Copositive matrices, sums of squares and the stability number of a graph

Vargas, Luis Felipe; Laurent, Monique

Published in:
Polynomial Optimization, Moments, and Applications

Publication date:
2023

Document Version
Peer reviewed version

Citation for published version (APA):
Vargas, L. F., & Laurent, M. (2023). Copositive matrices, sums of squares and the stability number of a graph. In M. Kočvara, B. Mourrain, & C. Riener (Eds.), *Polynomial Optimization, Moments, and Applications* (Vol. 206, pp. 99-132). (Springer Optimization and its Applications; Vol. 206). Springer Cham.

Copositive matrices, sums of squares and the stability number of a graph

Luis Felipe Vargas and Monique Laurent

Abstract This chapter investigates the cone of copositive matrices, with a focus on the design and analysis of conic inner approximations for it. These approximations are based on various sufficient conditions for matrix copositivity, relying on positivity certificates in terms of sums of squares of polynomials. Their application to the discrete optimization problem asking for a maximum stable set in a graph is also discussed. A central theme in this chapter is understanding when the conic approximations suffice for describing the full copositive cone, and when the corresponding bounds for the stable set problem admit finite convergence.

1 Introduction

An $n \times n$ symmetric matrix M is said to be *copositive* if the associated quadratic form $x^T M x = \sum_{i,j=1}^{n} M_{ij} x_i x_j$ is nonnegative over the nonnegative orthant $\mathbb{R}^n_+$. The set of copositive matrices is a cone, the *copositive cone* COP_n, thus defined as

$$\mathrm{COP}_n = \{M \in \mathcal{S}^n : x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n_+\}. \tag{1}$$

Copositive matrices are a fundamental class of matrices that play an important role in several areas, including linear algebra and combinatorial matrix theory (see the

Luis Felipe Vargas
Centrum Wiskunde & Informatica, Science Park 123 1098 XG Amsterdam, Netherlands, e-mail: luis.vargas@cwi.nl

Monique Laurent
Centrum Wiskunde & Informatica, Science Park 123 1098 XG Amsterdam, Netherlands, and Tilburg University, Warandelaan 2, 5037 AB Tilburg, Netherlands, e-mail: m.laurent@cwi.nl

This work is supported by the European Union's Framework Programme for Research and Innovation Horizon 2020 under the Marie SkÅĆodowska-Curie Actions Grant Agreement No. 813211 (POEMA).

monograph [53]) and optimization (see, e.g., the overview [17]). Their relevance in optimization is illustrated by the fact that many hard combinatorial optimization problems can be formulated as linear optimization problems over the copositive cone (see, e.g., [6, 7, 13, 16, 22]). This is the case, in particular, for the problem of determining the maximum stable set in a graph, a topic that we will discuss in this chapter (see Section 5).

Hence the copositive cone has a broad modeling power. As a consequence it is a computationally hard object to work with: linear optimization over COP_n is an NP-hard problem and checking whether a matrix is copositive is a co-NP-complete problem [39]. Motivated by these hardness results, several hierarchies of conic inner approximations for COP_n have been introduced in the literature. A key ingredient in these approximations is to design tractable certificates that permit to certify that the quadratic form $x^T Mx$ is nonnegative over $\mathbb{R}^n_+$ and thus that the matrix M is copositive. These certificates are based on using sums of squares of polynomials as a “proxy" for global nonnegativity, which is motivated by the fact that sums of squares of polynomials can be modeled using semidefinite optimization (as recalled later in relation (18)).

Another possible approach to certify copositivity of a matrix M is to consider the quartic form

$$(x^{\circ 2})^T M x^{\circ 2} := \sum_{i,j=1}^{n} M_{ij} x_i^2 x_j^2 \tag{2}$$

and to design sum-of-squares based certificates that certify that $(x^{\circ 2})^T M x^{\circ 2}$ is nonnegative on the full space $\mathbb{R}^n$. In other words, one may rely on the following alternative definition of the copositive cone

$$\mathrm{COP}_n = \{M \in \mathcal{S}^n : (x^{\circ 2})^T M x^{\circ 2} \geq 0 \text{ for all } x \in \mathbb{R}^n\}, \tag{3}$$

where we let $x^{\circ 2} = (x_1^2, \ldots, x_n^2)$ denote the vector of squared variables.

As we will see in this chapter, these two equivalent definitions (1) and (3) of the copositive cone offer the starting point for the definition of several hierarchies of conic approximations. Our objective in this chapter is to discuss the relationships between these various hierarchies, their convergence properties, and their application to the maximum stable set problem in graphs. We now briefly describe the contents of this chapter.

Organization of the chapter

In Section 2 we introduce some general background about polynomial optimization and sums of squares of polynomials. In particular, in Section 2.1, we recall some important positivity certificates that permit to certify the nonnegativity of a polynomial on the nonnegative orthant and on compact semialgebraic sets. In Section 2.2

we describe how these positivity certificates are used to define hierarchies of bounds for polynomial optimization problems and, in Section 2.3, we recall a criterion that can be used to detect when the bounds have finite convergence.

In Section 3 we present several hierarchies of conic inner approximations for the copositive cone COP_n. These conic approximations are based on using different types of positivity certificates for the quadratic form $x^T M x$, or for the quartic form $(x^{\circ 2})^T M x^{\circ 2}$ from (2). Moreover, one considers positivity on the full space $\mathbb{R}^n$, on the nonnegative orthant $\mathbb{R}^n_+$, on the standard simplex $\Delta_n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$, or on the unit sphere $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1\}$.

In Section 3.1 we introduce the cones $C_n^{(r)}$ and $\mathcal{K}_n^{(r)}$, where, for $C_n^{(r)}$, one requires that the polynomial $(\sum_{i=1}^n x_i)^r x^T M x$ has nonnegative coefficients, and, for $\mathcal{K}_n^{(r)}$, one requires that the polynomial $(\sum_{i=1}^n x_i^2)^r (x^{\circ 2})^T M x^{\circ 2}$ is a sum of squares of polynomials. These two conic hierarchies are motivated by the representation results by Reznick (for positive polynomials on $\mathbb{R}^n$, Theorem 1) and by Pólya (for positive polynomials on $\mathbb{R}^n_+$, Theorem 2). In addition, the cones $Q_n^{(r)}$ are introduced as a simpler, but weaker variation of the cones $\mathcal{K}_n^{(r)}$.

In Section 3.2 we introduce the Lasserre-type cones $\mathrm{LAS}^{(r)}_{\Delta_n}$, $\mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$ and $\mathrm{LAS}^{(r)}_{\mathbb{S}^{n-1}}$, where, respectively, one now uses positivity certificates for the polynomial $x^T M x$ on the standard simplex Δ_n (using representations in the quadratic module or the preordering of Δ_n), and positivity certificates for the polynomial $(x^{\circ 2})^T M x^{\circ 2}$ on the unit sphere $\mathbb{S}^{n-1}$. The motivation for these cones now stems from the representation results by Schmüdgen (Theorem 3) and by Putinar (Theorem 4).

In Section 3.3 we explain in detail the relationships between these various hierarchies of conic approximations of the copositive cone (see Theorem 7).

Each of the above hierarchies of conic approximations covers the *interior* of the copositive cone, which follows from the above mentioned representation results. This raises naturally the question of whether some of these hierarchies are able to cover the *full* copositive cone (i.e., also its boundary). This question is the central theme of Section 4.

Section 4 is devoted to investigating exactness properties of the above hierarchies of cones, i.e., for which matrix sizes the hierarchies are able to cover the full copositive cone COP_n. This question is studied for the cones $\mathcal{K}_n^{(r)}$ in Section 4.1 and for the cones $\mathrm{LAS}^{(r)}_{\Delta_n}$ in Section 4.2. Section 4.3 is devoted to the exceptional case $n = 5$, where one can show that the hierarchy of cones $\mathcal{K}_5^{(r)}$ covers the full copositive cone COP_5.

Section 5 discusses the application of the various conic approximation hierarchies for COP_n to the design of upper bounds for the graph parameter $\alpha(G)$, defined as the maximum cardinality of a stable set in a graph G. In particular, the cones $C_n^{(r)}$ lead to the linear programming based parameters $\zeta^{(r)}(G)$, discussed in Section 5.1, and the cones $\mathcal{K}_n^{(r)}$ lead to the semidefinite bounds $\vartheta^{(r)}(G)$, discussed in Section 5.2. The main theme in this section is to investigate whether the parameters $\vartheta^{(r)}(G)$ do admit finite convergence to $\alpha(G)$ or, equivalently, whether a class of associated copositive matrices M_G belong to the union $\bigcup_r \mathcal{K}_n^{(r)}$. This question, which relates to a long standing conjecture by de Klerk and Pasechnik [13], is now settled in the affirmative and a sketch of proof is offered in this section.

We conclude with some observations and further research directions in the last Section 6.

Notation

Throughout we will use the following notation. For $n \in \mathbb{N}$ we set $[n] = \{1, 2, \ldots, n\}$. The nonnegative orthant is $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_1, \ldots, x_n \geq 0\}$, the standard simplex in $\mathbb{R}^n$ is defined as $\Delta_n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$, and the unit sphere in $\mathbb{R}^n$ is defined as $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1\}$. For $x \in \mathbb{R}^n$, the support of x is the set $\{i \in [n] : x_i \neq 0\}$ and we let $x^{\circ 2} := (x_1^2, \ldots, x_n^2)$ denote the vector of squared entries. We use the notation e to denote the all-ones vector (of appropriate size), so $e = (1, \ldots, 1)^T$. For a sequence $\alpha \in \mathbb{N}^n$, we set $|\alpha| := \sum_{i=1}^n \alpha_i$.

Throughout, $\mathcal{S}^n$ denotes the set of $n \times n$ symmetric matrices. We say that a matrix $M \in \mathcal{S}^n$ is positive semidefinite (denoted as $M \succeq 0$) if $x^T M x \geq 0$ for all $x \in \mathbb{R}^n$. The set of $n \times n$ positive semidefinite matrices is denoted by $\mathcal{S}^n_+$. The set of diagonal matrices with strictly positive diagonal entries is denoted by $\mathcal{D}^n_{++}$. We let I_n, J_n (or simply I, J) denote the identity matrix and the all-ones matrix in $\mathcal{S}^n$.

We denote by $\mathbb{R}[x_1, x_2, \ldots, x_n]$ the set of polynomials with real coefficients in n variables. Throughout we abbreviate $\mathbb{R}[x_1, \ldots, x_n]$ by $\mathbb{R}[x]$ when there is no ambiguity. Any polynomial is of the form $p = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha$, where only finitely many coefficients p_α are nonzero. Then $|\alpha|$ is the degree of the monomial $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and the degree of p, denoted $\deg(p)$, is the maximum degree of its terms $p_\alpha x^\alpha$ with $p_\alpha \neq 0$. We denote by $\mathbb{R}[x]_r$ the set of polynomials of degree at most r. A form, also known as a homogeneous polynomial, is a polynomial in which all its terms have the same degree.

Given a polynomial $f \in \mathbb{R}[x]$ and a set $K \subseteq \mathbb{R}^n$, we say that f is *nonnegative* (or *positive*) on the set K if $f(x) \geq 0$ for all $x \in K$, and we say that f is *strictly positive* on K if $f(x) > 0$ for all $x \in K$. Given a tuple of polynomials $h = (h_1, \ldots, h_l)$, the ideal generated by h is defined as $I(h) := \{\sum_{i=1}^l q_i h_i : q_i \in \mathbb{R}[x]\}$. Its truncation at degree r is defined as $I(h)_r := \{\sum_{i=1}^l q_i h_i : \deg(q_i h_i) \leq r \text{ for } i \in [m]\}$. We will in particular consider the case when $h = \sum_{i=1}^n x_i - 1$ or $h = \sum_{i=1}^n x_i^2 - 1$, that define the simplex Δ_n and the unit sphere $\mathbb{S}^{n-1}$, respectively. Then we use the shorthand notation $I_{\Delta_n} := I(\sum_{i=1}^n x_i - 1)$ and $I_{\mathbb{S}^{n-1}} := I(\sum_{i=1}^n x_i^2 - 1)$. Finally, we let $\Sigma := \{\sum_{i=1}^m q_i^2 : q_i \in \mathbb{R}[x]\}$ denote the cone of sums of squares of polynomials, and, for an integer $r \in \mathbb{N}$, $\Sigma_r = \Sigma \cap \mathbb{R}[x]_r$ is the subcone consisting of the sums of squares that have degree at most r.

2 Preliminaries on polynomial optimization, nonnegative polynomials and sums of squares

Polynomial optimization asks for minimizing a polynomial over a semialgebraic set. That is, given polynomials $f, g_1, \ldots, g_m, h_1, \ldots, h_l \in \mathbb{R}[x]$, the task is to find (or approximate) the infimum of the following problem

$$f^* = \inf_{x \in K} f(x), \tag{4}$$

where

$$K = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \text{ for } i = 1, \ldots, m \text{ and } h_i(x) = 0 \text{ for } i = 1, \ldots, l\} \tag{5}$$

is a semialgebraic set. Problem (4) can be equivalently rewritten as

$$f^* = \sup\{\lambda : f(x) - \lambda \geq 0 \text{ for all } x \in K\}. \tag{6}$$

In view of this new formulation, finding lower bounds for a polynomial optimization problem amounts to finding certificates that certain polynomials are nonnegative on the semialgebraic set K.

2.1 Sum-of-squares certificates for nonnegativity

Testing whether a polynomial is nonnegative on a semialgebraic set is hard in general. Even testing whether a polynomial is globally nonnegative (nonnegative on $K = \mathbb{R}^n$) is a hard task in general. An easy *sufficient* condition for a polynomial to be globally nonnegative is being a sum of squares. A polynomial $p \in \mathbb{R}[x]$ is said to be a *sum of squares* if it can be written as a sum of squares of other polynomials, i.e., if $p = q_1^2 + \cdots + q_m^2$ for some $q_1, \ldots, q_m \in \mathbb{R}[x]$. Hilbert [24, 25] showed that every nonnegative polynomial of degree $2d$ in n variables is a sum of squares in the following cases: $(2d, n)$=$(2d, 1)$, $(2, n)$, or $(4, 2)$. Moreover, he showed that for any other pair $(2d, n)$ there exist nonnegative polynomials that are not sums of squares. The first explicit example of a nonnegative polynomial that is not a sum of squares was given by Motzkin [37] in 1967.

The Motzkin polynomial is nonnegative, but not a sum of squares

The following polynomial in two variables is known as the Motzkin polynomial:

$$h(x, y) = x^4 y^2 + x^2 y^4 - 3x^2 y^2 + 1. \tag{7}$$

The Motzkin polynomial is nonnegative in $\mathbb{R}^2$. This can be seen, e.g., by using the Arithmetic-Geometric Mean inequality, which gives

$$\frac{x^4y^2 + x^2y^4 + 1}{3} \geq \sqrt[3]{x^4y^2 \cdot x^2y^4 \cdot 1} = x^2y^2.$$

However, $h(x, y)$ cannot be written as a sum of squares. This can be checked using "brute force": assume $h = \sum_i q_i^2$ and examine the coefficients on both sides (starting from the coefficients of the monomials x^6, y^6, etc.; see, e.g., [48]).
The Motzkin form is the homogenization of h, thus the homogeneous polynomial in three variables:

$$m(x, y, z) = x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6. \tag{8}$$

Hence, the Motzkin form is nonnegative on $\mathbb{R}^3$ and it cannot be written as a sum of squares.

In 1927 Artin [1] proved that any globally nonnegative polynomial f can be written as a sum of squares of rational functions, i.e., $f = \sum_i (\frac{p_i}{q_i})^2$ for some $p_i, q_i \in \mathbb{R}[x]$, solving affirmatively Hilbert's 17th problem. Equivalently, Artin's result shows that for any nonnegative polynomial f there exists a polynomial q such that $q^2 f \in \Sigma$. Such certificates are sometimes referred to as certificates "with denominator". The following result shows that, when f is homogeneous and strictly positive on $\mathbb{R}^n \setminus \{0\}$, the denominator can be chosen to be a power of $(\sum_{i=1}^n x_i^2)$.

Theorem 1 (Reznick [47])

Let $f \in \mathbb{R}[x]$ be a homogeneous polynomial such that $f(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. Then the following holds:

$$\Big(\sum_{i=1}^n x_i^2\Big)^r f \in \Sigma \quad \textit{for some } r \in \mathbb{N}. \tag{9}$$

Scheiderer [49] shows that the *strict* positivity condition can be omitted for $n = 3$: any *nonnegative* form f in three variables admits a certificate as in (9). On the negative side, this is not the case for $n \geq 4$: there exist nonnegative forms in $n \geq 4$ variables that do not admit a positivity certificate as in (9) (an example is given below).

Certificate for nonnegativity of the Motzkin polynomial

Let $h(x, y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$ be the Motzkin polynomial, which is nonnegative and not a sum of squares. However,

$$(x^2 + y^2)^2 h(x, y) = x^2y^2(x^2 + y^2 + 1)(x^2 + y^2 - 2)^2 + (x^2 - y^2)^2$$

is a sum of squares. This sum-of-squares certificate thus shows (again) that h is nonnegative on $\mathbb{R}^2$.

A nonnegative polynomial f such that $(\sum_{i=1}^n x_i^2)^r f \notin \Sigma$ for all $r \in \mathbb{N}$

Let $q(x, y, z, w) := m^2 + w^6 m$, where m is the Motzkin form from (8). Clearly, q is nonnegative on $\mathbb{R}^4$, as m is nonnegative on $\mathbb{R}^3$. Assume that there exists $r \in \mathbb{N}$ such that $(x^2 + y^2 + z^2 + w^2)^r q \in \Sigma$. Then, $p' := (x^2 + y^2 + z^2 + 1)^r q(x, y, z, 1) = (x^2 + y^2 + z^2 + 1)^r (m^2 + m)$ is also a sum of squares. As p' is a sum of squares, one can check that also its lowest degree homogeneous part is a sum of squares (see [32, Lemma 4]). However, the lowest degree homogeneous part of p' is m, which is not a sum of squares. Hence this shows that $(x^2 + y^2 + z^2 + w^2)^r q \notin \Sigma$ for all $r \in \mathbb{N}$.

Next, we give some positivity certificates for polynomials on semialgebraic sets. The following result shows the existence of a positivity certificate for polynomials that are strictly positive on the nonnegative orthant $\mathbb{R}^n_+$.

Theorem 2 (Pólya [44])

Let f be a homogeneous polynomial such that $f(x) > 0$ for all $x \in \mathbb{R}^n_+ \setminus \{0\}$. Then the following holds:

$$\Big(\sum_{i=1}^n x_i\Big)^r f \text{ has nonnegative coefficients} \quad \text{for some } r \in \mathbb{N}. \tag{10}$$

In addition, Castle, Powers, and Reznick [9] show that nonnegative polynomials on $\mathbb{R}^n_+$ with finitely many zeros (satisfying some technical properties) also admit a certificate as in (10).

Now we consider positivity certificates for polynomials restricted to compact semialgebraic sets. Let $g = \{g_1, \ldots, g_m\}$ and $h = \{h_1, \ldots, h_l\}$ be sets of polynomials and consider the semialgebraic set K defined as in (5). The *quadratic module* generated by g, denoted by $\mathcal{M}(g)$, is defined as

$$\mathcal{M}(g) := \Big\{ \sum_{i=0}^m \sigma_i g_i : \sigma_i \in \Sigma \text{ for } i = 0, 1, \ldots, m, \text{ and } g_0 := 1 \Big\}, \tag{11}$$

and the *preordering* generated by g, denoted by $\mathcal{T}(g)$, is defined as

$$\mathcal{T}(g) := \Big\{ \sum_{J \subseteq [m]} \sigma_J \prod_{i \in J} g_i : \sigma_J \in \Sigma \text{ for } J \subseteq \{1, \ldots m\}, \text{ and } g_\emptyset := 1 \Big\}. \tag{12}$$

Observe that, if for a polynomial f we have

$$f \in \mathcal{M}(g) + I(h), \tag{13}$$

$$\text{or } \ f \in \mathcal{T}(g) + I(h), \tag{14}$$

then f is nonnegative on K. Moreover, if a polynomial admits a certificate as in (13), then it also admits a certificate as in (14), because $\mathcal{M}(g) \subseteq \mathcal{T}(g)$.

Example

Consider the polynomial $p(x, y) = x^2 + y^2 - xy$ in two variables x, y. We show that p is nonnegative on $\mathbb{R}^2_+$ in two different ways. The following identities hold:

$$(x + y)p(x, y) = x^3 + y^3,$$
$$p(x, y) = (x - y)^2 + xy,$$

which both certify that p is nonnegative on $\mathbb{R}^2_+$. The first identity is a certificate as in (10): $x^3 + y^3$ has nonnegative coefficients. The second identity shows that $p \in \mathcal{T}(\{x, y\})$, i.e., gives a certificate as in (14).

The following two theorems show that, under certain conditions on the semialgebraic set K (and on the tuples g and h defining it), every strictly positive polynomial admits certificates as in (13) or (14).

Theorem 3 (Schmüdgen [50])

Let $K = \{x \in \mathbb{R}^n : g_i(x) \geq 0$ for $i \in [m], h_j(x) = 0$ for $j \in [l]\}$ be a compact semialgebraic set. Let $f \in \mathbb{R}[x]$ such that $f(x) > 0$ for all $x \in K$. Then we have $f \in \mathcal{T}(g) + I(h)$.

We say that the sets of polynomials $g = \{g_1, \ldots, g_m\}$ and $h = \{h_1, \ldots, h_l\}$ satisfy the *Archimedean condition* if

$$N - \sum_{i=1}^{n} x_i^2 \in \mathcal{M}(g) + I(h) \quad \text{for some } N \in \mathbb{N}. \tag{15}$$

Note this implies that the associated set K is compact. We have the following result.

Theorem 4 (Putinar [45])

Let $K = \{x \in \mathbb{R}^n : g_i(x) \geq 0$ for $i \in [m], h_j(x) = 0$ for $j \in [l]\}$ be a semialgebraic set. Assume the sets of polynomials $g = \{g_1, \ldots, g_m\}$ and $h = \{h_1, \ldots, h_l\}$ satisfy the Archimedean condition (15). Let $f \in \mathbb{R}[x]$ be such that $f(x) > 0$ for all $x \in K$. Then we have $f \in \mathcal{M}(g) + I(h)$.

Note that positivity certificates for a polynomial f as in Theorem 3 and Theorem 4 involve a representation of the polynomial f “without denominators".

2.2 Approximation hierarchies for polynomial optimization

Based on the result in Putinar’s theorem, Lasserre [28] proposed a hierarchy of approximations $(f^{(r)})_{r\in\mathbb{N}}$ for problem (4). Given an integer $r \in \mathbb{N}$, the *quadratic module truncated at degree r* (generated by the set $g = \{g_1, \ldots, g_m\}$) is defined as

$$\mathcal{M}(g)_r := \Big\{ \sum_{i=0}^{m} \sigma_i g_i : \sigma_i \in \Sigma_{r-\deg(g_i)} \text{ for } i \in \{0, 1, \ldots, m\}, \text{and } g_0 = 1 \Big\}, \tag{16}$$

and the parameter $f^{(r)}$ as

$$f^{(r)} := \sup\{\lambda : f - \lambda \in \mathcal{M}(g)_r + I(h)_r\}. \tag{17}$$

Clearly, $f^{(r)} \le f^{(r+1)} \le f^*$ for all $r \in \mathbb{N}$. The hierarchy of parameters $f^{(r)}$ is also known as *Lasserre sum-of-squares hierarchy* for problem (4).

Semidefinite programming and sums of squares

Consider a polynomial $p \in \mathbb{R}[x]_{2d}$. The following observation was made in [10]:

$$p \in \Sigma_{2d} \iff p = [x]_d^T M [x]_d \text{ for some } M \succeq 0, \tag{18}$$

where $[x]_d = (x^\alpha)_{|\alpha| \le d}$ denotes the vector of monomials with degree at most d.

Indeed, if $p \in \Sigma_{2d}$ then $p = \sum_{i=1}^m q_i^2$ for some $q_i \in \mathbb{R}[x]_d$. We can write $q_i = [x]_d^T v_i$ for an appropriate vector v_i. Then, we obtain $p = \sum_{i=1}^m q_i^2 = [x]_d^T (\sum_{i=1}^m v_i v_i^T)[x]_d^T = [x]_d^T M [x]_d$, where $M := \sum_{i=1}^m v_i v_i^T$ is a positive semidefinite matrix.

Conversely, assume $p = [x]_d^T M [x]_d$ with $M \succeq 0$. Then $M = \sum_{i=1}^m v_i v_i^T$ for some vectors $v_1, \ldots, v_m$. Hence, $p = \sum_{i=1}^m ([x]_d^T v_i)^2$ is a sum of squares.

So relation (18) shows that testing whether a given polynomial is a sum of squares can be modeled as a semidefinite program. There exist efficient algorithms for solving semidefinite programs (up to any arbitrary precision, and under some technical assumptions). See, e.g., [2, 11].

Under the Archimedean condition, by Putinar's theorem, we have asymptotic convergence of the Lasserre hierarchy: $f^{(r)} \to f^*$ as $r \to \infty$. We say that *finite convergence* holds if $f^{(r)} = f^*$ for some $r \in \mathbb{N}$. In general, finite convergence does not hold, as the following example shows.

A polynomial optimization problem without finite convergence

Consider the problem

$$\min \quad x_1 x_2 \quad \text{s.t.} \quad x \in \Delta_3, \text{ i.e., } x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_1 + x_2 + x_3 = 1.$$

We show that the Lasserre hierarchy for this problem does not have finite convergence. The optimal value is clearly 0 and is attained, for example, in $x = (0, 0, 1)$. Assume the Lasserre hierarchy has finite convergence. Then,

$$x_1 x_2 = \sigma_0 + \sum_{i=1}^{3} x_i \sigma_i + q(\sum_{i=1}^{3} x_i - 1), \tag{19}$$

for some $\sigma_i \in \Sigma$ for $i = 0, 1, 2, 3$ and $q \in \mathbb{R}[x]$. For a scalar $t \in (0, 1)$ define the vector $u_t := (t, 0, 1 - t) \in \Delta_3$. Now we evaluate equation (19) at $x + u_t$ and obtain

$$\begin{aligned} x_1x_2 + tx_2 = \sigma_0(x + u_t) + (x_1 + t)\sigma_1(x + u_t) + x_2\sigma_2(x + u_t) \\ +(x_3 + 1 - t)\sigma_3(x + u_t) + q(x + u_t)(x_1 + x_2 + x_3). \end{aligned}$$

for any fixed $t \in (0, 1)$. We compare the coefficients of the polynomials in x at both sides of the above identity. Observe that there is no constant term in the left hand side, so $\sigma_0(u_t)+t\sigma_1(u_t)+(1-t)\sigma_3(u_t) = 0$, which implies $\sigma_i(u_t) = 0$ for $i = 0, 1, 3$ as $\sigma_i \in \Sigma$ and thus $\sigma_i(u_t) \geq 0$. Then, for $i = 0, 1, 3$, the polynomial $\sigma_i(x+u_t)$ has no constant term, and thus it has no linear terms. Now, by comparing the coefficient of x_1 at both sides, we get $q(u_t) = 0$. Finally, by comparing the coefficient of x_2 at both sides, we get $t = \sigma_2(u_t)$ for all $t \in (0, 1)$. This implies $\sigma_2(u_t) = t$ as polynomials in the variable t. This is a contradiction because $\sigma_2(u_t)$ is a sum of squares in t.

2.3 Optimality conditions and finite convergence

In this section we recall a result of Nie [40] that guarantees finite convergence of the Lasserre hierarchy (17) under some assumptions on the minimizers of problem (4). This result builds on a result of Marshall [35, 36].

Let u be a local minimizer of problem (4) and let $J(u) := \{j \in [m] : g_j(u) = 0\}$ be the set of inequality constraints that are active at u. We say that the *constraint qualification condition* (abbreviated as CQC) holds at u if the set

$$G(u) := \{\nabla g_j(u) : j \in J(u)\} \cup \{\nabla h_i(u) : i \in [l]\}$$

is linearly independent. If CQC holds at u then there exist $\lambda_1, \dots, \lambda_l, \mu_1, \dots, \mu_m \in \mathbb{R}$ satisfying

$$\begin{aligned} \nabla f(u) = \sum_{i=1}^{l} \lambda_i \nabla h_i(u) + \sum_{j\in J(u)} \mu_j \nabla g_j(u), \quad \mu_j \geq 0 \text{ for } j \in J(u), \\ \mu_j = 0 \text{ for } j \in [m] \setminus J(u). \end{aligned}$$

If we have $\mu_j > 0$ for all $j \in J(u)$, then we say that the *strict complementarity condition* (abbreviated as SCC) holds. The Lagrangian function $L(x)$ is defined as

$$L(x) := f(x) - \sum_{i=1}^{l} \lambda_i h_i(x) - \sum_{j\in J(u)} \mu_j g_j(x).$$

Another (second order) necessary condition for u to be a local minimizer is the following inequality

$$v^T \nabla^2 L(u) v \geq 0 \text{ for all } v \in G(u)^\perp. \tag{SONC}$$

If it happens that the inequality (SONC) is strict, i.e., if

$$v^T \nabla^2 L(u) v > 0 \text{ for all } 0 \neq v \in G(u)^\perp, \tag{SOSC}$$

then one says that the *second order sufficiency condition* (SOSC) holds at u.

We can now state the following result by Nie [40].

Theorem 5 (Nie [40])

Assume that the Archimedean condition (15) holds for the polynomial sets g and h in problem (4). If the constraint qualification condition (CQC), the strict complementarity condition (SCC), and the second order sufficiency condition (SOSC) hold at every global minimizer of (4), then the Lasserre hierarchy (17) has finite convergence, i.e., $f^{(r)} = f^*$ *for some* $r \in \mathbb{N}$.

Nie [40] uses Theorem 5 to show that finite convergence of Lasserre hierarchy (17) holds generically. Note that the conditions in the above theorem imply that problem (4) has finitely many minimizers. So this result may help to show finite convergence only when there are finitely many minimizers. It will be used later in this chapter (for the proof of Theorem 17 and Theorem 24).

3 Sum-of-squares approximations for COP_n

As mentioned in the Introduction, optimizing over the copositive cone is a hard problem, this motivates to design tractable conic inner approximations for it. One classical cone that is often used as inner relaxation of COP_n is the cone SPN_n, defined as

$$\mathrm{SPN}_n := \{M \in \mathcal{S}^n : M = P + N \text{ where } P \succeq 0, N \geq 0\}. \tag{20}$$

In this section we explore several conic approximations for COP_n, strengthening SPN_n, based on sums of squares of polynomials. They are inspired by the positivity certificates (9), (10), (13), and (14) introduced in Section 2.

3.1 Cones based on Pólya's nonnegativity certificate

In view of relation (1), a matrix is copositive if the homogeneous polynomial $x^T M x$ is nonnegative on $\mathbb{R}^n_+$. Motivated by the nonnegativity certificate (10) in Pólya's theorem, de Klerk and Pasechnik [13] introduced the cones $\mathcal{C}_n^{(r)}$, defined as

$$\mathcal{C}_n^{(r)} := \Big\{M \in \mathcal{S}^n : \Big(\sum_{i=1}^n x_i\Big)^r x^T M x \text{ has nonnegative coefficients}\Big\} \tag{21}$$

for any $r \in \mathbb{N}$. Clearly, $\mathcal{C}_n^{(r)} \subseteq \mathcal{C}_n^{(r+1)} \subseteq \mathrm{COP}_n$. By Pólya's theorem (Theorem 2), the cones $\mathcal{C}_n^{(r)}$ cover the interior of COP_n, i.e., $\mathrm{int}(\mathrm{COP}_n) \subseteq \bigcup_{r\geq 0} \mathcal{C}_n^{(r)}$. This follows from the fact that $M \in \mathrm{int}(\mathrm{COP}_n)$ precisely when $x^T M x > 0$ for all $x \in \mathbb{R}^n_+ \setminus \{0\}$. The cones $\mathcal{C}_n^{(r)}$ were introduced in [13] for approximating the stability number of a graph, as we will see in Section 5.

In a similar way, in view of relation (3), a matrix is copositive if the homogeneous polynomial $(x^{\circ 2})^T M x^{\circ 2}$ is globally nonnegative. Parrilo [42] introduced the cones $\mathcal{K}_n^{(r)}$, that are defined by using certificate (9) as

$$\mathcal{K}_n^{(r)} := \Big\{M \in \mathcal{S}^n : \Big(\sum_{i=1}^n x_i^2\Big)^r (x^{\circ 2})^T M x^{\circ 2} \in \Sigma\Big\}. \tag{22}$$

Clearly, $\mathcal{C}_n^{(r)} \subseteq \mathcal{K}_n^{(r)} \subseteq \mathrm{COP}_n$, and thus $\mathrm{int}(\mathrm{COP}_n) \subseteq \bigcup_{r\geq 0} \mathcal{K}_n^{(r)}$. This inclusion also follows from Reznick's theorem (Theorem 1).

The following result by Peña, Vera and Zuluaga [55] gives information about the structure of the homogeneous polynomials f for which $f(x^{\circ 2})$ is a sum of squares. As a byproduct, this gives the reformulation for the cones $\mathcal{K}_n^{(r)}$ from relation (24) below.

Theorem 6 (Peña, Vera, Zuluaga [55])

Let $f \in \mathbb{R}[x]$ be a homogeneous polynomial with degree d. Then the polynomial $f(x^{\circ 2})$ is a sum of squares if and only if f admits a decomposition of the form

$$f = \sum_{\substack{S\subseteq[n], |S|\leq d \\ |S|\equiv d \pmod 2}} \sigma_S x^S \quad \textit{for some } \sigma_S \in \Sigma_{d-|S|}. \tag{23}$$

In particular, for any $r \geq 0$, we have

$$\mathcal{K}_n^{(r)} = \Big\{M \in \mathcal{S}^n : \Big(\sum_{i=1}^n x_i\Big)^r x^T M x = \sum_{\substack{S\subseteq[n], |S|\leq r+2 \\ |S|\equiv r \pmod 2}} \sigma_S x^S \quad \textit{for some } \sigma_S \in \Sigma_{r+2-|S|}\Big\}. \tag{24}$$

Alternatively, the cones $\mathcal{K}_n^{(r)}$ may be defined as

$$\mathcal{K}_n^{(r)} = \Big\{M \in \mathcal{S}^n : \Big(\sum_{i=1}^n x_i\Big)^r x^T M x = \sum_{\substack{\beta\in\mathbb{N}^n \\ |\beta|\leq r+2}} \sigma_\beta x^\beta \quad \text{for some } \sigma_\beta \in \Sigma_{r+2-|\beta|}\Big\}, \tag{25}$$

where, in (24), one replaces square-free monomials by arbitrary monomials. Based on this reformulation of the cones $\mathcal{K}_n^{(r)}$, Peña et.al. [55] introduced the cones $\mathcal{Q}_n^{(r)}$, defined as

$$\mathcal{Q}_n^{(r)} := \Big\{ M \in \mathcal{S}^n : \Big(\sum_{i=1}^n x_i\Big)^r x^T M x = \sum_{\substack{\beta\in\mathbb{N}^n \\ |\beta|=r,r+2}} \sigma_\beta x^\beta \quad \text{for some } \sigma_\beta \in \Sigma_{r+2-|\beta|} \Big\}. \tag{26}$$

So $\mathcal{Q}_n^{(r)}$ is a restrictive version of the formulation (25) for the cone $\mathcal{K}_n^{(r)}$, in which the decomposition only allows sums of squares of degree 0 and 2. Then, we have

$$C_n^{(r)} \subseteq \mathcal{Q}_n^{(r)} \subseteq \mathcal{K}_n^{(r)}, \tag{27}$$

and thus

$$\text{int}(\text{COP}_n) \subseteq \bigcup_{r\geq 0} C_n^{(r)} \subseteq \bigcup_{r\geq 0} \mathcal{Q}_n^{(r)} \subseteq \bigcup_{r\geq 0} \mathcal{K}_n^{(r)}. \tag{28}$$

As an application of (24) we obtain the following characterization of the cones $\mathcal{K}_n^{(r)}$ for $r = 0, 1$. A matrix $M \in \mathcal{S}^n$ belongs to $\mathcal{K}_n^{(0)}$ if and only if

$$x^T M x = \sigma + \sum_{1\leq i<j\leq n} c_{ij} x_i x_j$$

for some $\sigma \in \Sigma_2$ and some scalars $c_{ij} \geq 0$ for $1 \leq i < j \leq n$, and M belongs to $\mathcal{K}_n^{(1)}$ if and only if

$$\Big(\sum_{i=1}^n x_i\Big) x^T M x = \sum_{i=1}^n x_i \sigma_i + \sum_{1\leq i\leq j\leq k\leq n} c_{ijk} x_i x_j x_k, \tag{29}$$

for some $\sigma_i \in \Sigma_2$ for $i \in [n]$ and some scalars c_{ijk} for $1 \leq i \leq j \leq n$. From this, one can also derive the following result.

Lemma 1 (Characterization of the cones $\mathcal{K}_n^{(0)}$ and $\mathcal{K}_n^{(1)}$)

Let $M \in \mathcal{S}^n$ be a symmetric matrix. Then the following holds.

(1) M belongs to the cone $\mathcal{K}_n^{(0)}$ if and only if there exists a positive semidefinite matrix $P \succeq 0$ such that $P \leq M$. In other words,

$$\mathcal{K}_n^{(0)} = \{M \in \mathcal{S}^n : M = P + N \text{ for some } P \succeq 0 \text{ and } N \geq 0\} = \text{SPN}_n. \tag{30}$$

(2) M belongs to the cone $\mathcal{K}_n^{(1)}$ if and only if there exist symmetric matrices $P(i)$ for $i \in [n]$ satisfying the following conditions:

(i) $P(i) \succeq 0$ for all $i \in [n]$,
(ii) $P(i)_{ii} = M_{ii}$ for all $i \in [n]$,
(iii) $2P(i)_{ij} + P(j)_{ii} = 2M_{ij} + M_{ii}$ for all $i \neq j \in [n]$,
(iv) $P(i)_{jk} + P(j)_{ik} + P(k)_{ij} \leq M_{ij} + M_{ik} + M_{jk}$ for all distinct $i, j, k \in [n]$.

Claim (1) and the "if" part in (2) in the above lemma were already proved by Parrilo in [42]. The "only if" part in (2) was proved by Bomze and de Klerk in [5].

A matrix P is called to be a $\mathcal{K}^{(0)}$-certificate for M if $P \succeq 0$ and $P \leq M$. Now we show a result that relates the zeros of the form $x^T Mx$ with the kernel of its $\mathcal{K}^{(0)}$-certificates, which will be used later in the chapter.

Lemma 2 ([32])

Let $M \in \mathcal{K}_n^{(0)}$ and let P be a $\mathcal{K}^{(0)}$-certificate of M. If $x \in \mathbb{R}^n_+$ and $x^T Mx = 0$, then $Px = 0$ and $P[S] = M[S]$, where $S = \{i \in [n] : x_i > 0\}$ is the support of x.

Proof Since P is a $\mathcal{K}^{(0)}$-certificate there exists a matrix $N \geq 0$ such that $M = P + N$. Hence, $0 = x^T Mx = x^T Px + x^T Nx$. Then $x^T Px = 0 = x^T Nx$ as $P \succeq 0$ and $N \geq 0$. This implies $Px = 0$ since $P \succeq 0$. On the other hand, since $x^T Nx = 0$ and $N \geq 0$, we get $N_{ij} = 0$ for $i, j \in S$. Hence, $M[S] = P[S]$, as $M = P + N$. □

3.2 Lasserre-type approximation cones

Recall the definitions (1) and (3) of the copositive cone. Clearly, in (1), the nonnegativity condition for $x^T Mx$ can be restricted to the simplex Δ_n and, in (3), the nonnegativity condition for $(x^{\circ 2})^T Mx^{\circ 2}$ can be restricted to the unit sphere $\mathbb{S}^{n-1}$. Based on these observations, one can now use the positivity certificate (13) or (14) to certify the nonnegativity on Δ_n or $\mathbb{S}^{n-1}$. This leads naturally to defining the following cones (as done in [33]): for an integer $r \in \mathbb{N}$,

$$\mathrm{LAS}^{(r)}_{\Delta_n} := \Big\{ M \in \mathcal{S}^n : x^T Mx = \sigma_0 + \sum_{i=1}^n \sigma_i x_i + q \ \text{ for } \sigma_0 \in \Sigma_r, \sigma_i \in \Sigma_{r-1},\ q \in I_{\Delta_n} \Big\}, \tag{31}$$

$$\mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}} = \Big\{ M \in \mathcal{S}^n : x^T Mx = \sum_{S\subseteq[n], |S|\leq r} \sigma_S x^S + q \ \text{ for } \sigma_S \in \Sigma_{r-|S|} \text{ and } q \in I_{\Delta_n} \Big\}, \tag{32}$$

$$\mathrm{LAS}^{(r)}_{\mathbb{S}^{n-1}} = \Big\{ M \in \mathcal{S}^n : (x^{\circ 2})^T Mx^{\circ 2} = \sigma + q \text{ for some } \sigma \in \Sigma_r, q \in I_{\mathbb{S}^{n-1}} \Big\}. \tag{33}$$

Clearly, we have $\mathrm{LAS}^{(r)}_{\Delta_n} \subseteq \mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$ and, by Putinar's theorem (Theorem 4),

$$\mathrm{int}(\mathrm{COP}_n) \subseteq \bigcup_{r\geq 0} \mathrm{LAS}^{(r)}_{\Delta_n}, \qquad \mathrm{int}(\mathrm{COP}_n) \subseteq \bigcup_{r\geq 0} \mathrm{LAS}^{(r)}_{\mathbb{S}^{n-1}}. \tag{34}$$

3.3 Links between the various approximation cones for COP_n

In this section, we link the various cones introduced in the previous sections.

Theorem 7 ([33])

Let $r \geq 2$ and $n \geq 1$. Then the following holds.

$$\mathrm{LAS}^{(r)}_{\Delta_n} \subseteq \mathcal{K}^{(r-2)}_n = \mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}} = \mathrm{LAS}^{(2r)}_{\mathbb{S}^{n-1}}. \tag{35}$$

So, this result shows that membership in the cones $\mathcal{K}^{(r)}_n$ can be characterized via positivity certificates on $\mathbb{R}^n_+$ or $\mathbb{R}^n$ of Pólya- and Reznick-type (using a 'denominator' of the form $(\sum_i x_i)^r$ for some $r \in \mathbb{N}$), or, alternatively, via 'denominator-free' positivity certificates on the simplex or the sphere of Schmüdgen- and Putinar-type.

Theorem 7 was implicitly shown in [31, Corollary 3.9]. We now sketch the proof. First, the equality $\mathcal{K}^{(r-2)}_n = \mathrm{LAS}^{(2r)}_{\mathbb{S}^{n-1}}$ follows from the following result.

Theorem 8 (de Klerk, Laurent, Parrilo [12])

Let f be a homogeneous polynomial of degree $2d$ and $r \in \mathbb{N}$. Then, we have $(\sum_{i=1}^n x_i^2)^r f \in \Sigma$ if and only if $f = \sigma + u(\sum_{i=1}^n x_i^2 - 1)$ for some $\sigma \in \Sigma_{2r+2d}$ and $u \in \mathbb{R}[x]$.

In particular, for any $r \geq 2$, we have

$$\mathrm{LAS}^{(2r)}_{\mathbb{S}^{n-1}} = \Big\{M \in \mathcal{S}^n : \Big(\sum_{i=1}^n x_i^2\Big)^{r-2} (x^{\circ 2})^T M x^{\circ 2} \in \Sigma\Big\} = \mathcal{K}^{(r-2)}_n. \tag{36}$$

Next, the inclusion $\mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}} \subseteq \mathrm{LAS}^{(2r)}_{\mathbb{S}^{n-1}}$ follows by replacing x by $x^{\circ 2}$ in the definition of $\mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$. Indeed, if $M \in \mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$, then

$$x^T M x = \sum_{S \subseteq [n], |S| \leq r} \sigma_S x^S + q\Big(\sum_{i=1}^n x_i - 1\Big) \text{ for } \sigma_S \in \Sigma_{|S|-r}, q \in \mathbb{R}[x].$$

Then, by replacing x by $x^{\circ 2}$, we obtain

$$(x^{\circ 2})^T M x^{\circ 2} = \sum_{\substack{S \subseteq [n] \\ |S| \leq r}} \sigma_S(x^{\circ 2}) \prod_{i \in S} x_i^2 + q(x^{\circ 2})\Big(\sum_{i=1}^n x_i^2 - 1\Big) \text{ for } \sigma_S \in \Sigma_{|S|-r}, q \in \mathbb{R}[x],$$

where the first summation is a sum of squares of degree at most $2r$, thus showing that $M \in \mathrm{LAS}^{(2r)}_{\mathbb{S}^{n-1}}$.

Finally, as the inclusion $\mathrm{LAS}^{(r)}_{\Delta_n} \subseteq \mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$ is clear, it remains to show that $\mathcal{K}^{(r-2)}_n \subseteq \mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$ in order to conclude the proof of Theorem 7. For this, we use the formulation (24) of the cones $\mathcal{K}^{(r)}_n$. Let $M \in \mathcal{K}^{(r-2)}_n$, then

$$\Big(\sum_{i=1}^n x_i\Big)^{r-2} x^T M x = \sum_{\substack{S\subseteq[n], |S|\le r\\ |S|\equiv r \pmod 2}} \sigma_S x^S \quad \text{for some } \sigma_S \in \Sigma_{r-|S|}.$$

Write $\sum_{i=1}^n x_i = (\sum_{i=1}^n x_i - 1) + 1$ and expand $(\sum_{i=1}^n x_i)^r$ as $1 + p(\sum_{i=1}^n x_i - 1)$ for some $p \in \mathbb{R}[x]$. From this, setting $q = -px^T Mx$, we obtain

$$x^T M x = \sum_{\substack{S\subseteq[n], |S|\le r\\ |S|\equiv r \pmod 2}} \sigma_S x^S + q\Big(\sum_{i=1}^n x_i - 1\Big) \quad \text{for some } \sigma_S \in \Sigma_{r+2-|S|},\ q \in \mathbb{R}[x],$$

which shows $M \in \mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$.

It is useful to note that, in the formulation (32) of $\mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$, we could equivalently require a decomposition of the form

$$x^T M x = \sum_{\beta\in\mathbb{N}^n, |\beta|\le r} \sigma_\beta x^\beta + q \quad \text{for some } \sigma_\beta \in \Sigma_{r-|\beta|} \text{ and } q \in I_{\Delta_n}, \tag{37}$$

thus using arbitrary monomials x^β instead of square-free monomials x^S. This allows to draw a parallel with the definitions of the cones $\mathcal{C}_n^{(r)}$ (in (21)) and $\mathcal{Q}_n^{(r)}$ (in (26)). Namely, using the same type of arguments as above, one can obtain the following analogous reformulations for the cones $\mathcal{C}_n^{(r)}$ and $\mathcal{Q}_n^{(r)}$:

$$\mathcal{Q}_n^{(r)} = \Big\{M \in \mathcal{S}^n : x^T M x = \sum_{\substack{\beta\in\mathbb{N}^n\\ |\beta|=r,r+2}} \sigma_\beta x^\beta + q \text{ for } \sigma_\beta \in \Sigma_{r+2-|\beta|} \text{ and } q \in I_{\Delta_n}\Big\}, \tag{38}$$

$$\mathcal{C}_n^{(r)} = \{M \in \mathcal{S}^n : x^T M x = \sum_{\substack{\beta\in\mathbb{N}^n\\ |\beta|=r+2}} c_\beta x^\beta + q \text{ for } c_\beta \ge 0 \text{ and } q \in I_{\Delta_n}\}. \tag{39}$$

Seeing all cones as restrictive Schmüdgen-type representations of $x^T Mx$

We illustrate how membership in the cones $\mathrm{LAS}^{(r)}_{\Delta_n}$, $\mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$, $\mathcal{C}_n^{(r)}$, and $\mathcal{Q}_n^{(r)}$ can also be viewed as 'restrictive' versions of membership in the cone $\mathcal{K}_n^{(r-2)}$. Indeed, as we saw above, $\mathcal{K}_n^{(r-2)} = \mathrm{LAS}^{(r)}_{\Delta_n,\mathcal{T}}$ and thus a matrix M belongs to $\mathcal{K}_n^{(r-2)}$ if and only if the form $x^T Mx$ has a decomposition of the form (37). Then, membership in the cones $\mathrm{LAS}^{(r)}_{\Delta_n}$, $\mathcal{C}_n^{(r-2)}$, and $\mathcal{Q}_n^{(r-2)}$ corresponds to restricting to decompositions that allow only some terms in (37):

$$\underbrace{\sigma_0 + \sum_{i=1}^{n} x_i\sigma_i}_{\text{for cones } \mathrm{LAS}^{(r)}_{\Delta_n}} + \cdots + \overbrace{\sum_{\beta\in\mathbb{N}^n,|\beta|=r-2} x^{\beta}\sigma_{\beta} + \underbrace{\sum_{\beta\in\mathbb{N}^n,|\beta|=r} x^{\beta} c_{\beta}}_{\text{for cones } \mathcal{C}_n^{(r-2)}}}^{\text{for cones } \mathcal{Q}_n^{(r-2)}} + \underbrace{q(\sum_{i=1}^{n} x_i - 1)}_{\text{for cones } \begin{cases}\mathrm{LAS}^{(r)}_{\Delta_n}\\ \mathcal{Q}_n^{(r-2)}\\ \mathcal{C}_n^{(r-2)}\end{cases}} \tag{40}$$

4 Exactness of sum-of-squares approximations for COP_n

We have discussed several hierarchies of conic inner approximations for the copositive cone COP_n. In particular, we have seen that each of them covers the interior of COP_n. In this section, we investigate the question of deciding exactness of these hierarchies, where we say that a hierarchy of conic inner approximations is *exact* if it covers the full copositive cone COP_n.

4.1 Exactness of the conic approximations $\mathcal{K}_n^{(r)}$

We first recall a result from [14], that shows equality in the inclusion $\mathcal{K}_n^{(0)} \subseteq \mathrm{COP}_n$ for $n \leq 4$.

Theorem 9 (Diananda [14])

For $n \leq 4$ we have

$$\mathrm{COP}_n = \{M \in \mathcal{S}^n : M = P + N \textit{ for some } P \succeq 0, N \geq 0\} = \mathcal{K}_n^{(0)} \ (= \mathrm{SPN}_n).$$

This result does not extend to matrix size $n \geq 5$. For instance, as we now see, the *Horn matrix H* in (41) is copositive, but it does not belong to $\mathcal{K}_5^{(0)}$.

The Horn matrix

The Horn matrix

$$H := \begin{pmatrix} 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 \end{pmatrix} \tag{41}$$

is copositive. A direct way to show this is to observe that $H \in \mathcal{K}_n^{(1)}$. Parrilo [42] shows this latter fact by giving the following explicit sum of squares decomposition:

$$\begin{aligned}\Big(\sum_{i=1}^5 x_i^2\Big)(x^{\circ 2})^T H x^{\circ 2} = \; & x_1^2(x_1^2+x_2^2+x_5^2-x_3^2-x_4^2)^2 \\ & + x_2^2(x_1^2+x_2^2+x_3^2-x_4^2-x_5^2)^2 \\ & + x_3^2(x_2^2+x_3^2+x_4^2-x_5^2-x_1^2)^2 \\ & + x_4^2(x_3^2+x_4^2+x_5^2-x_1^2-x_2^2)^2 \\ & + x_5^2(x_1^2+x_4^2+x_5^2-x_2^2-x_3^2)^2 \\ & + 4x_1^2x_2^2x_5^2 + 4x_1^2x_2^2x_3^2 \\ & + 4x_2^2x_3^2x_4^2 + 4x_3^2x_4^2x_5^2 \\ & + 4x_4^2x_5^2x_1^2.\end{aligned} \tag{42}$$

On the other hand, Hall and Newman [23] show that H does not belong to SPN_5 $(= \mathcal{K}_5^{(0)})$. We give a short proof of this fact, based on Lemma 2.

Theorem 10 (Hall, Newman [23])

The Horn matrix H does not belong to $\mathcal{K}_5^{(0)}$. Hence, the inclusion $\mathcal{K}_n^{(0)} \subseteq \mathrm{COP}_n$ is strict for any $n \geq 5$.

Proof Assume, by way of contradiction, that $H \in \mathcal{K}_5^{(0)}$. Let P be a $\mathcal{K}^{(0)}$-certificate for H, i.e., such that $P \geq 0$ and $P \leq H$, and let $C_1, C_2, \dots, C_5$ denote the columns of P. Observe that $u_1 = (1,0,1,0,0)$ and $u_2 = (1,0,0,1,0)$ are zeros of the form $x^T Hx$. Then, by Lemma 2, $Pu_1 = Pu_2 = 0$. Hence, $C_1 + C_3 = C_1 + C_4 = 0$, so that $C_3 = C_4$. Using an analogous argument we obtain that $C_1 = C_2 = \dots = C_5$, which implies $P = tJ$ for some scalar $t \geq 0$, where J is the all-ones matrix. This leads to a contradiction since $P \leq H$. ◻

Next, we recall a result of Dickinson, Dür, Gijben and Hildebrand [15] that shows exactness of the conic approximation $\mathcal{K}_5^{(1)}$ for copositive matrices with an all-ones diagonal.

Theorem 11 (Dickinson, Dür, Gijben, Hildebrand [15])

Let $M \in \mathrm{COP}_5$ with $M_{ii} = 1$ for all $i \in [5]$. Then $M \in \mathcal{K}_5^{(1)}$.

In contrast, the same authors show that the cone COP_n is never equal to a single cone $\mathcal{K}_n^{(r)}$ for $n \geq 5$.

Theorem 12 (Dickinson, Dür, Gijben, Hildebrand [15])

For any $n \geq 5$ and $r \geq 0$, we have $\mathrm{COP}_n \neq \mathcal{K}_n^{(r)}$.

Proof Let M be a copositive matrix that lies outside $\mathcal{K}_n^{(0)}$. Clearly, any positive diagonal scaling of M remains copositive, that is, $DMD \in \mathrm{COP}_n$ for any $D \in \mathcal{D}_{++}^n$. We will show that for any $r \geq 0$ there exists a diagonal matrix $D \in \mathcal{D}_{++}^n$ such that $DMD \notin \mathcal{K}_n^{(r)}$. Fix $r \geq 0$ and assume, by way of contradiction, that $DMD \in \mathcal{K}_n^{(r)}$ for any positive diagonal matrix D. Then, for all scalars $d_1, d_2, \ldots, d_n > 0$ the polynomial $(\sum_{i=1}^n x_i^2)^r(\sum_{i,j=1}^n M_{ij} d_i d_j x_i^2 x_j^2)$ is a sum of squares. Equivalently, the polynomial $(\sum_{i=1}^n d_i^{-1} z_i^2)^r(\sum_{i,j=1}^n M_{ij} z_i^2 z_j^2)$ is a sum of squares in the variables $z_i = \sqrt{d_i} x_i$ $(i = 1, \ldots, n)$. Now we fix $d_1 = 1$ and we let $d_i \to \infty$ for $i = 2, \ldots, n$. Since the cone of sums of squares of polynomials is closed (see, e.g., [30, Section 3.8]), the limit polynomial $(z_1^2)^r(\sum_{i,j=1}^n M_{i,j} z_i^2 z_j^2)$ is also a sum of squares in the variables $z_1, \ldots, z_n$. Say $(z_1^2)^r(\sum_{i,j=1}^n M_{i,j} z_i^2 z_j^2) = \sum_{k=1}^m q_k^2$. Then, for each k, we have $q_k(z) = 0$ whenever $z_1 = 0$. Hence, if $r \geq 1$, then z_1 can be factored out from q_k, and we obtain that $(z_1^2)^{r-1}(\sum_{i,j=1}^n M_{i,j} z_i^2 z_j^2)$ is also a sum of squares. After repeatedly using this argument we can conclude that $\sum_{i,j=1}^n M_{i,j} z_i^2 z_j^2$ is a sum of squares, that is, $M \in \mathcal{K}_n^{(0)}$, leading to a contradiction.

As was recalled earlier, sums of squares of polynomials can be expressed using semidefinite programming. Hence, the cone $\mathcal{K}_n^{(r)}$ is *semidefinite representable*, which means that membership in it can be modeled using semidefinite programming. In [4] it is shown that COP_5 is *not* semidefinite representable, which is thus a stronger result that implies Theorem 12. On the other hand, it was shown recently in [52] that every 5×5 copositive matrix belongs to the cone $\mathcal{K}_5^{(r)}$ for some $r \in \mathbb{N}$.

Theorem 13 (Laurent, Vargas [33]; Schweighofer, Vargas [52])
We have $\mathrm{COP}_5 = \bigcup_{r \geq 0} \mathcal{K}_5^{(r)}$.

We will return to this result in Section 4.3, where we will give some hints on the strategy and tools that are used for the proof.

It is known that the result from Theorem 13 does not extend to matrix size $n \geq 6$. To show this, we recall the following result.

Proposition 1 ([32])
Let $M_1 \in \mathrm{COP}_n$ *and* $M_2 \in \mathrm{COP}_m$ *be two copositive matrices. Assume* $M_1 \notin \mathcal{K}_n^{(0)}$ *and there exists* $0 \neq z \in \mathbb{R}_+^m$ *such that* $z^T M_2 z = 0$. *Then we have*

$$\begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} \in \mathrm{COP}_{n+m} \setminus \bigcup_{r \in \mathbb{N}} \mathcal{K}_{n+m}^{(r)}. \tag{43}$$

Now we give explicit examples of copositive matrices of size $n \geq 6$ that do not belong to any of the cones $\mathcal{K}_n^{(r)}$.

Examples of copositive matrices outside $\bigcup_{r \geq 0} \mathcal{K}_n^{(r)}$

Let $M_1 = H$ be the Horn matrix, known to be copositive with $H \notin \mathcal{K}_n^{(0)}$. For the matrix M_2 we first consider the 1×1 matrix $M_2 = 0$ and, as a second example, we consider $M_2 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \in \mathrm{COP}_2$. Then, as an application of Proposition 1, we obtain

$$\left(\begin{array}{c|c} H & 0 \\ \hline 0 & 0 \end{array}\right) \in \mathrm{COP}_6 \setminus \bigcup_{r\in\mathbb{N}} \mathcal{K}_6^{(r)}, \quad \left(\begin{array}{c|cc} H & \multicolumn{2}{c}{0} \\ \hline 0 & 1 & -1 \\ & -1 & 1 \end{array}\right) \in \mathrm{COP}_7 \setminus \bigcup_{r\in\mathbb{N}} \mathcal{K}_7^{(r)}. \tag{44}$$

The leftmost matrix in (44) is copositive, it has all its diagonal entries equal to 0 or 1, and it does not belong to any of the cones $\mathcal{K}_6^{(r)}$. Selecting for M_2 the zero matrix of size $m \geq 1$ gives a matrix in $\mathrm{COP}_n \setminus \bigcup_{r\geq 0} \mathcal{K}_n^{(r)}$ for any size $n \geq 6$. The rightmost matrix in (44) is copositive, it has all its diagonal entries equal to 1, and it does not lie in any of the cones $\mathcal{K}_7^{(r)}$. More generally, if we select the matrix $M_2 = \frac{1}{m-1}(mI_m - J_m)$, which is positive semidefinite with $e^T M_2 e = 0$, then we obtain a matrix in $\mathrm{COP}_n \setminus \bigcup_{r\geq 0} \mathcal{K}_n^{(r)}$ with an all-ones diagonal for any size $n \geq 7$. In contrast, as mentioned in Theorem 11, any copositive 5×5 matrix with an all-ones diagonal belongs to $\mathcal{K}_5^{(1)}$. The situation for the case of 6×6 copositive matrices remains open.

Question

Is it true that any 6×6 copositive matrix with an all-ones diagonal belongs to $\mathcal{K}_6^{(r)}$ for some $r \in \mathbb{N}$?

4.2 Exactness of the conic approximations $\mathrm{LAS}_{\Delta_n}^{(r)}$

We begin with the characterization of the matrix sizes n for which the hierarchy of cones $\mathrm{LAS}_{\Delta_n}^{(r)}$ is exact.

Theorem 14 (Laurent, Vargas [33])

We have $\mathrm{COP}_2 = \mathrm{LAS}_{\Delta_2}^{(3)}$*, and the inclusion* $\bigcup_{r\geq 0} \mathrm{LAS}_{\Delta_n}^{(r)} \subseteq \mathrm{COP}_n$ *is strict for any* $n \geq 3$.

Proof First, assume $M = \begin{pmatrix} a & c \\ c & b \end{pmatrix} \in \mathrm{COP}_2$, we show $M \in \mathrm{LAS}_{\Delta_2}^{(3)}$. Note that $a, b \geq 0$ and $c \geq -\sqrt{ab}$ (using the fact that $u^T Mu \geq 0$ with $u = (1,0)$, $(0,1)$, and $(\sqrt{b}, \sqrt{a})$). Then we can write $x^T Mx = (\sqrt{a}x_1 - \sqrt{b}x_2)^2 + 2(c + \sqrt{ab})x_1x_2$, which, modulo the ideal I_{Δ_2}, is equal to $(\sqrt{a}x_1 - \sqrt{b}x_2)^2(x_1 + x_2) + 2(c + \sqrt{ab})(x_2^2x_1 + x_1^2x_2)$, thus showing $M \in \mathrm{LAS}_{\Delta_2}^{(3)}$.

For $n = 3$, the matrix

$$M := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{45}$$

is copositive (since nonnegative), but does not belong to any of the cones $\mathrm{LAS}^{(r)}_{\Delta_3}$. To see this, assume, by way of contradiction, that $M \in \mathrm{LAS}^{(r)}_{\Delta_3}$ for some $r \in \mathbb{N}$. Then the polynomial $x^T M x = 2x_1x_2$ has a decomposition as in (19). However, we showed in the related example (end of Section 2.2) that such a decomposition does not exist. □

Some differences between the cones $\mathrm{LAS}^{(r)}_{\Delta_n}$ and $\mathcal{K}^{(r)}_n$

By Theorems 7 and 14, we have $\bigcup_r \mathrm{LAS}^{(r)}_{\Delta_n} \subseteq \bigcup_r \mathcal{K}^{(r)}_n$, with equality if $n = 2$. This inclusion is strict for any $n \geq 3$. Indeed, the matrix M in (45) is an example of a matrix that does not belong to any cone $\mathrm{LAS}^{(r)}_{\Delta_3}$ while it belongs to the cone $\mathcal{K}^{(0)}_3$ (because M is copositive and $\mathrm{COP}_3 = \mathcal{K}^{(0)}_3$, in view of Theorem 9).

Another example is the Horn Matrix H. As observed in (42), $H \in \mathcal{K}^{(1)}_5$ and it can be shown that $H \notin \mathrm{LAS}^{(r)}_{\Delta_5}$ for any r (see [33]). The proof exploits the structure of the (infinitely many) zeros of the form $x^T H x$ in Δ_5.

We just saw two examples of copositive matrices that do not belong to any cone $\mathrm{LAS}^{(r)}_{\Delta_n}$. In both cases, the structure of the *infinitely many* zeros plays a crucial role. We will now discuss some tools that can be used to show membership in some cone $\mathrm{LAS}^{(r)}_{\Delta_n}$ in the case when the quadratic form $x^T M x$ has *finitely many* zeros in Δ_n.

First, recall that, if a matrix M lies in the interior of the cone COP_n, then it belongs to some cone $\mathrm{LAS}^{(r)}_{\Delta_n}$ (see relation (34)). Therefore we now assume that M lies on the boundary of COP_n, denoted by $\partial\mathrm{COP}_n$. The next result shows that, if the quadratic form $x^T M x$ has finitely many zeros in Δ_n and if these zeros satisfy an additional technical condition, then M belongs to some cone $\mathrm{LAS}^{(r)}_{\Delta_n}$.

Theorem 15 (Laurent, Vargas [33])

Let $M \in \partial\mathrm{COP}_n$. Assume that the quadratic form $p_M := x^T M x$ has finitely many zeros in Δ_n and that, for every zero u of p_M in Δ_n, we have $(Mu)_i > 0$ for all $i \in [n] \setminus \mathrm{Supp}(u)$. Then, $M \in \bigcup_{r\geq 0} \mathrm{LAS}^{(r)}_{\Delta_n}$ and, moreover, $DMD \in \bigcup_{r\geq 0} \mathrm{LAS}^{(r)}_{\Delta_n}$ for all $D \in \mathcal{D}^n_{++}$.

The proof of Theorem 15 relies on following an optimization approach, which enables using the result from Theorem 5 about finite convergence of the Lasserre hierarchy. For this, consider the following standard quadratic program

$$\min\{x^T M x : x \in \Delta_n\}. \tag{46}$$

First, since $M \in \partial\mathrm{COP}_n$ the optimal value of problem (46) is zero and thus a vector $u \in \Delta_n$ is a global minimizer of problem (46) if and only if u is a zero of $x^T M x$.

Next, observe that, as a direct consequence of the definitions, showing membership in some cone $\mathrm{LAS}^{(r)}_{\Delta_n}$ amounts to showing finite convergence of the Lasserre hierarchy for problem (46).

Linking membership in $\mathrm{LAS}^{(r)}_{\Delta_n}$ to finite convergence of Lasserre hierarchy

Assume $M \in \partial\mathrm{COP}_n$. Then, $M \in \bigcup_{r\geq 0} \mathrm{LAS}^{(r)}_{\Delta_n}$ if and only if the Lasserre hierarchy (17) applied to problem (46) (for matrix M) has finite convergence.

Now, in order to study the finite convergence of the Lasserre hierarchy for problem (46), we will apply the result of Theorem 5 to the special case of problem (46). First, we observe that the Archimedean condition holds. For this, note that, for any $i \in [n]$, we have

$$1 - x_i = 1 - \sum_{k=1}^{n} x_k + \sum_{k\in[n]\setminus\{i\}} x_k, \quad 1 - x_i^2 = \frac{(1+x_i)^2}{2}(1-x_i) + \frac{(1-x_i)^2}{2}(1+x_i).$$

This implies $n - \sum_{i=1}^n x_i^2 \in \mathcal{M}(x_1, \ldots, x_n) + I_{\Delta_n}$, thus showing that the Archimedean condition holds.

In [33] it is shown that the strict complementarity condition (SCC) holds at a global minimizer u of problem (46) if and only if $(Mu)_i > 0$ for all $i \in [n] \setminus \mathrm{Supp}(u)$. It is also shown there that, if problem (46) has finitely many minimizers, then the second order sufficiency condition (SOSC) holds at each of them. These two facts (roughly) allow us to apply the result from Theorem 5 and to conclude the proof of Theorem 15. The exact technical details are summarized in the next result.

Proposition 2 ([33])

Let $M \in \partial\mathrm{COP}_n$ and $D \in \mathcal{D}^n_{++}$. Assume the form $x^T M x$ has finitely many zeros in Δ_n. Then the following holds.

(i) (SCC) holds at a minimizer u of problem (46) (for M) if $(Mu)_i > 0$ for all $i \in [n] \setminus \mathrm{Supp}(u)$.
(ii) (SOSC) holds at every minimizer of problem (46) (for M).

In addition, if the optimality conditions (SCC) and (SOSC) hold at every minimizer of problem (46) for the matrix M, then they also hold for every minimizer of problem (46) for the matrix DMD.

The following example shows a copositive matrix M for which the form $x^T M x$ has a unique zero in Δ_n; however M does not belong to $\bigcup_{r\geq 0} \mathcal{K}^{(r)}_n$, and thus it also does not belong to $\bigcup_{r\geq 0} \mathrm{LAS}^{(r)}_{\Delta_n}$ (in view of relation (35)). Hence, the condition on the support of the zeros in Theorem 15 cannot be omitted.

A copositive matrix with a unique zero, that does not belong to any cone $\mathcal{K}^{(r)}_n$

Let M_1 be a matrix lying in $\mathrm{int}(\mathrm{COP}_n) \setminus \mathcal{K}_n^{(0)}$. Such a matrix exists for any $n \geq 5$. As an example for M_1, one may take the Horn matrix H in (41), in which we replace all entries 1 by t, where t is a given scalar such that $1 < t < \sqrt{5} - 1$ (see [32]). By Theorem 1 we have

$$M := \left(\begin{array}{c|cc} M_1 & \multicolumn{2}{c}{0} \\ \hline 0 & 1 & -1 \\ & -1 & 1 \end{array}\right) \in \mathrm{COP}_{n+2} \setminus \bigcup_{r\geq 0} \mathcal{K}_{n+2}^{(r)}. \tag{47}$$

Now we prove that the quadratic form $x^T M x$ has a unique zero in the simplex. For this, let $x \in \Delta_{n+2}$ such that $x^T M x = 0$. As M_1 is strictly copositive and $y := (x_1, \ldots, x_n)$ is a zero of the quadratic form $y^T M_1 y$ it follows that $x_1 = \ldots = x_n = 0$. Hence (x_{n+1}, x_{n+2}) is a zero of the quadratic form $x_{n+1}^2 - 2x_{n+1}x_{n+2} + x_{n+2}^2$ in the simplex Δ_2 and thus $x_{n+1} = x_{n+2} = 1/2$. This shows that the only zero of the quadratic form $x^T M x$ in the simplex Δ_n is $x = (0, 0, \ldots, 0, \frac{1}{2}, \frac{1}{2})$, as desired.

4.3 The cone of 5 × 5 copositive matrices

In this section we return to the cone COP_5, more specifically, to the result in Theorem 13 claiming that $\mathrm{COP}_5 = \bigcup_r \mathcal{K}_5^{(r)}$. Here we give a sketch of proof for (some of) the main arguments that are used to show this result.

As a starting point, observe that it suffices to show that every 5×5 copositive matrix that lies on an extreme ray of COP_5 (for short, call such a matrix *extreme*) belongs to some cone $\mathcal{K}_5^{(r)}$. Then, as a crucial ingredient, we use the fact that the extreme matrices in COP_5 have been fully characterized by Hildebrand [26]. Note that, if M is an extreme matrix in COP_n, then the same holds for all its positive diagonal scalings DMD where $D \in \mathcal{D}_{++}^n$. Hildebrand [26] introduced the following matrices

$$T(\psi) = \begin{pmatrix} 1 & -\cos\psi_4 & \cos(\psi_4+\psi_5) & \cos(\psi_2+\psi_3) & -\cos\psi_3 \\ -\cos\psi_4 & 1 & -\cos\psi_5 & \cos(\psi_5+\psi_1) & \cos(\psi_3+\psi_4) \\ \cos(\psi_4+\psi_5) & -\cos\psi_5 & 1 & -\cos\psi_1 & \cos(\psi_1+\psi_2) \\ \cos(\psi_2+\psi_3) & \cos(\psi_5+\psi_1) & -\cos\psi_1 & 1 & -\cos\psi_2 \\ -\cos\psi_3 & \cos(\psi_3+\psi_4) & \cos(\psi_1+\psi_2) & -\cos\psi_2 & 1 \end{pmatrix},$$

where $\psi \in \mathbb{R}^5$, which he used to prove the following theorem.

Theorem 16 (Hildebrand [26])

The extreme matrices M in COP_5 *can be divided into the following three categories:*

(i) $M \in \mathcal{K}_n^{(0)}$,

(ii) *M is (up to row/column permutation) a positive diagonal scaling of the Horn matrix H,*

(iii) M is (up to row/column permutation) a positive diagonal scaling of a matrix $T(\psi)$ for some $\psi \in \Psi$, where the set Ψ is defined by

$$\Psi = \Big\{\psi \in \mathbb{R}^5 : \sum_{i=1}^{5} \psi_i < \pi,\ \psi_i > 0 \text{ for } i \in [5]\Big\}. \tag{48}$$

As a direct consequence, in order to show equality $\mathrm{COP}_5 = \bigcup_{r\geq 0} \mathcal{K}_n^{(r)}$, it suffices to show that every positive diagonal scaling of the matrices $T(\psi)$ ($\psi \in \Psi$) and H lies in some cone $\mathcal{K}_n^{(r)}$. It turns out that a different proof strategy is needed for the class of matrices $T(\psi)$ and for the Horn matrix H. The main reason lies in the fact that the form $x^T Mx$ has finitely many zeros in the simplex when $M = T(\psi)$, but infinitely many zeros when $M = H$. We will next discuss these two cases separately.

Proof strategy for the matrices $T(\psi)$

Here we show that any positive diagonal scaling of a matrix $T(\psi)$ (with $\psi \in \Psi$) belongs to some cone $\mathcal{K}_5^{(r)}$. We, in fact, show a stronger result, namely membership in some cone $\mathrm{LAS}_{\Delta_n}^{(r)}$. For this, the strategy is to apply the result of Theorem 15 to the matrix $T(\psi)$. So we need to verify that the required conditions on the zeros of $x^T T(\psi)x$ are satisfied. First, we recall a characterization of the (finitely many) zeros of $x^T T(\psi)x$, which follows from results in [26].

Lemma 3 ([26])

For any $\psi \in \Psi$, the zeros of the quadratic form $x^T T(\psi)x$ in the simplex Δ_5 are the vectors $v_i = \frac{u_i}{\|u_i\|_1}$ for $i \in [5]$, where the u_i's are defined by

$$u_1 = \begin{pmatrix} \sin\psi_5 \\ \sin(\psi_4+\psi_5) \\ \sin\psi_4 \\ 0 \\ 0 \end{pmatrix},\ u_2 = \begin{pmatrix} \sin(\psi_3+\psi_4) \\ \sin\psi_3 \\ 0 \\ 0 \\ \sin\psi_4 \end{pmatrix},\ u_3 = \begin{pmatrix} 0 \\ \sin\psi_1 \\ \sin(\psi_1+\psi_5) \\ \sin\psi_5 \\ 0 \end{pmatrix},\ u_4 = \begin{pmatrix} 0 \\ 0 \\ \sin\psi_2 \\ \sin(\psi_1+\psi_2) \\ \sin\psi_1 \end{pmatrix},\ u_5 = \begin{pmatrix} \sin\psi_2 \\ 0 \\ 0 \\ \sin\psi_3 \\ \sin(\psi_2+\psi_3) \end{pmatrix}.$$

Then, it is straightforward to check that the conditions in Theorem 15 are satisfied and so we obtain the following result for the extreme matrices of type (iii) in Theorem 16.

Theorem 17 (Laurent, Vargas [33])

We have $DT(\psi)D \in \bigcup_{r\geq 0} \mathrm{LAS}_{\Delta_n}^{(r)}$ for all $D \in \mathcal{D}_{++}^5$ and $\psi \in \Psi$.

Proof strategy for the Horn matrix H

As already mentioned, the above strategy cannot be applied to the positive diagonal scalings of H (extreme matrices of type (ii) in Theorem 16), because the form $x^T Hx$ has infinitely many zeros in Δ_5; e.g., any $x = (\frac{1}{2}, 0, \frac{t}{2}, \frac{1-t}{2}, 0)$ with $t \in [0,1]$ is a zero. In fact, as mentioned earlier, the Horn matrix H does not belong to any of the cones

$\mathrm{LAS}^{(r)}_{\Delta_n}$ (see [33]). Then, another strategy should be applied for showing that all its positive diagonal scalings belong to some cone $\mathcal{K}^{(r)}_5$.

The starting point is to use the fact that $\bigcup_r \mathcal{K}^{(r)}_n = \bigcup_r \mathrm{LAS}^{(r)}_{\mathbb{S}^{n-1}}$ (recall Theorem 7) and to change variables. This enables us to rephrase the question of whether all positive diagonal scalings of H belong to $\bigcup_r \mathcal{K}^{(r)}_5$ as the question of deciding whether, for all positive scalars $d_1, \ldots, d_5$, the form $(x^{\circ 2})^T H x^{\circ 2}$ can be written as a sum of squares modulo the ideal generated by $\sum_{i=1}^5 d_i x_i^2 - 1$. This latter question was recently answered in the affirmative by Schweighofer and Vargas [52].

Theorem 18 (Schweighofer, Vargas [52])

Let $d_1, d_2, \ldots, d_5 > 0$ be positive real numbers. Then we have

$$(x^{\circ 2})^T H x^{\circ 2} = \sigma + q\Big(1 - \sum_{i=1}^{5} d_i x_i^2\Big) \text{ for some } \sigma \in \Sigma \text{ and } q \in \mathbb{R}[x].$$

Therefore, $DHD \in \bigcup_r \mathcal{K}^{(r)}_5$ for all $D \in \mathcal{D}^5_{++}$.

The proof of this theorem uses the theory of pure states in real algebraic geometry (as described in [8]), combined with a characterization of the diagonal scalings of the Horn matrix that belong to the cone $\mathcal{K}^{(1)}_n$ (given in [32]). The technical details go beyond the scope of this chapter, so we refer to [52] for details.

5 The stability number of a graph $\alpha(G)$

In this section, we investigate a class of copositive matrices that arise naturally from graphs. Consider a graph $G = (V = [n], E)$, where $V = [n]$ is the set of vertices and E is the set of edges, consisting of the pairs of distinct vertices that are adjacent in G. A set $S \subseteq V$ is called *stable* (or *independent*) if it does not contain any edge of G. Then, the *stability number* of G, denoted by $\alpha(G)$, is defined as the maximum cardinality of a stable set in G. Computing $\alpha(G)$ is a well-known NP-hard problem (see [27]), with many applications, e.g., in operations research, social networks analysis, and chemistry. There is a vast literature on this problem, dealing among other things with how to define linear and/or semidefinite approximations for $\alpha(G)$ (see, e.g., [13, 29, 55] and further references therein).

Lasserre hierarchy for $\alpha(G)$ via polynomial optimization on the binary cube

The stability number of $G = ([n], E)$ can be formulated as a polynomial optimization problem on the binary cube $\{0, 1\}^n$:

$$\alpha(G) = \max\Big\{\sum_{i\in V} x_i : \; x_i x_j = 0 \text{ for } \{i,j\} \in E,\; x_i^2 - x_i = 0 \text{ for } i \in V\Big\}. \tag{49}$$

We can consider the Lasserre hierarchy (17) for problem (49) and obtain the following bounds

$$\mathrm{las}^{(r)}(G) := \min \Big\{ \lambda : \lambda - \sum_{i \in V} x_i = \sigma + \sum_{\{i,j\} \in E} p_{ij} x_i x_j + \sum_{i \in V} q_i (x_i^2 - x_i) \tag{50}$$

$$\text{for some } \sigma \in \Sigma_{2r} \text{ and } p_{ij}, q_i \in \mathbb{R}[x]_{2r-2} \Big\}. \tag{51}$$

Clearly, we have $\alpha(G) \leq \mathrm{las}^{(r)}(G)$. Moreover, the bound is exact at order $r = \alpha(G)$, that is, $\alpha(G) = \mathrm{las}^{(\alpha(G))}(G)$ (see [29]). The proof is not difficult and exploits the fact that in the definition of these parameters one works modulo the ideal generated by the polynomials $x_i^2 - x_i$ $(i \in V)$ and the edge monomials $x_i x_j$ $(\{i, j\} \in E)$. At order $r = 1$, the bound $\mathrm{las}^{(1)}(G)$ coincides with the parameter $\vartheta(G)$ introduced in 1979 by Lovász in his seminal paper [34].

In this section we focus on the hierarchies of approximations that naturally arise when considering the following copositive reformulation for $\alpha(G)$, given by de Klerk and Pasechnik [13]:

$$\alpha(G) = \min\{t : \; t(A_G + I) - J \in \mathrm{COP}_n\}. \tag{52}$$

Here, A_G, I, and J are, respectively, the adjacency matrix of G (whose entries are all 0 except 1 at the positions corresponding to the edges of G), the identity, and the all-ones matrix. As a consequence, it follows from (52) that the following *graph matrix*

$$M_G := \alpha(G)(I + A_G) - J \tag{53}$$

belongs to COP_n. The copositive reformulation (52) for $\alpha(G)$ can be seen as an application of the following quadratic formulation by Motzkin and Straus [38]:

$$\frac{1}{\alpha(G)} = \min\{x^T(I + A_G)x : x \in \Delta_n\}.$$

The Horn matrix coincides with the graph matrix of the graph C_5.

When $G = C_5$ is the 5-cycle, its adjacency matrix A_G is given by

$$A_{C_5} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

As $\alpha(C_5) = 2$, it follows that the graph matrix $M_{C_5} = 2(I + A_{C_5}) - J$ of C_5 coincides with the Horn matrix H.

Based on the formulation (52), de Klerk and Pasechnik [13] proposed two hierarchies $\zeta^{(r)}(G)$ and $\vartheta^{(r)}(G)$ of upper bounds for $\alpha(G)$, that are obtained by replacing in (52) the cone COP_n by its subcones $\mathcal{C}_n^{(r)}$ and $\mathcal{K}_n^{(r)}$, respectively. In this section, we present several known results about these two hierarchies and related results for the graph matrices M_G. One of the central questions is whether the hierarchy $\vartheta^{(r)}(G)$ converges to $\alpha(G)$ in finitely many steps or, equivalently, whether the matrix M_G belongs to $\bigcup_r \mathcal{K}_n^{(r)}$, and what can be said about the minimum number of steps where finite convergence takes place.

5.1 The hierarchy $\zeta^{(r)}(G)$

As mentioned above, for an integer $r \geq 0$, the parameter $\zeta^{(r)}(G)$ is defined as

$$\zeta^{(r)}(G) := \min\{t : \ t(A_G + I) - J \in \mathcal{C}_n^{(r)}\}. \tag{54}$$

Since $\mathrm{int}(\mathrm{COP}_n) \subseteq \bigcup_{r\geq 0} \mathcal{C}_n^{(r)}$, it follows directly that the parameters $\zeta^{(r)}(G)$ converge asymptotically to $\alpha(G)$ as $r \to \infty$. Note that, if $G = K_n$ is a complete graph, then $\alpha(G) = 1$ and the matrix $I + A_G - J$ is the zero matrix, thus belonging trivially to the cone $\mathcal{C}_n^{(0)}$, so that $1 = \alpha(K_n) = \zeta^{(0)}(K_n)$. However, finite convergence does not hold if G is not a complete graph.

Theorem 19 (de Klerk, Pasechnik [13])

Assume G is not a complete graph. Then, we have $\zeta^{(r)}(G) > \alpha(G)$ for all $r \in \mathbb{N}$.

By the definition of the cone $\mathcal{C}_n^{(r)}$, the parameter $\zeta^{(r)}(G)$ can be formulated as a linear program, asking for the smallest scalar t for which all the coefficients of the polynomial $(\sum_{i=1}^n x_i)^r \ x^T(t(I + A_G) - J)x$ are nonnegative. The parameter $\zeta^{(r)}(G)$ is very well understood. Indeed, Peña, Vera and Zuluaga [43] give a closed-form expression for it in terms of $\alpha(G)$.

Theorem 20 (Peña, Vera, Zuluaga [43])

Write $r+2 = u\alpha(G)+v$, where u, v are nonnegative integers such that $v \leq \alpha(G)-1$. Then we have

$$\zeta^{(r)}(G) = \frac{\binom{r+2}{2}}{\binom{u}{2}\alpha(G) + uv},$$

where we set $\zeta^{(r)}(G) = \infty$ if $r \leq \alpha(G) - 2$ (since then the denominator in the above formula is equal to 0).

So the above result shows that the bound $\zeta^{(r)}$ is useless for $r \leq \alpha(G) - 2$. Another consequence is that after $r = \alpha(G)^2 - 1$ steps we find $\alpha(G)$ up to rounding. (See also [13] where this result is shown for $r = \alpha(G)^2$).

Corollary 1 ([43])

We have $\lfloor \zeta^{(r)}(G) \rfloor = \alpha(G)$ if and only if $r \geq \alpha(G)^2 - 1$.

5.2 The hierarchy $\vartheta^{(r)}(G)$

We now consider the parameter $\vartheta^{(r)}(G)$, for $r \in \mathbb{N}$, defined as follows in [13]:

$$\vartheta^{(r)}(G) := \min\{t : t(A_G + I) - J \in \mathcal{K}_n^{(r)}\}. \tag{55}$$

Since $C_n^{(r)} \subseteq \mathcal{K}_n^{(r)} \subseteq \mathrm{COP}_n$ we have $\alpha(G) \le \vartheta^{(r)}(G) \le \zeta^{(r)}(G)$ for any $r \ge 0$, and thus the parameters $\vartheta^{(r)}(G)$ converge asymptotically to $\alpha(G)$ as $r \to \infty$.

At order $r = 0$, while the parameter $\zeta^{(0)}(G) = \infty$ is useless, the parameter $\vartheta^{(0)}(G)$ provides a useful bound for $\alpha(G)$. Indeed, it is shown in [13] that $\vartheta^{(0)}(G)$ coincides with the variation $\vartheta'(G)$ of the Lovász theta number $\vartheta(G)$ (obtained by adding some nonnegativity constraints); so we have the inequalities $\alpha(G) \le \vartheta'(G) = \vartheta^{(0)}(G) \le \vartheta(G)$ (see [34, 51]). This connection in fact motivates the choice of the notation $\vartheta^{(r)}(G)$. For instance, if G is a perfect graph[1], then we have $\vartheta(G) = \vartheta^{(0)}(G) = \alpha(G)$ (see [20] for a broad exposition). We also have $\vartheta(C_5) = \vartheta^{(0)}(C_5)$ (note that C_5 is not a perfect graph since $\omega(C_5) = 2 < \chi(C_5) = 3$). But there exist graphs for which $\alpha(G) = \vartheta^{(0)}(G) < \vartheta(G)$ (see, e.g., [3]).

In Theorem 19 we saw that the bounds $\zeta^{(r)}(G)$ are never exact. This raises naturally the question of whether the (stronger) bonds $\vartheta^{(r)}(G)$ may be exact. Recall the definition of the graph matrix $M_G = \alpha(G)(A_G + I) - J$ in (53), and define the associated polynomial $p_G := (x^{\circ 2})^T M_G x^{\circ 2}$. Then, for any $r \in \mathbb{N}$, we have

$$\vartheta^{(r)}(G) = \alpha(G) \iff M_G \in \mathcal{K}_n^{(r)} \iff \Big(\sum_{i=1}^n x_i^2\Big)^r p_G \in \Sigma.$$

As M_G is copositive the polynomial p_G is globally nonnegative. The point however is that p_G has zeros in $\mathbb{R}^n \setminus \{0\}$. In particular, every stable set $S \subseteq V$ of cardinality $\alpha(G)$ provides a zero $x = \chi^S$. Thus the question of whether p_G admits a positivity certificate of the form $(\sum_{I=1}^n x_i^2)^r p_G \in \Sigma$ for some $r \in \mathbb{N}$ (as in (9)) is nontrivial. In [13] it was in fact conjectured that such a certificate exists at order $r = \alpha(G) - 1$; in other words, that the parameter $\vartheta^{(r)}(G)$ is exact at order $r = \alpha(G) - 1$.

Conjecture 1 (de Klerk and Pasechnik [13])
For any graph G, we have $\vartheta^{(\alpha(G)-1)}(G) = \alpha(G)$, or, equivalently, we have $M_G \in \mathcal{K}_n^{(\alpha(G)-1)}$.

Comparison of the parameters $\vartheta^{(r)}(G)$ and $\mathrm{las}^{(r)}(G)$

At the beginning of Section 5 we introduced the parameters $\mathrm{las}^{(r)}(G)$. In [21] it is shown that, for any integer $r \ge 1$, a slight strengthening of the parameter

[1] A graph G is called *perfect* if its clique number $\omega(G)$ coincides with its chromatic number $\chi(G)$, and the same holds for any induced subgraph G' of G. Here $\omega(G)$ denotes the maximum cardinality of a clique (a set of pairwise adjacent vertices) in G and $\chi(G)$ is the minimum number of colors that are needed to color the vertices of G in such a way that adjacent vertices receive distinct colors. An induced subgraph G' of G is any subgraph of G of the form $G' = G[U]$, obtained by selecting a subset $U \subseteq V$ and keeping only the edges of G that are contained in U.

$\mathrm{las}^{(r)}(G)$ (obtained by adding some nonnegativity constraints) is at least as good as the parameter $\vartheta^{(r-1)}(G)$. The bounds $\mathrm{las}^{(r)}(G)$ are known to converge to $\alpha(G)$ in $\alpha(G)$ steps, i.e., $\mathrm{las}^{(\alpha(G))}(G) = \alpha(G)$. Thus Conjecture 1 asks whether a similar property holds for the parameters $\vartheta^{(r)}(G)$. While the finite convergence property for the Lasserre-type bounds is relatively easy to prove (by exploiting the fact that one works modulo the ideal generated by $x_i^2 - x_i$ for $i \in V$ and $x_i x_j$ for $\{i, j\} \in E$)), proving Conjecture 1 seems much more challenging.

Conjecture 1 is known to hold for some graph classes. For instance, we saw above that it holds for perfect graphs (with $r = 0$), but it also holds for odd cycles and their complements – that are not perfect (with $r = 1$, see [13]). In [21] Conjecture 1 was shown to hold for all graphs G with $\alpha(G) \leq 8$ (see also [43] for the case $\alpha(G) \leq 6$). In fact, a stronger result is shown there: the proof relies on a technical construction of matrices that permit to certify membership of M_G in the cones $\mathcal{Q}_n^{(r)}$ (and thus in the cones $\mathcal{K}_n^{(r)}$).

Theorem 21 (Gvozdenović, Laurent [21])

Let G be a graph with $\alpha(G) \leq 8$. Then we have $\vartheta^{(\alpha(G)-1)}(G) = \alpha(G)$, or, equivalently, $M_G \in \mathcal{K}_n^{(\alpha(G)-1)}$.

Whether Conjecture 1 holds in general is still an open problem. However, a weaker form of it has been recently settled; namely finite convergence of the hierarchy $\vartheta^{(r)}(G)$ to $\alpha(G)$, or, equivalently, membership of the graph matrices M_G in $\bigcup_r \mathcal{K}_n^{(r)}$.

Theorem 22 (Schweighofer, Vargas [52])

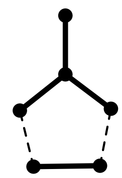
For any graph G, we have $\vartheta^{(r)}(G) = \alpha(G)$ for some $r \in \mathbb{N}$. Equivalently, we have $M_G \in \bigcup_r \mathcal{K}_n^{(r)}$.

In what follows we discuss some of the ingredients that are used for the proof of this result. Here too, we will use the fact that $\bigcup_r \mathrm{LAS}_{\Delta_n}^{(r)} \subseteq \bigcup_r \mathcal{K}_n^{(r)} = \bigcup_r \mathrm{LAS}_{\mathbb{S}^{n-1}}^{(r)}$ (recall Theorem 7) and so we we will consider the quadratic form $x^T M_G x$ instead of the quartic form $p_G = (x^{\circ 2})^T M_G x^{\circ 2}$. Whether the quadratic form $x^T M_G x$ has finitely many zeros in the simplex plays an important role. We will first discuss the case when there are finitely many zeros, in which case one can show a stronger result, namely membership of M_G in $\bigcup_r \mathrm{LAS}_{\Delta_n}^{(r)}$ (see Theorem 24 below).

As we will see in Corollary 2 below, whether the number of zeros of $x^T M_G x$ in Δ_n is finite is directly related to the notion of critical edges in the graph G. We first introduce this graph notion.

Critical edges

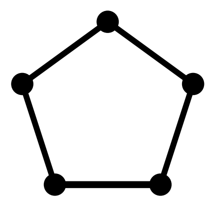
Let $G = (V, E)$ be a graph. The edge $e \in E$ is *critical* is $\alpha(G \setminus e) = \alpha(G) + 1$. Here $G \setminus e$ denotes the graph $(V, E \setminus \{e\})$.



For example, for the above graph, the two dashed edges are its critical edges.

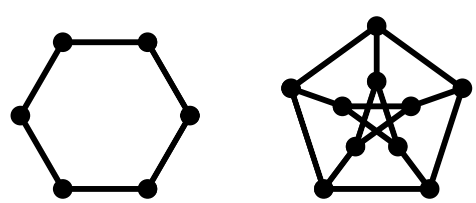
Critical graphs

We say that G is *critical* if all its edges are critical. For example, odd cycles are critical graphs. The next figure shows the 5-cycle C_5.



Acritical graphs

We say that G is *acritical* if it does not have critical edges. Every even cycle is acritical, as well as the Petersen graph. The next figure shows the 6-cycle C_6 and the Petersen graph.



We now explain the role played by the critical edges in the description of the zeros of the form $x^T M_G x$ in the simplex Δ_n. First, note that, if S is a stable set of size $\alpha(G)$, then $x = \chi^S/|S|$ is a zero. However, in general, there are more zeros. A characterization of the zeros was given in [31] (see also [19]).

Theorem 23 ([31])

Let $x \in \Delta_n$ with support $S := \{i \in V : x_i > 0\}$ and let $V_1, V_2, \dots, V_k$ denote the connected components of $G[S]$, the subgraph of G induced by the support S of x. Then x is a zero of the form $x^T M_G x$ if and only if $k = \alpha(G)$ and, for all $h \in [k]$, V_h is a clique of G and $\sum_{i \in V_h} x_i = \frac{1}{\alpha(G)}$. In addition, the edges that are contained in S are critical edges of G.

In particular, we can characterize the graphs G for which the form $x^T M_G x$ has finitely many zeros in Δ_n.

Corollary 2 ([31])

Let G be a graph. The form $x^T M_G x$ has finitely many zeros in Δ_n if and only if G is acritical (i.e., G has no critical edge). In that case, the zeros are the vectors of the form $\chi^S/|S|$, where S is a stable set of size $\alpha(G)$.

Zeros of the form $x^T M_G x$ for the cycles C_4 and C_5

The 4-cycle C_4 has vertex set $\{1,2,3,4\}$ and edges $\{1,2\}$, $\{2,3\}$, $\{3,4\}$, and $\{4,1\}$. It has stability number $\alpha(C_4) = 2$, it is acritical, and its maximum stable sets are the sets $\{1,3\}$ and $\{2,4\}$. Then, in view of Corollary 2, the only zeros of the form $x^T M_{C_4} x$ in Δ_4 are $(\frac{1}{2}, 0, \frac{1}{2}, 0)$ and $(0, \frac{1}{2}, 0, \frac{1}{2})$.

The 5-cycle C_5 has vertex set $\{1,2,3,4,5\}$ and edges $\{1,2\}$, $\{2,3\}$, $\{3,4\}$, $\{4,5\}$, and $\{5,1\}$. It has stability number $\alpha(C_5) = 2$ and it is critical. Then, in view of Theorem 23, the form $x^T M_{C_5} x$ has infinitely many zeros in Δ_5. For example, for any $t \in (0,1)$, the point $x_t = (\frac{1}{2}, 0, \frac{t}{2}, \frac{1-t}{2}, 0)$ is a zero supported in the two cliques $\{1\}$ and $\{3,4\}$ (indeed a critical edge). It can be checked that (up to symmetry) all zeros take the shape of x_t for $t \in [0,1]$.

When G is an acritical graph one can show that its graph matrix M_G belongs to one of the cones $\mathrm{LAS}^{(r)}_{\Delta_n}$, thus a stronger result than the result from Theorem 22.

Theorem 24 (Laurent, Vargas [31])

Let G be an acritical graph. Then we have $M_G \in \bigcup_{r \geq 0} \mathrm{LAS}^{(r)}_{\Delta_n}$.

As $\mathrm{LAS}^{(r)}_{\Delta_n} \subseteq \mathcal{K}^{(r)}_n$ for any $r \in \mathbb{N}$, this result implies finite convergence of the hierarchy of bounds $\vartheta^{(r)}(G)$ to $\alpha(G)$ for the class of acritical graphs.

The proof of Theorem 24 relies on applying Theorem 5. By assumption, G is acritical, and thus the quadratic form $x^T M_G x$ has finitely many zeros in Δ_n, as described in Corollary 2. Now it suffices to verify that the zeros satisfy the conditions of Theorem 5. We next give the (easy) details for the sake of concreteness.

Lemma 4 ([31])

Let G be an acritical graph and let S be a stable set of size $\alpha(G)$. Then, for $x = \chi^S/\alpha(G)$, we have $(M_G x)_i > 0$ for $i \notin S$.

Proof For a vertex $i \in V \setminus S$, let $N_S(i)$ denote the number of neighbours of i in S. We have $N_S(i) \geq 1$ because $S \cup \{i\}$ is not stable, as S is a stable set of size $\alpha(G)$. Since G is acritical we must have $N_S(i) \geq 2$. Indeed, if $N_S(i) = 1$ and $j \in S$ is the only neighbour of i in S, then $\{i, j\}$ is a critical edge, contradicting the assumption on G. Now we compute $(M_G x)_i$:

$$\begin{aligned}(M_G x)_i &= \frac{1}{\alpha(G)}((\alpha(G) - 1)N_S(i) - (\alpha(G) - N_S(i)))\\ &= \frac{1}{\alpha(G)}(\alpha(G)N_S(i) - \alpha(G)) > 0,\end{aligned}$$

where the last inequality holds as $N_S(i) \geq 2$. □

The above strategy does not extend for general graphs (having some critical edges) and also the result of Theorem 24 does not extend. For example, if $G = C_5$ is the 5-cycle (whose edges are all critical), then M_G is the Horn matrix that does not belong to any of the cones $\text{LAS}^{(r)}_{\Delta_n}$ (as we saw in Section 4.2). Hence another strategy is needed to show membership of M_G in $\bigcup_r \mathcal{K}^{(r)}_n$ for general graphs. We now sketch some of the key ingredients that are used to show this result.

Some key ingredients for the proof for Theorem 22

For studying Conjecture 1 and, in general, the membership of the graph matrices M_G in the cones $\mathcal{K}^{(r)}_n$, it turns out that the graph notion of isolated nodes plays a crucial role.

A node i of a graph G is said to be an *isolated node* of G if i is not adjacent to any other node of G. Given a graph $G = (V, E)$ and a new node $i_0 \notin V$, the graph $G \oplus i_0$ is the graph $(V \cup \{i_0\}, E)$ obtained by adding i_0 as an isolated node to G. The following result makes the link to Conjecture 1 clear.

Theorem 25 (Gvozdenović, Laurent [21])

Assume that, for any graph $G = ([n], E)$ and $r \in \mathbb{N}$, we have

$$M_G \in \mathcal{K}^{(r)}_n \Longrightarrow M_{G\oplus i_0} \in \mathcal{K}^{(r)}_{n+1}. \tag{56}$$

Then Conjecture 1 holds.

Moreover, it was conjectured in [21] that (56) holds for each $r \in \mathbb{N}$ (which, if true, would thus imply Conjecture 1). However, this conjecture was disproved in [31].

Adding an isolated node may not preserve membership in $\mathcal{K}^{(r)}$

Consider the 5-cycle C_5, whose graph matrix coincides with the Hall matrix: $M_{C_5} = H$. As we have seen earlier, $M_{C_5} \in \mathcal{K}_5^{(1)}$. In [31] it is shown that, if $G = C_5 \oplus i_1 \oplus \cdots \oplus i_8$ is the graph obtained by adding eight isolated nodes to the 5-cycle, then $M_G \in \mathcal{K}_{13}^{(1)}$, but, if we add one more isolated node i_0 to G (thus we add nine isolated nodes to C_5), then we have $M_{G\oplus i_0} \notin \mathcal{K}_{14}^{(1)}$.

Hence, one cannot rely on the result of Theorem 25 and a new strategy is needed for solving Conjecture 1. The following variation of Theorem 25 is shown in [32], which can serve as a basis for proving a weaker form of Conjecture 1, namely membership of M_G in $\bigcup_r \mathcal{K}_n^{(r)}$.

Theorem 26 (Laurent and Vargas [32])

The following two assertions are equivalent.

(i) *For any graph* $G = ([n], E)$*,* $M_G \in \bigcup_{r\geq 0} \mathcal{K}_n^{(r)}$ *implies* $M_{G\oplus i_0} \in \bigcup_{r\geq 0} \mathcal{K}_{n+1}^{(r)}$.

(ii) *For any graph* $G = ([n], E)$*, we have* $M_G \in \bigcup_{r\geq 0} \mathcal{K}_n^{(r)}$.

This result is used as a crucial ingredient in [52] for showing Theorem 22; namely, the authors of [52] show that Theorem 26 (i) holds. The starting point of their proof is to use the fact that $\bigcup_{r\geq 0} \mathcal{K}_n^{(r)} = \bigcup_{r\geq 0} \mathrm{LAS}_{\mathbb{S}^{n-1}}^{(r)}$ (by Theorem 7) and then to show that membership of the graph matrices in $\bigcup_{r\geq 0} \mathrm{LAS}_{\mathbb{S}^{n-1}}^{(r)}$ is preserved after adding isolated nodes. Recall that $p_G = (x^{\circ 2})^T M_G x^{\circ 2} = \sum_{i,j\in V} x_i^2 x_j^2 (M_G)_{ij}$.

Theorem 27 (Schweighofer and Vargas [52])

Let $G = ([n], E)$ *be a graph. Assume that* $p_G = \sigma_0 + q(\sum_{i=1}^n x_i^2 - 1)$ *for some* $\sigma_0 \in \Sigma$ *and* $q_0 \in \mathbb{R}[x_1, \ldots, x_n]$*. Then* $p_{G\oplus i_0} = \sigma_1 + q_1(x_{i_0}^2 + \sum_{i=1}^n x_i^2 - 1)$ *for some* $\sigma_1 \in \Sigma$ *and* $q_1 \in \mathbb{R}[x_{i_0}, x_1, \ldots, x_n]$.

Here too, the proof of this theorem uses the theory of pure states in real algebraic geometry (as described in [8]). The technical details are too involved and thus go beyond the scope of this chapter, we refer to [52] for the full details. As explained above, this theorem implies Theorem 22. The result (and proof) of Theorem 27, however, does not give any explicit bound on the degree of σ_1 in terms of the degree of σ_0. Hence one cannot infer any information on the degree of a representation of p_G in $\Sigma + I(\sum_{i=1}^n x_i^2 - 1)$. In other words, this result gives no information on the number of steps at which finite convergence of $\vartheta^{(r)}(G)$ to $\alpha(G)$ takes place.

Therefore, the status of Conjecture 1 remains widely open and its resolution likely requires new techniques. There is some evidence for its validity; for instance, Conjecture 1 holds for perfect graphs and for graphs G with $\alpha(G) \leq 8$ (Theorem 25), and any graph matrix M_G belongs to some cone $\mathcal{K}_n^{(r)}$ (Theorem 22). These facts also make the search for a possible counterexample a rather difficult task.

6 Concluding remarks

In this chapter we have discussed several hierarchies of conic inner approximations for the copositive cone COP_n, motivated by various sum-of-squares certificates for positive polynomials on $\mathbb{R}^n$, $\mathbb{R}^n_+$, the simplex Δ_n, and the unit sphere $\mathbb{S}^{n-1}$. The main players are Parrilo's cones $\mathcal{K}_n^{(r)}$, originally defined as the sets of matrices M for which the polynomial $(\sum_{i=1}^n x_i^2)^r (x^{\circ 2})^T M x^{\circ 2}$ is a sum of squares of polynomials, thus having a certificate "with denominator" (for positivity on $\mathbb{R}^n$). The question whether these cones cover the full copositive cone is completely settled: the answer is positive for $n \leq 5$ and negative for $n \geq 6$. The cones $\mathcal{K}_n^{(r)}$ also capture the class of copositive graph matrices, of the form $M_G = \alpha(G)(A_G + I) - J$ for some graph G. The challenge in settling these questions lies in the fact that, for any copositive matrix lying on the border of COP_n, the associated form has (nontrivial) zeros (and thus is not *strictly* positive), so that the classical positivity certificates do not suffice to claim membership in the conic approximations, and thus other techniques are needed.

A useful step is understanding the links to other certificates "without denominators" for positivity on the simplex or the sphere, which lead to the Lasserre-type cones $\mathrm{LAS}_{\Delta_n}^{(r)}$ and $\mathrm{LAS}_{\mathbb{S}^{n-1}}^{(r)}$. Roughly speaking, the simplex-based cones form a weaker hierarchy, while the sphere-based cones provide an equivalent formulation for Parrilo's cones (see Theorem 7 and relation (40) for the exact relationships). Membership in the simplex-based cones can be shown for some classes of copositive matrices, which thus implies membership in Parrilo's cones.

We recall Conjecture 1 that asks whether any graph matrix M_G belongs to the cone $\mathcal{K}_n^{(r)}$ of order $r = \alpha(G) - 1$, still widely open for graphs with $\alpha(G) \geq 9$. The resolution of Conjecture 1 would offer an interesting result that is relevant to the intersection of combinatorial optimization (about the computation of $\alpha(G)$), matrix copositivity (membership of a class of structured copositive matrices in one of Parrilo's approximation cones), and real algebraic geometry (a sum-of-squares representation result with an explicit degree bound for a polynomial with zeros).

Matrix copositivity revolves around the question of deciding whether a quadratic form is nonnegative on $\mathbb{R}^n_+$. This fits, more generally, within the study of copositive tensors, thus going from quadratic forms to forms with degree $d \geq 2$. There is a wide literature on copositive tensors; we refer, e.g., to [41, 46, 54] and further references therein. The relationships between the various types of positivity certificates discussed in this chapter for the case $d = 2$ extend to the case $d \geq 2$. (Note indeed that Theorems 6 and 8 hold for general homogeneous polynomials.) An interesting research direction may be to understand classes of structured symmetric tensors that are captured by some of the corresponding conic hierarchies.

References

1. E. Artin. Über die Zerlegung definiter Funktionen in Quadrate. *Abh. Math. Sem. Univ. Hamburg* 5: 100-115, 1927; In: Collected Papers (Lang, S., Tate, J. eds), Addison-Wesley (1965),

reprinted by Springer, pp. 273-288.

2. A. Ben-Tal, A. Nemirovski. *Lectures on Modern Convex Optimization - Analysis, Algorithms, and Engineering Applications, MPS-SIAM Series on Optimization, 2001.*
3. M.R. Best, A.E. Brouwer, F.J. MacWilliams, A.M. Odlyzko, N.J.A Sloane. Bounds for binary codes of length less than 25. *IEEE Transactions on Information Theory* 24(1), 81-93, 1978.
4. M. Bodirsky, M. Kummer, A. Thom. Spectrahedral Shadows and Completely Positive Maps on Real Closed Fields. arXiv:2206.06312, 2022.
5. I. Bomze, E. de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24:163–185, 2002.
6. I.M. Bomze, M. Dür, E. de Klerk, C. Roos, A.J. Quist, T. Terlaky. On copositive programming and standard quadratic optimization problems. *J. Global Optim.*, 18(4):301–320, 2000.
7. S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming, Ser. A*, 120:479–495, 2009.
8. S. Burgdorf, C. Scheiderer, M. Schweighofer. Pure states, nonnegative polynomials and sums of squares. *Comment. Math. Helv. 87, no. 1: 113–140, 2012.*
9. M. Castle, V. Powers, B. Reznick. A quantitative Pólya's Theorem with zeros, *J. Symbolic Comput.*, 44: 1285-1290, 2009.
10. M.D. Choi, T.Y Lam, B. Reznick. Sums of squares of real polynomials. *In: K-Theory and Algebraic Geometry: Connections with Quadratic Forms and Division Algebras* (Santa Barbara, CA, 1992). Proceedings of Symposia in Pure Mathematics, vol. 58: 103-126. American Mathematical Society, Providence, 1995.
11. E. de Klerk. *Aspects of Semidefinite Programming - Interior Point Algorithms and Selected Applications*. Kluwer, 2002.
12. E. de Klerk, M. Laurent, P. Parrilo. On the equivalence of algebraic approaches to the minimization of forms on the simplex. *Positive Polynomials in Control*, Number 312 in *LNCIS. Springer Verlag*, Germany, 2005.
13. E. de Klerk, D. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, 12:875–892, 2002.
14. P. Diananda. On non-negative forms in real variables some or all of which are non-negative. *Mathematical Proceedings of the Cambridge Philosophical Society*, 58(1): 17-25, 1962.
15. P.J.C. Dickinson, M. Dür, L. Gijben, R. Hildebrand. Scaling relationship between the copositive cone and Parrilo's first level approximation. *Optimization Letters*, 7(8):1669–1679. 2013.
16. I. Dukanovic, F. Rendl. Copositive programming motivated bounds on the clique and the chromatic number. *Mathematical Programming*, 121:249–268. 2010.
17. M. Dür. Copositive Programming – a Survey. In: Diehl, M., Glineur, F., Jarlebring, E., Michiels, W. (eds) *Recent Advances in Optimization and its Applications in Engineering*, pages 3–20. Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-642-12598-01
18. M.R. Garey, D.S. Johnson, L. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1:237–267, 1976.
19. L.E. Gibbons, D.W. Hearn, P.M. Pardalos and M.V. Ramana. Continuous characterizations of the maximum clique problem, *Math. Oper. Res.*, 22(3):754–768. 1997.
20. M. Grötschel, L. Lovász, A. Schrijver. Geometric algorithms and combinatorial optimization. Second edition. *Algorithms and Combinatorics*, 2. Springer-Verlag, Berlin, 1993.
21. N. Gvozdenović, M. Laurent. Semidefinite bounds for the stability number of a graph via sums of squares of polynomials. *Mathematical Programming*, 110:145–173, 2007.
22. M. Gvozdenović, M. Laurent. The operator Ψ for the chromatic number of a graph. *SIAM Journal on Optimization*, 19(2):572-591, 2008.
23. M. Hall, M. Newman. Copositive and completely positive quadratic forms. *Mathematical Proceedings of the Cambridge Philosophical Society*, 59(2), 329-339. 1963.
24. D. Hilbert. Über die Darstellung definiter Formen als Summe von Formenquadraten. *Math. Ann*. 32: 342-350, 1888. See: Ges. Abh. 2, Springer, Berlin (1933): 339-364, reprinted by Chelsea, New York,1981
25. D. Hilbert. Über ternäre definite Formen, *Acta. Math.* 17: 169-197,1893. See: Ges. Abh. 2: 345-366, Springer, Berlin (1933), reprinted by Chelsea, New York , 1981

26. R. Hildebrand R. The extreme rays of the 5 × 5 copositive cone. *Linear Algebra and its Applications*, 437(7):1538–1547, 2012.
27. R. Karp. Reducibility among combinatorial problems. *Plenum Press*, New York, 85–103, 1972.
28. J.B. Lasserre. Global optimization with polynomials and the problem of moments, *SIAM J. Optim.*, 11:796–817, 2001.
29. M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming, *Math. Oper. Res.*, 28(3): 470–496, 2003.
30. M. Laurent. Sums of squares, moment matrices and optimization over polynomials, in IMA Volumes in *Mathematics and its Applications*, M. Putinar and S. Sullivant (eds.), vol. 149:157–270, 2009.
31. M. Laurent, L.F. Vargas. Finite Convergence of Sum-of-Squares Hierarchies for the Stability Number of a Graph. *SIAM Journal on Optimization*, 32(2):491-518, 2022.
32. M. Laurent, L.F. Vargas. Exactness of Parrilo's conic approximations for copositive matrices and associated low order bounds for the stability number of a graph, *Math. Oper. Res.*, to appear, published online [https://doi.org/10.1287/moor.2022.1290, 2022.]
33. M. Laurent, L.F. Vargas. On the exactness of sum-of-squares approximations for the cone of 5×5 copositive matrices, *Linear Algebra Appl.*, 651: 26–50, 2022.
34. L Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25:1–7, 1979.
35. M. Marshall. Representations of non-negative polynomials having finitely many zeros. *Annales de la Faculté des Sciences de Toulouse*, 15(3):599–609. 2006.
36. M. Marshall. Representation of non-negative polynomials, degree bounds and applications to optimization. *Canad. J. Math.*, 61(1):205–221, 2009.
37. T. S. Motzkin. The arithmetic-geometric mean. In: *Inequalities* (Proceedings of a Symposium Wright-Patterson Air Force Base,Ohio), 1965, pp. 205-224. Academic Press, NewYork (1967)
38. T.S. Motzkin, E.G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540. 1965.
39. K.G. Murty, S.N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Math. Programming,* 39(2):117–129, 1987.
40. J. Nie. Optimality conditions and finite convergence of Lasserre's hierarchy. *Mathematical Programming*, 146(1-2):97–121, 2012.
41. J. Nie, Z. Yang, X. Zhang. A complete semidefinite algorithm for detecting copositive matrices and tensors. *SIAM Journal on Optimization*, 28(4): 2902–2921, 2018.
42. P.A. Parrilo. Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. PhD thesis, California Institute of Technology, 2000.
43. J. Peña, J.Vera, L.F. Zuluaga. Computing the stability number of a graph via linear and semidefinite programming. *SIAM Journal on Optimization*, 18(1):87–105, 2007.
44. G. Pólya. Über positive Daarstellung von Polynomen. *Naturforsch. Ges. Zurich*, (73):141–145, 1928.
45. M. Putinar. Positive polynomials on compact semi-algebraic sets. *Ind. Univ. Math. J.*, 42:969–984, 1993.
46. L. Qi. Symmetric nonnegative tensors and copositive tensors. *Linear Algebra and its Applications*, 439(1):228–238, 2013.
47. B. Reznick. Uniform denominators in Hilbert's seventeenth problem, *Math. Z.*, 220:75–97, 1995.
48. B. Reznick. Some concrete aspects of Hilbert's 17th problem, In *Real Algebraic Geometry and Ordered Structures*, C.N. Delzell and J.J. Madden (eds.), Contemporary Mathematics, 253:251–272, 2000.
49. C. Scheiderer. Sums of squares on real algebraic surfaces. *Manuscripta Mathematica*, 119:395–410, 2006.
50. K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Math. Ann.*, 289:203–206, 1991.
51. A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Trans. Inform. Theory*, 25: 425–429, 1979.

52. M. Schweighofer, L.F. Vargas. Sum-of-squares representations for copositive matrices and the stability number of a graph. In preparation, 2023+.
53. N. Shaked-Monderer, A. Berman. Copositive and completely positive matrices. *World Scientific* Publishing Co. Pte. Ltd., Hackensack, NJ, 2021.
54. Y. Song and L. Qi. Necessary and sufficient conditions for copositive tensors. *Linear Multilinear Algebra,* 63:120–131, 2015.
55. L.F. Zuluaga, J. Vera, J. Peña. LMI approximations for cones of positive semidefinite forms, *SIAM J. Optim.*, 16:1076–1091, 2006.