in [15]. Occasionally we can use this approach trivially in a non-adaptive way by uniformly dividing the simplex in equal simplices reduced by a factor of k, where k is a positive integer. We call that a k-regular partition. As we increase k we get convergence as well.

4.5.2 Lasserre approximation

Another inner approximation to $\mathcal{CP}$ cone is the one proposed by Lasserre in [34] that we briefly explain it here.

Recall that a Borel measure on $\mathbb{R}^n$ is a nonnegative set function on Borel sets of $\mathbb{R}^n$, such that $\mu(\emptyset) = 0$ and $\mu(\bigcup_{i=1}^p E_i) = \sum_{i=1}^p \mu(E_i)$ for any countable collection of disjoint Borel sets $E_1, \dots, E_p \subseteq \mathbb{R}^n$. A Borel measure μ is a probability measure if and only if $\mu(\mathbb{R}^n) = 1$. Define the support of a Borel measure μ as the minimal closed set $U \subseteq \mathbb{R}^n$ such that $\mu(\mathbb{R}^n \setminus U) = 0$ and denote it by $supp(\mu)$. For any $\alpha \in \mathbb{N}^n$, define the α moment of a Borel measure μ as

$$y_\alpha := \int_{\mathbb{R}^n} x^\alpha d\mu(x).$$

Recall that P_n is the ring of real polynomials in n variables $x = (x_1, \dots, x_n)$ and $P_{n,d}$ is the vector space of real polynomials in n variables and of degree less than or equal to d. Also let $\mathbb{N}_d^n = \{\alpha \in \mathbb{N}^n \mid \sum_i \alpha_i \leq d\}$. Then the moment matrix is defined as following.

Definition 8. *With a sequence $y = (y_\alpha)$, $\alpha \in \mathbb{N}_n$, let $L_y : P_n \to R$ be the linear functional*

$$h = \sum_\alpha h_\alpha x^\alpha \mapsto L_y(h) = \sum_\alpha h_\alpha y_\alpha, \quad h \in P_n.$$

With $d \in \mathbb{N}$, let $M_d(y)$ be the symmetric matrix with rows and columns indexed in $\mathbb{N}_d^n$, and defined by

$$M_d(y)(\alpha, \beta) := L_y(x^{\alpha+\beta}) = y_{\alpha+\beta}, \quad (\alpha, \beta) \in \mathbb{N}_d^n \times \mathbb{N}_d^n.$$

The matrix $M_d(y)$ is called the moment matrix associated with y.

It can be easily checked that

$$L_y(g^2) \geq 0 \quad \text{for all } g \in P_n \Leftrightarrow M_d(y) \succeq 0, \quad d = 0, 1, \dots$$

Let us now consider a polynomial $f(x) = \sum_{\gamma \in \mathbb{N}^n} f_\gamma x^\gamma$ and let $M_d(fy)$ be the symmetric matrix with rows and columns indexed in $\mathbb{N}_d^n$. We define the localizing matrix associated with f and y as

$$M_d(fy)(\alpha, \beta) := L_y(f(x) x^{\alpha+\beta}) = \sum_\gamma f_\gamma y_{\alpha+\beta+\gamma}, \quad (\alpha, \beta) \in \mathbb{N}_d^n \times \mathbb{N}_d^n. \tag{4.12}$$

It can be seen that

$$\langle g, M_d(fy) g \rangle = L_y(g^2 f), \quad \text{for all } g \in P_{n,d}.$$

A well-known theorem when considering moments in connection to polynomials is the following.

Theorem 4.5.1. *[33, Theorem 3.2] Consider a finite Borel measure μ and a polynomial f, such that the support of μ is equal to $K \subseteq R^n$. The polynomial $f(x)$ is nonnegative if and only if*

$$\int_K g^2 f d\mu \geq 0, \quad \textit{for all } g \in P_n. \tag{4.13}$$

If $y = (y_\alpha)$, $\alpha \in \mathbb{N}^n$ is the sequence of moments of μ then (4.13) *is in turn equivalent to*

$$M_d(fy) \succeq 0 \quad \textit{for all } d = 0, 1, \ldots$$

where $M_d(fy)$ is the localizing matrix associated with f and y, defined in (4.12).

In [35], the authors particularize this result to the case of copositive matrices viewed as homogeneous forms of degree 2, nonnegative on the closed set $K = \mathbb{R}^n_+$.

So with $A = (a_{ij}) \in \mathcal{S}^n$, let $f_A = x^T A x$ be a quadratic form and let μ be the joint probability measure associated with n i.i.d. exponential variables (with mean 1), with $supp(\mu) = \mathbb{R}^n_+$, and with moments $y = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, given by

$$y_\alpha = \int_{\mathbb{R}^n_+} x^\alpha d\mu(x) = \int_{\mathbb{R}^n_+} x^\alpha exp(-\sum_{i=1}^n x_i) dx = \prod_{i=1}^n \alpha_i!, \quad \text{for all } \alpha \in \mathbb{N}^n. \tag{4.14}$$

Recall that a matrix $A \in \mathcal{S}^n$ is copositive if $f_A(x) \geq 0$ for all $x \in \mathbb{R}^n_+$ and denote by $\mathcal{COP} \subset \mathcal{S}^n$ the cone of copositive matrices, i.e.,

$$\mathcal{COP} = \{A \in \mathcal{S}^n \mid f_A(x) \geq 0, \quad \text{for all } x \in \mathbb{R}^n_+\}.$$

Next, introduce the following sets $C_d \subset \mathcal{S}^n$, $d = 0, 1, \ldots$ defined by

$$C_d = \{A \in \mathcal{S}^n \mid M_d(f_A y) \succeq 0\}, \quad d = 0, 1, \ldots \tag{4.15}$$

where $M_d(f_A y)$ is the localizing matrix defined in (4.12), associated with the quadratic form f_A and the sequence y in (4.14).

It can be easily seen that the entries of the matrix $M_d(f_A y)$ are homogeneous and linear in A. Therefore, the condition $M_d(f_A y) \succeq 0$ is a homogeneous linear matrix inequality and hence defines a spectrahedron of $\mathcal{S}^n$. Each $C_d \subset \mathcal{S}^n$ is a convex cone defined solely in terms of the entries of $A \in \mathcal{S}^n$, and the hierarchy of spectrahedra C_d, with $d \in \mathbb{N}$, provides a nested sequence of outer approximations for $\mathcal{COP}$.

Theorem 4.5.2. *[35, Theorem 2.1] Let y be as in* (4.14) *and let $C_d \subset \mathcal{S}^n$, $d = 0, 1, \ldots$ be the hierarchy of convex cones defined in* (4.15). *Then $C_0 \supset C_1 \supset \ldots \supset C_d \ldots \supset \mathcal{COP}$ and $\mathcal{COP} = \bigcap_{d=0}^{\infty} C_d$.*

In Figure 4.3 we compared two outer approximations for $\mathcal{COP}$ cone given by Bundfuss-Dür and Lasserre approaches for a 2 dimensional slice of the 5×5 symmetric matrix. In Lasserre approach, we considered the moment vectors to be in the form (4.14). We did the test for $d = 2, 3$ and plotted two dimensional slices of each which in the Figure 4.3 corresponds to two outermost bluish areas. The innermost sets are those approximated by Bundfuss-Dür approach using k-regular partitions with $k = 2, 6, 10$ and the green area inside is $\mathcal{S}^5_+ + \mathcal{N}^5$. It is clear that the approximation given by

Bundfuss-Dür approach is better since it is closer to $\mathcal{S}_+^5 + \mathcal{N}^5$ compared to the approximation given by Lasserre. Also in Lasserre approach the underlying SDP problem grows very fast as d increases and this approximation becomes inefficient.

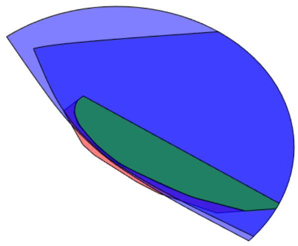

Fig. 4.3 Comparison of dual Lasserre ($d = 2,3$) with dual linear approach of Bundfuss-Dür with partition numbers $K = 2,6,10$ and $\mathcal{S}_+^5 + \mathcal{N}^5$

4.6 Inner approximating the completely positive cone via the cone of scaled diagonally dominant matrices

In this section, we propose a new inner approximation scheme to $\mathcal{CP}^n$ that is based on second-order cone programming (SOCP) problems and can be either uniform or problem-dependent. Our approach is motivated by the recent work in [2, 3] that uses the cone of scaled diagonally dominant matrices for inner-approximating the cone of positive semidefinite matrices which we reviewed in Chapter 2. Specifically, we use the cones of nonnegative scaled diagonally dominant matrices (SDD_+) and their projections as a natural inner approximation to $\mathcal{CP}^n$, and derive a new SOCP-based approximation scheme for completely positive and copositive programming. Our approximation scheme has a natural graphical interpretation. By exploiting this interpretation, we can flexibly expand or trim the SOCP problems in our hierarchy, leading to both uniform and problem-dependent approximation schemes. The use of SOCP offers a compromise between the expressive power of SDP, that comes at a significant computational cost, and the speed of LP approaches, that have inherently lower expressive power. Numerical experiments on solving random instances and the stable set problem demonstrate the effectiveness of our approximation schemes.

4.6.1 Blanket assumptions

We make the following **blanket assumptions** concerning (4.4) and (4.5):

A1. Problem (4.4) is feasible.

A2. The mapping $X \mapsto (\text{tr}(A_1X), \dots, \text{tr}(A_mX))$ is surjective.

A3. Problem (4.5) is strictly feasible, i.e., there exists $\bar{y}$ satisfying

$$C - \sum_{i=1}^{m} \bar{y}_i A_i \in \text{int}\,\mathcal{COP}^n.$$

Under these assumptions, the dual Slater condition holds [13]. Therefore we have $v_p = v_d$, with both values being finite and the primal optimal value v_p being attained.

4.7 The scaled diagonally dominant cone and beyond

In this section, we present the basis for our construction of inner approximations in Sections 4.8 and 4.9. Our construction is motivated by the work in [2, 3], which studied inner approximations of the cone of positive semidefinite matrices based on the cones of diagonally dominant and scaled diagonally dominant matrices which we reviewed in Chapter 2. While their work can be directly applied to the existing SOS hierarchies to yield outer approximations of $\mathcal{CP}^n$ (see [3, Section 4.2]) we show an alternative approach, based on the same cones but using them in a fundamentally different way, in order to obtain an inner approximation to $\mathcal{CP}^n$.

Recall from Chapter 2 that, the cone SDD^n of $n\times n$ sdd matrices is given by

$$\mathrm{SDD}^n := \sum_{1\le i<j\le n} \iota_{ij}(\mathcal{S}^2_+), \tag{4.16}$$

where $\iota_{ij} : \mathcal{S}^2 \to \mathcal{S}^n$ is the map that sends an $S \in \mathcal{S}^2$ to the matrix D given by

$$d_{rs} := \begin{cases} s_{11} & \text{if } (r,s) = (i,i), \\ s_{12} & \text{if } (r,s) = (i,j), \\ s_{21} & \text{if } (r,s) = (j,i), \\ s_{22} & \text{if } (r,s) = (j,j), \\ 0 & \text{otherwise.} \end{cases}$$

This cone is therefore given in terms of 2×2 semidefinite constraints or, in other words, second-order cone constraints, which makes it quite suitable to use in convex optimization.

One can prove the following basic properties of SDD^n, and of the set $\mathrm{SDD}^n_+ := \mathrm{SDD}^n \cap \mathcal{N}^n$. Note that item (i) in Proposition 4.7.1 below can be found in [2], and a more general version of it can be found in [43, Lemma 5]. We include it here for completeness. In what follows ι^*_{ij} denotes the adjoint of the map ι_{ij}, which in this case can be defined by saying that $\iota^*_{ij}(S)$ is the 2×2 submatrix of S indexed by rows and columns i and j.

Proposition 4.7.1. *The following statements hold.*

(i) $(\mathrm{SDD}^n)^* = \{Q \in \mathcal{S}^n : \ \iota^*_{ij}(Q) \succeq 0, \ \ \forall 1 \le i < j \le n\}$.

(ii) $(\mathrm{SDD}^n_+)^* = (\mathrm{SDD}^n)^* + \mathcal{N}^n$.

(iii) $\mathrm{SDD}^n_+ = \sum_{1\le i<j\le n} \iota_{ij}(\mathcal{S}^2_+ \cap \mathcal{N}^2)$.

Proof. We first prove (i). Recall from (4.16) that $\mathrm{SDD}^n = \sum_{1\le i<j\le n} \iota_{ij}(\mathcal{S}^2_+)$. Thus, we have from [47, Corollary 16.3.2] that

$$(\mathrm{SDD}^n)^* = \bigcap_{1\le i<j\le n} (\iota_{ij}(\mathcal{S}^2_+))^*,$$

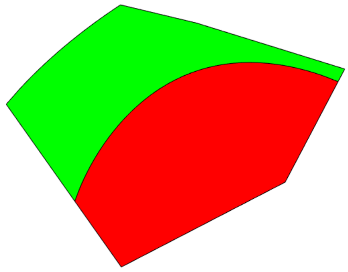

Fig. 4.4 Comparison of $\mathcal{S}_+^5 \cap \mathcal{N}^5$ with SDD_+^5

from which the desired equality follows immediately.

Next, we prove (ii). Note that

$$\sum_{1\le i<j\le n} \iota_{ij}(E) \in \mathrm{SDD}^n \cap \operatorname{int}\mathcal{N}^n.$$

Thus, we conclude from [47, Corollary 16.3.2] that

$$(\mathrm{SDD}_+^n)^* = (\mathrm{SDD}^n \cap \mathcal{N}^n)^* = (\mathrm{SDD}^n)^* + \mathcal{N}^n.$$

Finally, we prove (iii). It is clear that $\mathrm{SDD}_+^n \supseteq \sum_{1\le i<j\le n} \iota_{ij}(\mathcal{S}_+^2 \cap \mathcal{N}^2)$. For the converse inclusion, consider any $Q \in \mathrm{SDD}_+^n$. Then Q is nonnegative and can be written as $\sum_{1\le i<j\le n} \iota_{ij}(S_{ij})$ for some $S_{ij} \in \mathcal{S}_+^2$, $1 \le i < j \le n$. Observe that each S_{ij} has nonnegative diagonal entries, and moreover, its nondiagonal entry equals the (i,j)th entry of Q, which is also nonnegative. Thus, $S_{ij} \in \mathcal{S}_+^2 \cap \mathcal{N}^2$ and hence $Q \in \sum_{1\le i<j\le n} \iota_{ij}(\mathcal{S}_+^2 \cap \mathcal{N}^2)$. This completes the proof. ■

Since 2×2 nonnegative positive semidefinite matrices are completely positive, we see from Proposition 4.7.1(iii) that SDD_+^n is an inner approximation to $\mathcal{CP}^n$. In Figure 4.4 we show a random 2 dimensional slice of the cone of doubly nonnegative 5×5 matrices (i.e., $\mathcal{S}_+^5 \cap \mathcal{N}^5$) with the slice of SDD_+^5 highlighted in red. The cone $\mathcal{CP}^5$ is sandwiched between them. This simple inner approximation can be used as a basis to construct more general inner second-order cone approximations for $\mathcal{CP}^n$. To do that we consider a useful variant of SDD_+^n that will help us construct inner approximations of $\mathcal{CP}^n$.

Definition 9. *Let $U \in \mathbb{R}_+^{t\times n}$ have row sum* 1. *Define*

$$\mathrm{SDD}_+^n(U) := \{U^TYU : Y \in \mathrm{SDD}_+^t\} = U^T(\mathrm{SDD}_+^t)U. \tag{4.17}$$

The above definition is similar to the development in [2, Section 3.1], which makes use of the so-called $\mathrm{DD}(U)$. Here we assume that U has nonnegative entries so that $\mathrm{SDD}_+^n(U)$ will be a subcone of $\mathcal{CP}^n$; see Proposition 4.7.3 below. In addition, we assume that the rows of U have sum one: we can then always think of the rows of U as points in the simplex Δ^n. This is no less general than just considering $U \in \mathbb{R}_+^{t\times n}$ with nonzero rows, because scaling rows of U by positive scalars does not

change $\mathrm{SDD}_+^n(U)$. Note that $\mathrm{SDD}_+^n(U)$ is simply a linear image of SDD_+^t into $\mathcal{S}^n$.
Some basic properties of this set are that $\mathrm{SDD}_+^n(I_n) = \mathrm{SDD}_+^n$, and that if $U \in \mathbb{R}_+^{t\times n}$ is a submatrix of $\tilde{U} \in \mathbb{R}_+^{s\times n}$ then $\mathrm{SDD}_+^n(U) \subseteq \mathrm{SDD}_+^n(\tilde{U})$. We show in the next example that $\mathrm{SDD}_+^n(U)$ can be strictly larger than SDD_+^n in general.

Example 4.7.1. *One can see that the matrix*

$$M = \begin{bmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{bmatrix}$$

is in $\mathcal{S}_+^3 \cap \mathcal{N}^3$. However, $M \notin \mathrm{SDD}_+^3$; indeed, if we define

$$W := \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix},$$

then $\mathrm{tr}(WM) < 0$ but $W \in (\mathrm{SDD}_+^3)^*$ thanks to Proposition 4.7.1(ii), showing that $M \notin \mathrm{SDD}_+^3$.

Now, suppose we set U to be the 4×3 matrix constructed from concatenating the identity I_3 with an all $\frac{1}{3}$ row vector, i.e.,

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix},$$

and consider the set $\mathrm{SDD}_+^3(U)$. Then we know $\mathrm{SDD}_+^3 \subseteq \mathrm{SDD}_+^3(U)$ because I_3 is a submatrix of U. Furthermore, we have

$$M = U^T \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 3 \\ 3 & 3 & 3 & 27 \end{bmatrix} U \in \mathrm{SDD}_+^3(U),$$

where the inclusion holds because

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 3 \\ 3 & 3 & 3 & 27 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 9 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 9 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 9 \end{bmatrix} \in \mathrm{SDD}_+^4.$$

Consequently, $\mathrm{SDD}_+^3(U)$ is a strictly larger set than SDD_+^3.

We next give an important characterization of $\mathrm{SDD}_+^n(U)$ that is crucial in our development of inner approximation schemes in Sections 4.8 and 4.9. Recall from (4.2) that $\mathcal{CP}^n$ can be seen as the convex hull of all vv^T with $v \in \mathbb{R}_+^n$. The next theorem shows that one can think of $\mathrm{SDD}_+^n(U)$ similarly.

Theorem 4.7.2. *Let $U \in \mathbb{R}_+^{t\times n}$ have row sum 1. Then $\mathrm{SDD}_+^n(U)$ is the conic hull of all vv^T with v belonging to some line segment $[u_i, u_j]$, where u_i is the i-th row of U.*

Proof. Note from Proposition 4.7.1(iii) and (4.17) that any matrix in $\mathrm{SDD}_+^n(U)$ can be written as

$$\sum_{1\le i<j\le n} U^T \iota_{ij}(S_{ij})U$$

for some $S_{ij} \in \mathcal{S}_+^2 \cap \mathcal{N}^2$. Moreover, any matrix $S \in \mathcal{S}_+^2 \cap \mathcal{N}^2$ can be written as $S = v_1v_1^T + v_2v_2^T$ for some nonnegative vectors $v_i \in \mathbb{R}_+^2$. Furthermore, we know that for any $v \in \mathbb{R}^2$, it holds that $\iota_{ij}(vv^T) = ww^T$ where $w \in \mathbb{R}^t$ is the vector whose ith entry is v_1, jth entry equals v_2, and is zero otherwise. Hence, we deduce that any matrix in $\mathrm{SDD}_+^n(U)$ can be written as

$$\sum_{k=1}^{N} U^T w_k w_k^T U,$$

where each $w_k \in \mathbb{R}^t$ is nonnegative and has a support of cardinality at most 2. Conversely, it is easy to see that any matrix that can be written as such a sum is in $\mathrm{SDD}_+^n(U)$. But each $U^T w$, with $w \neq 0$, is simply a (nonzero) conic combination of two rows of U, u_i and u_j; so, up to positive scaling, it is in $[u_i,u_j]$, proving our claim. ∎

We can now prove the following properties of $\mathrm{SDD}_+^n(U)$.

Proposition 4.7.3. *Let* $U \in \mathbb{R}_+^{t\times n}$ *have row sum* 1. *Then the following statements hold.*

(i) *The cone* $\mathrm{SDD}_+^n(U)$ *is a closed sub-cone of* $\mathcal{CP}^n$.

(ii) $(\mathrm{SDD}_+^n(U))^* = \{Y : UYU^T \in (\mathrm{SDD}_+^t)^*\} = \{Y : UYU^T \in (\mathrm{SDD}^t)^* + \mathcal{N}^t\}$.

Proof. From Theorem 4.7.2, it follows that $\mathrm{SDD}_+^n(U)$ is a sub-cone of $\mathcal{CP}^n$. It remains to prove closedness. Since U is nonnegative and has no zero rows, the origin is not in the convex hull of vv^T, where v belongs to some $[u_i,u_j]$, and u_i is the i-th row of U. Hence $\mathrm{SDD}_+^n(U)$ is the conic hull of a compact convex set not containing the origin. Thus, it is closed.

To prove (ii), recall that $\mathrm{SDD}_+^n(U) = U^T(\mathrm{SDD}_+^t)U$. From this we see that $Y \in (\mathrm{SDD}_+^n(U))^*$ if and only if

$$\mathrm{tr}(Y(U^TWU)) \ge 0 \quad \forall W \in \mathrm{SDD}_+^t,$$

which is the same as $UYU^T \in (\mathrm{SDD}_+^t)^*$. This proves the first equality. The second equality in (ii) follows from Proposition 4.7.1(ii). This completes the proof. ∎

Note that the construction of $\mathrm{SDD}_+^n(U)$ is fairly general. Anytime we have a cone $\mathcal{C} \subseteq \mathcal{CP}^t$ and a matrix $U \in \mathbb{R}_+^{t\times n}$ whose rows have sum one, with $t \ge n$, one can define the cone

$$\mathcal{C}(U) := \{U^TYU : Y \in \mathcal{C}\} = U^T\mathcal{C}U. \tag{4.18}$$

This is easily seen to always verify $\mathcal{C}(U) \subseteq \mathcal{CP}^n$, since $\mathcal{C}(U) \subseteq U^T\mathcal{CP}^tU \subseteq \mathcal{CP}^n$. It is helpful to state in this language the usual LP inner approximations to $\mathcal{CP}^n$. Let Diag_+^n be the set of nonnegative $n\times n$ diagonal matrices. Clearly $\mathrm{Diag}_+^n \subseteq \mathcal{CP}^n$, so we can define

$$\mathrm{Diag}_+^n(U) := \{U^TYU : Y \in \mathrm{Diag}_+^t\}. \tag{4.19}$$

This is nothing more than the conic hull of the matrices $u_i u_i^T$, $i = 1,\dots,t$, where u_i is the i-th row of U. The use of (4.19) for inner approximation corresponds to the standard LP approximation strategy used, for example, in [15], where strategies for efficient choices of U were explored.

Another possibility for obtaining an LP relaxation would be to use the cone of $n \times n$ symmetric nonnegative diagonally dominant matrices, denoted by DD_+^n. We have $\mathrm{Diag}_+^n \subseteq \mathrm{DD}_+^n \subseteq \mathrm{SDD}_+^n$. So, if we define

$$\mathrm{DD}_+^n(U) := \{U^T Y U : \; Y \in \mathrm{DD}_+^t\}, \tag{4.20}$$

we would get $\mathrm{Diag}_+^n(U) \subseteq \mathrm{DD}_+^n(U) \subseteq \mathrm{SDD}_+^n(U)$. However, since one can easily see that DD_+^n is the conic hull of $(e_i+e_j)(e_i+e_j)^T$ for $1 \le i \le j \le n$, it is not hard to see that $\mathrm{DD}_+^n(U)$ is simply the conic hull of $(u_i+u_j)(u_i+u_j)^T$ for $1 \le i \le j \le t$, and hence can be expressed in terms of $\mathrm{Diag}_+^n(U')$ for some U' that contains U as a submatrix.

Other choices would be to use not submatrices in $\mathcal{S}_+^2$, as we did for SDD_+^n, but matrices in $\mathcal{S}_+^3$ or $\mathcal{S}_+^4$. Note that it is still true in these two cases that $\mathcal{S}_+^i \cap \mathcal{N}^i \subseteq \mathcal{CP}^i$. These cones would give better approximations, but we would get a much higher number of constraints that would not be second-order cone constraints but fully semidefinite. While the semidefinite constraints would still be small, the process would become more cumbersome and significantly less tractable.

4.7.1 A graphical refinement

We saw above that $\mathrm{SDD}_+^n(U)$ is a natural inner approximation to $\mathcal{CP}^n$. Furthermore, Theorem 4.7.2 suggests that the fundamental property of U that guides the approximation is the collection of segments $[u_i, u_j]$. We might associate to the points u_i vertices of a graph, and to the segments its edges, and think of the collection of points and segments as a concrete realization of the graph in $\mathbb{R}^n$. This insight can be used to refine the approximation, making it more flexible. We start by generalizing the notion of SDD.

Given a graph G with vertex set $\{1,\dots,n\}$ and edge set $\mathcal{E}$, we define

$$\mathrm{SDD}^G := \sum_{\{i,j\}\in\mathcal{E}} \iota_{ij}(\mathcal{S}_+^2),$$

and we set $\mathrm{SDD}^G := \{0\}$ if $\mathcal{E} = \emptyset$ by convention. The graph G simply encodes which principal 2×2 submatrices will be required to be semidefinite. In particular, if we consider G to be the complete graph K^n, this is simply SDD^n. We can define SDD_+^G as the nonnegative matrices in SDD^G, similarly as before. Then we can naturally define a generalization of $\mathrm{SDD}_+^n(U)$:

Definition 10. *For a graph G with t vertices and a matrix $U \in \mathbb{R}_+^{t\times n}$ whose rows have sum one, we define the cone $\mathrm{SDD}_+^G(U)$ as*

$$\mathrm{SDD}_+^G(U) := \{U^T Y U : \; Y \in \mathrm{SDD}_+^G\} = U^T(\mathrm{SDD}_+^G)U.$$

It will be helpful to think of the rows of U as points in the standard simplex Δ^n (i.e. with nonnegative coordinates summing to one). These points correspond to vertices of the graph G, and the edge set of G simply encodes which pairs of rows of U (vertices) are “connected". In other words, the pair (G,U) is a realization of the graph G inside Δ^n with segments for edges. We will denote by

$\mathrm{seg}(G,U)$ the set of points in some of the segments, i.e,

$$\mathrm{seg}(G,U) = \bigcup_{\{i,j\}\in\mathcal{E}} [u_i, u_j],$$

where u_i is the i-th row of U. This set completely controls the geometry of the cone. Based on this notion and the proof of Theorem 4.7.2, we can immediately obtain the following refinement of Theorem 4.7.2 for the representation of $\mathrm{SDD}_+^G(U)$.

Theorem 4.7.4. *Let G be a graph with t vertices and $U \in \mathbb{R}_+^{t\times n}$ be a matrix whose rows have sum one. Then $\mathrm{SDD}_+^G(U)$ is the conic hull of all vv^T with $v \in \mathrm{seg}(G,U)$.*

Theorem 4.7.4 gives a simple way of translating results from the graph language to results about cones. In particular if we have $\mathrm{seg}(G,U) \subseteq \mathrm{seg}(G',U')$, we have $\mathrm{SDD}_+^G(U) \subseteq \mathrm{SDD}_+^{G'}(U')$, and furthermore $\mathrm{SDD}_+^G(U) \subseteq \mathrm{SDD}_+^{K^t}(U) = \mathrm{SDD}_+^n(U) \subseteq \mathcal{CP}^n$, for all graphs G with t vertices and matrices $U \in \mathbb{R}_+^{t\times n}$ whose rows have sum one.On the other hand, if every node of the graph G is covered by some edges, then $\mathrm{SDD}_+^G(U) \supseteq \mathrm{Diag}_+^n(U)$, the usual LP inner approximation. Thus, the graphical notation allows us to construct intermediate approximations somewhere in between the simple LP inner approximation and the full $\mathrm{SDD}_+^n(U)$ version.

We end the section by noting that most of our other previous results concerning SDD_+^n and $\mathrm{SDD}_+^n(U)$ can be adapted with no effort to this new cone.

Theorem 4.7.5. *Given a graph G with $t \geq n$ vertices and edge set $\mathcal{E}$, and a matrix $U \in \mathbb{R}_+^{t\times n}$ whose rows have sum one, we have the following properties.*

(i) $(\mathrm{SDD}^G)^* = \{Q \in \mathcal{S}^n : \iota_{ij}^*(Q) \succeq 0 \;\; \forall \{i,j\} \in \mathcal{E}\}$;

(ii) $\mathrm{SDD}_+^G = \sum_{\{i,j\}\in\mathcal{E}} \iota_{ij}(\mathcal{S}_+^2 \cap \mathcal{N}^2)$;

(iii) $(\mathrm{SDD}_+^G(U))^* = \{Y : UYU^T \in (\mathrm{SDD}_+^G)^*\}$;

(iv) $\mathrm{SDD}_+^G(U)$ *is a closed sub-cone of* $\mathcal{CP}^n$.

Proof. Immediate from the proofs of Proposition 4.7.1 and Proposition 4.7.3. ■

4.8 Inner approximation schemes for the completely positive cone

The main idea of this section is to approximate the solution to (4.4) by using the cones $\mathrm{SDD}_+^G(U)$ to replace $\mathcal{CP}^n$. More concretely our scheme is based on the following family of optimization problems, which depends on a graph G on $t \geq n$ vertices and a $U \in \mathbb{R}_+^{t\times n}$ whose rows have sum one:

$$\begin{aligned} v_p(G,U) := \min \quad & \mathrm{tr}(CX) \\ \text{s.t.} \quad & \mathrm{tr}(A_iX) = b_i,\ i = 1,\ldots,m, \\ & X \in \mathrm{SDD}_+^G(U), \end{aligned} \tag{4.21}$$

and its dual problem given by

$$\begin{aligned} v_d(G,U) := \max \quad & b^T y \\ \text{s.t.} \quad & C - \textstyle\sum_{i=1}^m y_iA_i \in (\mathrm{SDD}_+^G(U))^*. \end{aligned} \tag{4.22}$$

Note that the semidefinite constraints in (4.21) are imposed only on 2×2 matrices. Thus, these problems are SOCP problems.

Recall from Theorem 4.7.5 that $\mathrm{SDD}_+^G(U)$ and $(\mathrm{SDD}_+^G(U))^*$ are both closed convex cones. Also, notice that (4.22) has a strictly feasible point due to Assumption **A**3 and the fact that $\mathcal{COP}^n \subseteq (\mathrm{SDD}_+^G(U))^*$ (which follows from $\mathrm{SDD}_+^G(U) \subseteq \mathcal{CP}^n$). Consequently, if Problem (4.21) is feasible, then $v_p(G,U) = v_d(G,U)$, both values are finite and $v_p(G,U)$ is attained. Moreover, we conclude from $\mathrm{SDD}_+^G(U) \subseteq \mathcal{CP}^n$ that $v_p(G,U) \geq v_p$. Furthermore, we have already pointed out that augmenting the embedded graph (G,U) leads to an enlargement in $\mathrm{SDD}_+^G(U)$. In view of these observations, we will discuss strategies for constructing an "enlarging" sequence of graphs $\{(G^k,U^k)\}$ to possibly tighten the gap $v_p(G^k,U^k) - v_p$ as k increases.

To simplify our terminology, we make the following definition.

Definition 11. *A sequence of embedded graphs $\{(G^k,U^k)\}$ is called a positively enlarging sequence if* $\mathrm{seg}(G^k,U^k) \subseteq \mathrm{seg}(G^{k+1},U^{k+1})$*, each U is a nonnegative matrix having at least n rows, each row of U (the realizations of vertices of G) sums to one, and each node of G is covered by at least one edge.*

Positively enlarging sequences verify $v_p(G^k,U^k) \geq v_p(G^{k+1},U^{k+1}) \geq v_p$ by construction. Furthermore, once (4.21) is feasible for some $k = k_0$, it will remain feasible whenever $k \geq k_0$, since the sequence of sets $\{\mathrm{SDD}_+^{G_k}(U^k)\}$ are monotonically increasing. Moreover, we have noted above that we might think of the rows of U to be in the simplex Δ^n so that we can think of this as an enlarging family of graphs embedded in Δ^n.

We next study convergence of our inner approximation schemes for (4.4) based on (4.21) when $\{(G^k,U^k)\}$ is a positively enlarging sequence. We first prove a convergence result concerning a similar approximation scheme, which uses $\mathrm{Diag}_+^n(U)$ (as defined in (4.19)) in place of $\mathrm{SDD}_+^G(U)$ in (4.21). This strategy was used in [15], which studied the pairs (4.21) and (4.22) with $\mathrm{Diag}_+^n(U)$ in place of $\mathrm{SDD}_+^G(U)$, and constructed an "enlarging" sequence $\{U^k\}$ by adding new rows to U^k from Δ^n at each step. To determine what rows to add, they solve another LP approximation scheme based on U, which they see as the set of vertices of a simplicial partition of Δ^n, and use its results to construct a sequence of $\{U^k\}$ with an increasing number of rows. In studying the convergence of that method they proved a version of the following result for copositive programming problems in [15, Theorem 4.2]. The version presented below will be useful for studying convergence of our inner approximation schemes for (4.4).

Theorem 4.8.1. *Assume that* (4.4) *is strictly feasible. Let $\{U^k\}$ be a sequence of matrices whose rows have sum one, where for each k, $U^k \in \mathbb{R}_+^{t_k\times n}$ for some $t_k \geq n$. Suppose that*

$$\lim_{k\to\infty} \max_{x\in\Delta^n} \min_{i=1,\dots,t_k} \|x - u_i^k\| = 0, \tag{4.23}$$

where u_i^k is the i-th row of U^k. Consider for each k the following problem

$$\begin{aligned} \tilde{v}_p(U^k) := \min \quad & \mathrm{tr}(CX) \\ \text{s.t.} \quad & \mathrm{tr}(A_iX) = b_i,\ i = 1,\dots,m, \\ & X \in \mathrm{Diag}_+^n(U^k). \end{aligned} \tag{4.24}$$

Then the following statements hold.

(i) $\tilde{v}_p(U^k)$ *is finite for all sufficiently large* k *and* $\lim_{k\to\infty}\tilde{v}_p(U^k) = v_p$.

(ii) *The solution set of* (4.24) *is nonempty and uniformly bounded for all sufficiently large* k.

(iii) *Let* X^k *be a solution of* (4.24) *whenever the solution set is nonempty. Then any accumulation point of* $\{X^k\}$ *is a solution of* (4.4).

Proof. Note that the $\mathrm{Diag}_+^n(U^k)$ defined in (4.19) is the conic hull of $u_i^k {u_i^k}^T$, where u_i^k are rows of U^k. Note also that any element X in $\mathcal{CP}^n$ can be written as the conic combination of $\frac{n(n+1)}{2}$ matrices vv^T, with $v \in \Delta^n$. Thus, in view of (4.23), X can then be written as the limit of a sequence $\{X^k\}$, where $X^k \in \mathrm{Diag}_+^n(U^k)$ for each k. This together with $\mathrm{Diag}_+^n(U^k) \subseteq \mathcal{CP}^n$ shows that the sequence of sets $\{\mathrm{Diag}_+^n(U^k)\}$ converges to $\mathcal{CP}^n$ in the sense of Painlevé-Kuratowski [48, Chapter 4B].

Since the mapping $X \mapsto \mathcal{A}(X) := (\mathrm{tr}(A_1X),\ldots,\mathrm{tr}(A_mX))$ is surjective by Assumption **A**2 and (4.4) is strictly feasible, the vector b and the set $\mathcal{A}(\mathcal{CP}^n)$ cannot be separated in the sense of [48, Theorem 2.39]. Thus, [48, Theorem 4.32] shows that the sequence of feasible sets of (4.24) converges to the feasible set of (4.4) in the sense of Painlevé-Kuratowski.

It now follows from [48, Theorem 4.10(a)] and the nonemptiness of the feasible set of (4.4) that the feasible sets of (4.24) are nonempty for all sufficiently large k. Hence $\tilde{v}_p(U^k) < \infty$ for all sufficiently large k. Note that for each k, the dual problem to (4.24) is dual strictly feasible because of Assumption **A**3 and $\mathcal{COP}^n \subseteq (\mathrm{Diag}_+^n(U^k))^*$. Thus, $\tilde{v}_p(U^k)$ is indeed finite for all sufficiently large k. Moreover, thanks to the dual strict feasibility, the solution sets of (4.24) are nonempty whenever $\tilde{v}_p(U^k)$ is finite hence, in particular, are nonempty for all sufficiently large k.

Next, note that by Assumption **A**3 the dual problems of (4.24) for each k actually have a *common* Slater point, i.e., there exists a matrix

$$\bar{Y} := C - \sum_{i=1}^m \bar{y}_i A_i \in \mathrm{int}\,\mathcal{COP}^n \subseteq \mathrm{int}\,(\mathrm{Diag}_+^n(U^k))^*.$$

Therefore, there exists $\varepsilon > 0$ so that $\bar{Y} + \varepsilon\mathbf{B} \subseteq \mathrm{int}\,\mathcal{COP}^n$, where $\mathbf{B}$ is the unit closed ball centered at the origin (in Fröbenius norm). Consequently, for any $X \in \mathcal{CP}^n$, it holds that $\mathrm{tr}(\bar{Y}X) \geq \varepsilon\|X\|_F$. We now argue that the solution sets of (4.24) are uniformly bounded for all k. Indeed, fix any k so that the solution set of (4.24) is nonempty, and let X^k be a solution. Then X^k is a Lagrange multiplier for the dual problem. In particular,

$$\tilde{v}_p(U^k) = \max_y \left\{ b^T y + \mathrm{tr}\left(X^k \left[C - \sum_{i=1}^m y_i A_i \right] \right) \right\} \geq b^T\bar{y} + \mathrm{tr}(X^k\bar{Y}) \geq b^T\bar{y} + \varepsilon\|X^k\|_F,$$

where the last inequality holds because $X^k \in \mathrm{Diag}_+^n(U^k) \subseteq \mathcal{CP}^n$. Since $\{\tilde{v}_p(U^k)\}$ is nonincreasing, we conclude from the above inequality that $\{X^k\}$ can be bounded above by a constant independent of k. Thus, the solution sets of (4.24) are uniformly bounded for all k.

Finally, since the sequence of sets $\{\mathrm{Diag}_+^n(U^k)\}$ is monotonically increasing, we see from [48, Proposition 7.4(c)] that the objective function (with the constraint considered as the indicator function) of (4.24) epi-converges to that of (4.4) in the sense of [48, Definition 7.1]. The desired conclusion concerning limits of $\{\tilde{v}_p(U^k)\}$ and $\{X^k\}$ now follows from [48, Theorem 7.31(b)]. ∎

Since $\mathrm{Diag}_+^n(U) \subseteq \mathrm{SDD}_+^G(U)$ if the edges of G cover all nodes, we get the convergence of the sequence of problems (4.21) for a positively enlarging sequence $\{(G^k, U^k)\}$ under the same assumptions on U^k. But we can actually obtain the desired convergence result under a weaker condition.

Theorem 4.8.2. *Assume that* (4.4) *is strictly feasible. Let* $\{(G^k, U^k)\}$ *be a positively enlarging sequence such that*

$$\lim_{k\to\infty} \max_{x\in\Delta^n} \min_{y\in\mathrm{seg}(G^k,U^k)} \|x-y\| = 0. \tag{4.25}$$

Then it holds that:

(i) $v_p(G^k, U^k)$ *is finite for all sufficiently large* k *and* $\lim_{k\to\infty} v_p(G^k, U^k) = v_p$.

(ii) *The solution set of* (4.21) *with* $(G, U) = (G^k, U^k)$ *is nonempty and uniformly bounded for all sufficiently large* k.

(iii) *Let* X^k *be a solution of* (4.21) *with* $(G, U) = (G^k, U^k)$ *whenever the solution set is nonempty. Then any accumulation point of* $\{X^k\}$ *is a solution of* (4.4).

Proof. Note that from Theorem 4.7.4 and the description of $\mathrm{Diag}_+^n(U)$ as the conic hull of all matrices $u_i u_i^T$ where u_i is a row of U, if every node of G is covered by some edges and if we construct U' by adding rows such that each new row lies in $[u_i, u_j]$ for some $\{i, j\} \in \mathcal{E}$, we have $\mathrm{Diag}_+^n(U') \subseteq \mathrm{SDD}_+^G(U)$.

For each U^k, subdivide each segment $[u_i^k, u_j^k]$ into segments no longer than $1/k$, and adding these new points to U^k to form $\tilde{U}^k \in \mathbb{R}_+^{\tilde{t}_k \times n}$. Then for each $x \in \Delta^n$, we have

$$\min_{i=1,\ldots,\tilde{t}_k} \|x - \tilde{u}_i^k\| \le \min_{y\in\mathrm{seg}(G^k,U^k)} \|x-y\| + \frac{1}{k},$$

where $\tilde{u}_i^k$ is the i-th row of $\tilde{U}^k$. Thus, the sequence $\{\tilde{U}^k\}$ satisfies the conditions of Theorem 4.8.1. Consequently, from the proof of Theorem 4.8.1, the sequence of sets $\{\mathrm{Diag}_+^n(\tilde{U}^k)\}$ converges to $\mathcal{CP}^n$ in the sense of Painlevé-Kuratowski. In view of this and [48, Exercise 4.3(c)], $\{\mathrm{SDD}_+^n(U^k)\}$ converges to $\mathcal{CP}^n$. The rest of the proof follows exactly the same arguments as in the proof of Theorem 4.8.1. ■

An obvious way of guaranteeing the satisfaction of the condition (4.25) in Theorem 4.8.2 is to consider the rows of U^k to be the set of points in $x \in \Delta^n$ such that $kx \in \mathbb{Z}^n$, i.e. an equally spaced distribution of points in the simplex, with a growing number of points. This is in fact the strategy explored in [53] with the linear programming approach. As guaranteed by Theorem 4.8.2, this is sufficient to get convergence in our case, independently of the edges considered, but we can get away with much less. Indeed, it is easy to see, for example, that we do not need to map vertices to the interior of the simplex to get convergence and, in fact, it is enough to uniformly sample the *boundary* of the simplex, and form a graph with all possible edges between the chosen vertices. Finding embedded graphs that optimally cover Δ^n in the sense of minimizing the maximum distance to a point of the simplex seems to be a hard problem with no obvious answer, but many different strategies can be attempted. For practical purposes, it might be helpful to use the problem structure to design strategies for constructing $\{(G^k,U^k)\}$; these may not satisfy condition (4.25) and hence the convergence behavior can be compromised, but their corresponding problem (4.21) may be easier to solve. Indeed, as discussed in [36, Section 1.4], the amount of work per iteration for solving (4.22) is $\mathcal{O}((m+t_k^2)^2(4|\mathcal{E}|+t_k^2))$ when $(G,U)=(G^k,U^k)$. Hence, we will explore some problem-dependent inner approximation schemes in the next section.

Before ending this section, we would like to point out that the approach in [53] using $\mathrm{Diag}_+^n(U)$ for (rows of) U equally distributed in the simplex is one of the few problem-independent inner approximations to $\mathcal{CP}^n$ presented in the literature. The only other approach is that of [34], which leads to SDP problems. Although conceptually very interesting and with guaranteed convergence, this latter approach performs poorly in practice, because the size of the constraints grows very fast and the small instances that can be reasonably computed give weak approximations. In some sense, our SOCP based approximation schemes may lend some of the power of semidefinite programming to the LP approximation without completely sacrificing computability.

4.9 Problem-dependent inner approximation schemes

In this section, we propose some problem-dependent heuristic schemes for constructing $\{(G^k,U^k)\}$. They typically lead to computationally more tractable problems than a positively enlarging sequence satisfying (4.25). As we shall see later in our numerical experiments, these problem-dependent schemes in general return solutions with reasonable quality, though their convergence behaviors are still unknown. A related problem-dependent approach was developed in [1] for semidefinite programming. In there, they proposed the use of the cone $\mathrm{SDD}^n(U)$ and progressively enlarge the U to obtain efficient inner approximations to $\mathcal{S}_+^n$. We propose in this section a related approach. The main difference is that in the semidefinite case considered in [1], enlarging the U is relatively simple, as we can always separate the dual solution to the inner approximation from $\mathcal{S}_+^n$, if it is not there. In the case of completely positive cone, however, there is no realistic way of even checking if the dual solution is copositive. Thus, a direct separation procedure, like the one proposed in [1], is not viable.

4.9.1 Problem-dependent positively enlarging sequence

In this section, we describe a problem-dependent strategy for constructing a positively enlarging sequence $\{(G^k,U^k)\}$ that can potentially perform better on specific problem instances.

After solving (4.21) with a choice of (G^k,U^k), if the problem is feasible, one will obtain a solution $X\in\mathrm{SDD}_+^G(U)$. By Theorem 4.7.4, this X can be written as a conic combination of vv^T for $v\in\mathrm{seg}(G,U)$. Our plan here is to add these v as vertices to G and add some new edges from them, in order to increment the graph. The decomposition is not unique, so one has to carefully define what is meant by it.

First, note that for an $M\in\mathcal{S}_+^2\cap\mathcal{N}^2$, there exist $a\geq 0$, $b\geq 0$ and $v\in\mathbb{R}_+^2$ so that

$$M=vv^T+\begin{bmatrix} a & 0 \\ 0 & b\end{bmatrix}. \tag{4.26}$$

This is trivially true if any element in the diagonal of M is zero. For other matrices, the above decomposition can be realized by taking for example $v = (\sqrt{m_{11}}, m_{12}/\sqrt{m_{11}})$, implying $a = 0$ and $b = m_{22} - m_{12}^2/m_{11}$, which is greater than or equal to zero since $M \succeq 0$.

Now, for any $U \in \mathbb{R}_+^{t\times n}$, one can see that $U^T \iota_{ij}(M)U = au_iu_i^T + bu_ju_j^T + (v_1u_i + v_2u_j)(v_1u_i + v_2u_j)^T$, where u_i is the ith row of U, $1 \le i < j \le t$. So, besides the vertices u_i and u_j, we need at most one point coming from each edge $[u_i, u_j]$ to describe $U^T \iota_{ij}(M)U$. Since elements of $\mathrm{SDD}_+^G(U)$ are sums of matrices of this type for $\{i, j\} \in \mathcal{E}$ by Theorem 4.7.4, we have the following lemma refining Theorem 4.7.4.

Lemma 4.9.1. *Any element $X \in \mathrm{SDD}_+^G(U)$ can be written as*

$$X = \sum_{i=1}^{t} \lambda_i u_i u_i^T + \sum_{\{i,j\}\in\mathcal{E}} \gamma_{ij} w_{ij} w_{ij}^T$$

where u_i is the i-th row of $U \in \mathbb{R}_+^{t\times n}$, $w_{ij} \in [u_i, u_j]$ and $\lambda_i, \gamma_{ij} \ge 0$.

A natural question to ask is which points we can pick in each segment. To answer this question, we assume without loss of generality that $m_{12} > 0$ (and hence $m_{11} > 0$ and $m_{22} > 0$) in (4.26) and demonstrate how the v there can be chosen. Note that $U^T vv^T U$ is supposed to correspond to a $\gamma_{ij} w_{ij} w_{ij}^T$ in the decomposition in Lemma 4.9.1.

Since $m_{12} > 0$, we must have $v_1 > 0$ and $v_2 > 0$. Then we just need to see what the ratio $r = v_1/v_2$ can be. What we saw above right after (4.26) was the largest case, where we get $r = m_{11}/m_{12}$. The smallest it can get is attained by setting $v = (m_{12}/\sqrt{m_{22}}, \sqrt{m_{22}})$, which gives us $r = m_{12}/m_{22}$. These two values for r can be seen by noting that any extremal ratio v_1/v_2 for the v in (4.26) must correspond to $a = 0$ or $b = 0$. A balanced option, defined in a way that the ratio between diagonal entries of vv^T preserves the ratio between the diagonal entries of M, is to take

$$v = \sqrt{m_{12}} \begin{bmatrix} \left(\frac{m_{11}}{m_{22}}\right)^{\frac{1}{4}} \\ \left(\frac{m_{22}}{m_{11}}\right)^{\frac{1}{4}} \end{bmatrix}, \tag{4.27}$$

which corresponds to $r = \sqrt{m_{11}/m_{22}}$, the geometric mean of the largest and smallest possible ratios.

Based on these observations, we can now describe a general strategy for an iterative procedure to obtain upper bounds for (4.4).

Scheme 1: Successive upper bound scheme for (4.4)

Step 0. Start with a *complete* graph G^0 and its embedding (G^0, I) in Δ^n. Set $k = 0$ and $U^0 = I$.

Step 1. For an optimal solution X^k of (4.21) with $(G,U) = (G^k, U^k)$, apply Lemma 4.9.1 to obtain points w_{ij} for some $\{i, j\} \in \mathcal{E}' \subseteq \mathcal{E}$ such that X is a conic combination of $w_{ij}w_{ij}^T$ for $\{i, j\} \in \mathcal{E}'$ and $u_i u_i^T$ for the vertex i of G.

Step 2. Define a new graph embedding (G^{k+1}, U^{k+1}) by adding new vertices at the points w_{ij} (or at least some subset of them) and some new edges connecting those vertices to some of the previously defined ones, and possibly remove redundant edges and go to Step 1.

The general idea is therefore to, augment the graph at each step by adding some vertices in the edges that were active in the optimal solution and some edges incident with them. All the steps have, however, some subtleties that need to be addressed.

The initial embedding (G^0, U^0) is currently taken to be simply the embedding of K^n into the vertices of Δ^n, so that $\mathrm{SDD}_+^{G^0}(U^0) = \mathrm{SDD}_+^n$. If that is infeasible, however, the strategy does not work. Nevertheless, assuming strict feasibility of (4.4), we know from Theorem 4.8.1 that there is some small enough uniform simplicial partition of Δ^n that will make the problem feasible.

The decomposition obtained in Step 1 is not unique. There are two sources of variations. First, as discussed above, given a 2×2 semidefinite matrix M such that $\iota_{ij}(M)$ appears in the decomposition of X, we have some leeway on which point to pick in the edge $[u_i, u_j]$. Second, notice that even these matrices M are not uniquely defined. Since the matrices M will be a side result of the solution to (4.21), the choice of algorithm and the way the problem is encoded will have some impact in the decomposition. As for defining the v given the matrix M, we will use the balanced approach described above in (4.27) as it seems to perform well in practice.

The augmenting step (Step 2) is the most delicate of all. Different augmenting techniques will give rise to very different procedures. Here and in our numerical experiments, we consider two different approaches. We will present more implementation details in Section 4.10.

The Maximalist Approach: In this approach, we add some new vertices and then connect all vertices to form a complete graph. This is memory consuming and induces some redundancies: every node we add is in the middle of an already existing edge. Adding edges to those does not enlarge the cone $\mathrm{SDD}_+^G(U)$ and might lead to numerical inaccuracies, as we create multiple ways of writing points in a segment. Some pruning techniques could be applied.

The Adaptive Simplicial Partition Approach: This is mimicking the technique introduced in [15], which maintains the set of edges as that of a simplicial partition. At every step we would pick edges to subdivide and subdivide all the simplices containing that edge. The choice of nodes and edges to add to G^k in our approach is based on the solution we obtain from solving (4.21) for $(G,U) = (G^{k-1}, U^{k-1})$. This is different from [15], which relies solely on an outer approximation to guide the subdivision process.

Note that we do not have any guarantee of convergence for Scheme 1. However, geometrically one can see what must happen in order for the method to get stuck, i.e., for $\mathrm{SDD}_+^{G^k}(U^k) = \mathrm{SDD}_+^{G^{k+1}}(U^{k+1})$. As an immediate consequence of Theorem 4.7.4, this happens if and only if all the newly added edges in the embedding are contained in previously existing edges. This is because rank one nonnegative matrices are on the extreme rays of $\mathcal{CP}^n$ (see [6]). Thus, we see from Theorem 4.7.4 that $\mathrm{SDD}_+^{G^k}(U^k) = \mathrm{SDD}_+^{G^{k+1}}(U^{k+1})$ if and only if $\mathrm{seg}(G^k, U^k) = \mathrm{seg}(G^{k+1}, U^{k+1})$. This is an extremely strong condition, that implies essentially (depending on the scheme chosen to enlarge the graph) that the scheme gets stuck if for some iteration the optimal solution can be attained as a combination of only the nodes, and no elements from the edges. Or, in other words, the problem (4.21) has the same solution if we replace $\mathrm{SDD}_+^{G^k}(U^k)$ by $\mathrm{Diag}_+^n(U^k)$. On passing, we would like to point out that, in occasions where convergence is a serious concern, one can modify Step 2 of Scheme 1 by adding a random vertex in Δ^n in addition to those w_{ij}: this resulting scheme is guaranteed to converge in view of Theorem 4.8.2 if (4.4) is also strictly feasible.

4.9.2 A forgetfulness scheme

The use of a positively enlarging sequence $\{(G^k, U^k)\}$ can lead to large-scale SOCP problems when k is huge. As a heuristic to alleviate the computational complexity, we propose a simple forgetfulness scheme.

In this approach, we maintain the complete graph throughout. However, we always form U^k by appending only the newly generated vertices to U^0, which we choose to be the identity matrix. The details are described below.

Scheme 2: A forgetfulness upper bound scheme for (4.4)

Step 0. Start with a *complete* graph G^0 and its embedding (G^0, I) in Δ^n. Set $k = 0$ and $U^0 = I$.

Step 1. For an optimal solution X^k of (4.21) with $(G, U) = (G^k, U^k)$, apply Lemma 4.9.1 to obtain points w_{ij} for some $\{i, j\} \in \mathcal{E}' \subseteq \mathcal{E}$ such that X is a conic combination of $w_{ij}w_{ij}^T$ for $\{i, j\} \in \mathcal{E}'$ and $u_iu_i^T$ for the vertex i of G.

Step 2. Define a new graph embedding (G^{k+1}, U^{k+1}): starting with (G^0, I), add new vertices at the points w_{ij} and then add edges between each new vertex and all vertices in G^0. Go to Step 1.

Note that, in general, one cannot guarantee that the forgetfulness scheme is even monotonous, as we are dropping the factors $u_iu_i^T$ that were a part of the representation of the optimal solution X in Step 1. However, in most studied random instances in our numerical experiments, the forgetfulness scheme appears to be monotonous. The main reason could be that the algorithm tends to write X as a conic combination of just the matrices $w_{ij}w_{ij}^T$ for $\{i, j\} \in \mathcal{E}'$. When this happens, we are guaranteed that the next iteration will be non-increasing, but this need not always be the case.

4.10 Numerical simulations

In this section, we report on numerical experiments to test our proposed approaches. All experiments were performed in Matlab (R2017a) on a 64-bit PC with an Intel(R) Core(TM) i7-6700 CPU (3.40GHz) and 16GB RAM. We used the convex optimization software CVX [28], running the solver MOSEK to solve the conic optimization problems that arise. In our tests, we specifically consider the following strategies:

Δ-partition: In this approach, controlled by a parameter $k \geq 2$, we generate the vertices of the graph G^k as the $\binom{n+k-1}{k}$ vertices in the uniform subdivision of the simplex Δ^n into simplices of size $\frac{1}{k}\Delta^n$. We then add edges between two vertices whenever their supports differ by 2.

Note that by Theorem 4.8.2, if (4.4) is in addition strictly feasible, then $v_p(G^k, U^k)$ will be close to v_p for all sufficiently large k, so this strategy is guaranteed to converge as k increases.

Max: This is a variant of Scheme 1. Specifically, in Step 1, we decompose X^k as described in Lemma 4.9.1 using the balanced option given in (4.27). Then, in Step 2, we add all w_{ij} whose X^k_{ij} is sufficiently large as new vertices, and add edges between all vertices so that the new graph G^{k+1} is complete.

Max1: This is another variant of Scheme 1. Step 1 is the same as in **Max**. However, in Step 2, we *only* add the w_{ij} corresponding to the largest X^k_{ij} (if X^k_{ij} exceeds a certain threshold) as a new vertex. We then add edges between all vertices so that the new graph G^{k+1} is complete.

Adaptive Δ-partition: This is also an variant of Scheme 1. Step 1 is the same as in **Max**. For Step 2, the way of adding vertices is the same as in **Max1**. However, the way we add edges mimics the approach introduced in [15], which maintains the set of edges as that of a simplicial partition. Specifically, we subdivide the edge corresponding to the w_{ij} we added, and subdivide all the simplices containing that edge.

Forgetfulness: This is a variant of Scheme 2. We perform Step 1 as in **Max**. As for Step 2, we add all w_{ij} whose X^k_{ij} is sufficiently large as new vertices to the original graph G^0. We then add edges to join each newly added vertex to all vertices in G^0.

In Section 4.10.1, we compare the strategies **Max**, **Adaptive Δ-partition** and **Forgetfulness** on random instances of (4.4). We will also present results obtained via **Δ-partition** (with $k = 2$) as benchmark. In Section 4.10.2, we will first review the standard completely positive programming formulation of the stable set problem, and then examine how **Max1** performs for some standard test graphs.

4.10.1 Random instances

In order to test the performance of our method in a generic setting, we test it for randomly generated instances of problem (4.4). We generate our objective function by setting $C = M^T M$ where M is an

$n \times n$ matrix with i.i.d. standard Gaussian entries, guaranteeing strict feasibility of (4.5). Furthermore, we generate the constraints by setting $A_i = (M_i + M_i^T)/2$, where the M_i are also $n \times n$ matrices with i.i.d. standard Gaussian entries, and choosing b_i such that $b_i = \operatorname{tr}(A_i(E + nI))$. This guarantees strict feasibility of (4.4).

For the first of our tests we varied the number of variables, n, and the number of constraints m, so that n is either 10 or 25 and m is either 5, 10 or 15. Given the complexity of copositive programming, there is actually no reliable way to find the true solution for these problems and there is no available implemented method that can generate lower bounds with which to compare our results. As a work-around, throughout this section we will compare the results we obtain with the classical (and somewhat coarse) lower bound provided by replacing $\mathcal{CP}^n$ by $\mathcal{S}_+^n \cap \mathcal{N}^n$ in problem (4.4). We will use the difference of our approximations to this lower bound, normalized by dividing it by the bound, as a proxy for the quality of the methods, and will simply denote it by relative gap. Precisely, this quantity is defined by $\operatorname{gap}(x) = \frac{x - x^*}{|x^*|}$, where x is the objective value attained by the method being studied and x^* the doubly nonnegative lower bound. This makes it somewhat easier to compare different methods across different instances of the problem. The drawback is that the gap we compute is actually the sum of the gaps of the proposed method and the doubly nonnegative approximation, which we don't know how to independently estimate.

		Max		Adaptive Δ-Partition		Forgetfulness		Δ-Partition	
n	m	time (sec)	Relative Gap	time (sec)	Relative Gap	time (sec)	Relative Gap	time (sec)	Relative Gap
10	5	8.2	5.035e-02	25.3	6.362e-02	13.0	2.006e-02	4.6	4.620e-01
10	10	19.5	2.281e-02	25.3	7.920e-02	23.0	1.849e-02	4.5	4.095e-01
10	15	41.7	1.212e-02	27.6	8.207e-02	27.0	1.179e-02	5.0	2.995e-01
25	5	23.0	6.748e-01	55.1	5.828e-01	38.4	2.975e-01	—	—
25	10	45.8	4.660e-01	62.9	7.841e-01	52.7	2.020e-01	—	—
25	15	71.8	3.715e-01	56.1	8.565e-01	61.5	1.545e-01	—	—

Table 4.1 Comparison of different iterative approaches

The results obtained can be seen in Table 4.1, where we present both the average gaps and the average running time for the studied methods. A few technical details are needed to be able to replicate the experiment. The results presented are averages of 30 instances per parameter pair. Moreover we fix the maximum number of iterations for the Max, Adaptive Δ-partition and Forgetfulness schemes as, respectively, 5, 20 and 15 for $n = 10$ and 5, 15 and 12 for $n = 25$. This was done (in an ad hoc way) to try to keep the average execution time as similar as possible across iterative methods, so that a fair comparison can be made. Also, since the maximalist approach can occasionally explode in size, we also stop this approach early when $t_{k+1} > 200$ (Recall that $U^k \in \mathbb{R}_+^{t_k \times n}$ for all k). For the forgetfulness approach, we prune the U^k in each step by removing redundant rows: we compute $\delta_{ij}^k := \|u_i^k - u_j^k\|_1$, where u_i^k and u_j^k are the i-th and the j-th rows of U^k respectively, $j > i$, and discard u_j^k if $\delta_{ij}^k < 10^{-6}$. We also stop this approach early when $t_{k+1} > 200$ for the U^{k+1} after pruning. The static Δ-partition is not computed for $n = 25$ as it takes too long.

These results show that the forgetfulness scheme dominates the others in all categories as far as the relation quality/time is concerned. The relative gaps of the attained solutions jumps from between 1% and 2% for $n = 10$ to between 15% and 30% for $n = 25$. Once again, we stress that these are upper bounds for the forgetfulness scheme quality as well as for the doubly nonnegative approximation quality, and we cannot separate the contributions from each method.

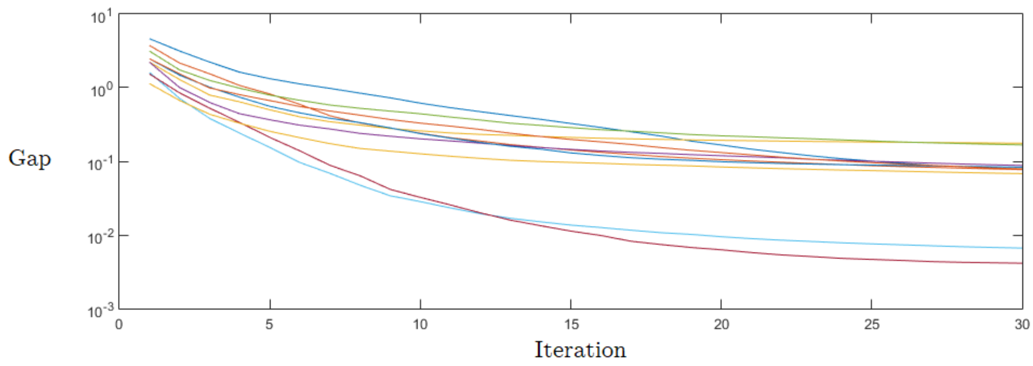

Fig. 4.5 Evolution of the gap for the forgetfulness scheme as iterations increase

We also plot in Figure 4.5 the evolutions of the gaps for the forgetfulness scheme for 10 random instances of the problem (4.4) with $n = 25$ and $m = 10$. We can see the logarithmic scale plot of the gap as iterations increase, and the diminishing returns in improvement percentage. Again, note that the true gap might actually be decreasing faster, as what we are seeing is the gap to the doubly nonnegative lower bound.

4.10.2 Stable set problems

While in the previous section we focus on random problems, the main focus of the completely positive/copositive programming literature has been in highly structured combinatorial optimization problems. One of the most common applications is to the stable set problem, i.e., the problem of finding in a graph G a set of vertices of maximal cardinality such that no two are connected with an edge. The cardinality of such a set is known as the independence number of G, denoted by $\alpha(G)$. In [18, Equation (8)], the following completely positive formulation was introduced for that problem.

$$\begin{aligned} \alpha(G) = \max \quad & \operatorname{tr}(EX) \\ \text{s.t.} \quad & \operatorname{tr}((A_G + I)X) = 1, \\ & X \in \mathcal{CP}^n, \end{aligned} \tag{4.28}$$

where A_G is the adjacency matrix of G.

In this setting we have a single constraint, so $m = 1$. Our inner approximations of $\mathcal{CP}^n$ will yield in this case lower bounds, from which one might be able to extract an actual feasible stable set with given cardinality. There are a number of good heuristic approaches to the stable set problem with good results, as there exist implementations of exact algorithms that can handle small to medium sized graphs, all performing necessarily much better than our all-purpose conic programming approach. However, we can still see how our approach performs on its own, to get some indication of its performance on low codimension structured problems.

In this class of problems, symmetry and structure likely imply that the growth of the matrix U in the greedier maximalist approach but also in the forgetfulness approach is too fast and adds too much redundancy. To avoid this phenomenon we take the Max1 approach: at every iteration we only add to U the vertex that has the largest weight in the solution found. This yields a greedy sort of algorithm, that in practice tends to grow the stable set greedily one by one. We stopped as soon as the greedy process got stuck and there was no improvement in two consecutive iterations.

We computed both stability numbers, $\alpha(G)$, and clique numbers, $\omega(G)$, which are simply the stability numbers of the complementary graph. Following [15], we started by computing the clique numbers of the graphs where their method was tested. Our method yields the correct answers in a relatively short time, as can be seen in Table 4.2, where our results are presented under the column “result", and the column “$\omega(G)$" corresponds to the known clique numbers. Note that this is not too surprising, as finding a large stable set, or clique, is in a general sense computationally easier than proving that a larger one does not exist. In other words, lower bounding the stable set and clique numbers of particular graphs tends to be easier than upper bounding them, so our problem has a smaller scope than what was attempted in [15], leading to much faster times. The graphs in the table come from two sources, the first is a 17 vertex graph from [42] that is notoriously hard for upper bounding by convex approximations, the other five come from the 2nd DIMACS implementation challenge test instances [32], and only hamming6-4 and johnson8-2-4 could be solved by Bundfuss and Dür’s method in less than two hours as reported in their paper [15].

graph	vertices	iterations	time(sec)	result	$\omega(G)$
pena17	17	5	13.8	6.0000	6
hamming6-2	64	31	836.7	32.0000	32
hamming6-4	64	3	64.0	4.0000	4
johnson8-2-4	28	3	11.7	4.0000	4
johnson8-4-4	70	13	322.5	14.0000	14
johnson16-2-4	120	7	637.0	8.0000	8

Table 4.2 Clique number for different graphs

To explore the limits of our approach we tried a few more instances of the stable set problem. We tried Paley graphs, known to mimic some properties of random graphs, with some degree of success, and a few small-sized instances of graphs derived from error correcting codes, available at [50]. The results are much worse in this family, with our algorithm failing in small instances, as can be seen in Table 4.3, where our results are reported under the column “result", and the true stability numbers are presented under the column “$\alpha(G)$". One word of caution is that the entire procedure is highly unstable, and simply changing the solver from MOSEK to SDPT3 can result in changes in the result, e.g. Paley$_{137}$ becomes exact in SDPT3.

graph	vertices	iterations	time(sec)	result	$\alpha(G)$
Paley_{137}	137	4	977.4	5.0000	7
Paley_{149}	149	6	1841.6	7.0000	7
Paley_{157}	157	6	2254.1	7.0000	7
1tc.16	16	6	15.7	7.0000	8
1tc.32	32	10	85.5	11.0000	12
1dc.64	64	7	235.8	8.0000	10
1dc.128	128	13	2491.0	14.0000	16
2dc.128	128	4	823.6	5.0000	5

Table 4.3 Stability number for different graphs

References

[1] Ahmadi, A. A., Dash, S., and Hall, G. (2017). Optimization over structured subsets of positive semidefinite matrices via column generation. *Discrete Optim.*, 24:129–151.

[2] Ahmadi, A. A. and Hall, G. (2017). Sum of squares basis pursuit with linear and second order cone programming. In *Algebraic and geometric methods in discrete mathematics*, volume 685 of *Contemp. Math.*, pages 27–53. Amer. Math. Soc., Providence, RI.

[3] Ahmadi, A. A. and Majumdar, A. (2019). DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization. *SIAM J. Appl. Algebra Geom.*, 3(2):193–230.

[4] Barker, G. P. and Carlson, D. (1975). Cones of diagonally dominant matrices. *Pacific J. Math.*, 57(1):15–32.

[5] Barvinok, A. (2002). *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.

[6] Berman, A. and Shaked-Monderer, N. (2003). *Completely positive matrices*. World Scientific Publishing Co., Inc., River Edge, NJ.

[7] Bertsimas, D., Freund, R. M., and Sun, X. A. (2013). An accelerated first-order method for solving SOS relaxations of unconstrained polynomial optimization problems. *Optim. Methods Softw.*, 28(3):424–441.

[8] Blekherman, G. (2006). There are significantly more nonnegative polynomials than sums of squares. *Israel J. Math.*, 153:355–380.

[9] Blekherman, G., Parrilo, P. A., and Thomas, R. R., editors (2013). *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA.

[10] Boman, E. G., Chen, D., Parekh, O., and Toledo, S. (2005). On factor width and symmetric H-matrices. *Linear Algebra Appl.*, 405:239–248.

[11] Bomze, I. M. and De Klerk, E. (2002). Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. volume 24, pages 163–185. Dedicated to Professor Naum Z. Shor on his 65th birthday.

[12] Bomze, I. M., Schachinger, W., and Uchida, G. (2012). Think co(mpletely)positive! Matrix properties, examples and a clustered bibliography on copositive optimization. *J. Global Optim.*, 52(3):423–445.

[13] Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press, Cambridge.

[14] Brualdi, R. A. (1998). *Introductory Combinatorics*. Prentice Hall, 3 edition.

[15] Bundfuss, S. and Dür, M. (2009). An adaptive linear approximation algorithm for copositive programs. *SIAM J. Optim.*, 20(1):30–53.

[16] Burer, S. (2009). On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.*, 120(2, Ser. A):479–495.

[17] Burer, S. (2012). Copositive programming. In *Handbook on semidefinite, conic and polynomial optimization*, volume 166 of *Internat. Ser. Oper. Res. Management Sci.*, pages 201–218. Springer, New York.

[18] de Klerk, E. and Pasechnik, D. V. (2002). Approximation of the stability number of a graph via copositive programming. *SIAM J. Optim.*, 12(4):875–892.

[19] de Klerk, E. and Sotirov, R. (2010). Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. *Math. Program.*, 122(2, Ser. A):225–246.

[20] Dickinson, P. J. C. (2011). Geometry of the copositive and completely positive cones. *J. Math. Anal. Appl.*, 380(1):377–395.

[21] Dickinson, P. J. C. (2013). *The copositive cone, the completely positive cone and their generalisations*. PhD thesis. Relation: https://www.rug.nl/ Rights: University of Groningen.

[22] Dür, M. (2010). Copositive programming – a survey. In Diehl, M., Glineur, F., Jarlebring, E., and Michiels, W., editors, *Recent Advances in Optimization and its Applications in Engineering*, pages 3–20, Berlin, Heidelberg. Springer Berlin Heidelberg.

[23] Fidalgo, C. and Kovačec, A. (2011). Positive semidefinite diagonal minus tail forms are sums of squares. *Math. Z.*, 269(3-4):629–645.

[24] Gatermann, K. and Parrilo, P. A. (2004). Symmetry groups, semidefinite programs, and sums of squares. *J. Pure Appl. Algebra*, 192(1-3):95–128.

[25] Gershgorin, S. (1931). Über die Abgrenzung der Eigenwerte einer Matrix. *Bull. Acad. Sci. URSS*, 1931(6):749–754.

[26] Gouveia, J., Kovačec, A., and Saee, M. (2019). On sums of k-nomial squares. *In preparation*.

[27] Gouveia, J., Pong, T. K., and Saee, M. (2018). Inner approximating the completely positive cone via the cone of scaled diagonally dominant matrices. *J. Global Optim, to appear, https://arxiv.org/abs/1807.00379*.

[28] Grant, M. and Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx.

[29] Hilbert, D. (1888). Ueber die Darstellung definiter Formen als Summe von Formenquadraten. *Math. Ann.*, 32(3):342–350.

[30] Horn, R. A. and Johnson, C. R. (2013). *Matrix analysis*. Cambridge University Press, Cambridge, second edition.

[31] Hurwitz, A. (1891). Ueber den Vergleich des arithmetischen und des geometrischen Mittels. *J. Reine Angew. Math.*, 108:266–268.

[32] Johnson, D. S. and Trick, M. A., editors (1996). *Cliques, coloring, and satisfiability*, volume 26 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*. American Mathematical Society, Providence, RI. Papers from the workshop held as part of the 2nd DIMACS Implementation Challenge in New Brunswick, NJ, October 11–13, 1993.

[33] Lasserre, J. B. (2011). A new look at nonnegativity on closed sets and polynomial optimization. *SIAM J. Optim.*, 21(3):864–885.

[34] Lasserre, J. B. (2014). New approximations for the cone of copositive matrices and its dual. *Math. Program.*, 144(1-2, Ser. A):265–276.

[35] Lasserre, J. B., Toh, K.-C., and Yang, S. (2017). A bounded degree SOS hierarchy for polynomial optimization. *EURO J. Comput. Optim.*, 5(1-2):87–117.

[36] Lobo, M. S., Vandenberghe, L., Boyd, S., and Lebret, H. (1998). Applications of second-order cone programming. volume 284, pages 193–228. ILAS Symposium on Fast Algorithms for Control, Signals and Image Processing (Winnipeg, MB, 1997).

[37] Lofberg, J. (2004). Yalmip : A toolbox for modeling and optimization in matlab. *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, pages 284–289.

[38] Motzkin, T. S. (1967). The arithmetic-geometric inequality. In *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)*, pages 205–224. Academic Press, New York.

[39] Muir, T. (1960). *A treatise on the theory of determinants*. Revised and enlarged by William H. Metzler. Dover Publications, Inc., New York.

[40] Nie, J. and Wang, L. (2012). Regularization methods for SDP relaxations in large-scale polynomial optimization. *SIAM J. Optim.*, 22(2):408–428.

[41] Parrilo, P. (2000). Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. *PhD Thesis, California Institute of Technology*.

[42] Peña, J., Vera, J., and Zuluaga, L. F. (2007). Computing the stability number of a graph via linear and semidefinite programming. *SIAM J. Optim.*, 18(1):87–105.

[43] Permenter, F. and Parrilo, P. (2018). Partial facial reduction: simplified, equivalent SDPs via approximations of the PSD cone. *Math. Program.*, 171(1-2, Ser. A):1–54.

[44] Ramana, M. and Goldman, A. J. (1995). Some geometric results in semidefinite programming. *J. Global Optim.*, 7(1):33–50.

[45] Reznick, B. (1978). Extremal PSD forms with few terms. *Duke Math. J.*, 45(2):363–374.

[46] Reznick, B. (1995). Uniform denominators in Hilbert's seventeenth problem. *Math. Z.*, 220(1):75–97.

[47] Rockafellar, R. T. (1970). *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J.

[48] Rockafellar, R. T. and Wets, R. J.-B. (1998). *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin.

[49] Schneider, R. (1993). *Convex bodies: the Brunn-Minkowski theory*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

[50] Sloane, N. (2005). Challenge problems: Independent sets in graphs. *Information Sciences Research Center, https://oeis.org/A265032/a265032.html*.

[51] Vallentin, F. (2009). Symmetry in semidefinite programs. *Linear Algebra Appl.*, 430(1):360–369.

[52] Vandenberghe, L. and Boyd, S. (1996). Semidefinite programming. *SIAM Rev.*, 38(1):49–95.

[53] Yıldırım, E. A. (2012). On the accuracy of uniform polyhedral approximations of the copositive cone. *Optim. Methods Softw.*, 27(1):155–173.

[54] Yıldırım, E. A. (2017). Inner approximations of completely positive reformulations of mixed binary quadratic programs: a unified analysis. *Optim. Methods Softw.*, 32(6):1163–1186.

[55] Zhao, X.-Y., Sun, D., and Toh, K.-C. (2010). A Newton-CG augmented Lagrangian method for semidefinite programming. *SIAM J. Optim.*, 20(4):1737–1765.