On approximations, complexity, and applications for copositive programming

Gijben, Luuk

Chapter 1

Introduction to copositive programming

In relatively recent years copositive programming has become a useful tool to deal with a wide variety of optimization problems, in particular with respect to the class of NP-hard problems, as the problem of checking for copositivity is itself NP-hard. This complexity result is due to Murty and Kabadi [MK87] who in fact showed that the problem of seeing whether a matrix is in the copositive cone is co-NP-complete. In Chapter 1 we will review some of the results obtained within copositive programming so far, providing a thorough introduction and overview of the field as well as giving a number of results that will be needed throughout this text. We shall start the introduction in Section 1.2 by focusing on some of the properties of the copositive cone and its dual, the completely positive cone. In Section 1.3 we will review a number of approximation techniques that are currently known for the copositive cone as well as the completely positive cone. In Section 1.4 we will give an overview of the field of copositive programming, describing optimization techniques as well as a number of problems known to be rewritable as a copositive (completely positive) program. However, first we will need to explain some of the notation used in this thesis, which we will do in Section 1.1. In particular, we will define most of the general notation that will be used throughout this text. More specific notation that will only be used once or twice, or only during a specific chapter or section, will further be defined throughout the rest of the text as it is needed. We included a nomenclature and an index at the end of the text to aid in the readability of this thesis.

1.1 Notation

We denote the space of n-dimensional real vectors by $\mathbb{R}^n$, rational vectors by $\mathbb{Q}^n$, vectors containing natural numbers by $\mathbb{N}^n$, and integer vectors by $\mathbb{Z}^n$. By $\mathbb{R}^n_+$ we denote the set of nonnegative real vectors, while $\mathbb{R}^n_{++}$ denotes the set of strictly positive real vectors. Moreover the space of $n \times m$ real matrices is denoted by $\mathbb{R}^{n\times m}$. Similar notation will be used for the other sets previously defined, e.g. $\mathbb{Q}^{n\times m}_+$ is the set of nonnegative rational $n\times m$ matrices. The set of symmetric real matrices of order n shall be denoted by $\mathbb{S}^n$ and the cone of $n\times n$ nonnegative real matrices by $\mathcal{N}^n$. We let Δ_n be the n-dimensional standard simplex. From now on we will omit the 'n' for any set if the dimension is equal to one, or when the dimension is clear from the context, in an effort to increase readability. All vectors will be printed in bold text to distinguish them from other variables and parameters. Likewise, we will generally use lowercase letters to denote scalars or vectors and capital letters to denote matrices. By $\mathbb{R}[\mathbf{x}]$ we denote the ring of polynomials in $\mathbf{x}$ with coefficients in $\mathbb{R}$.

The identity matrix and the all-ones matrix of order n will be written as I_n and E_n respectively. Similarly we let $\mathbf{e}_n$ be the vector of all-ones. Whenever we say $A \in \mathbb{R}^{n\times m}$ is nonnegative, or write $A \geq 0$, we mean that A is entry-wise nonnegative.

For any matrix $A \in \mathbb{R}^{n\times n}$, the operator $\mathrm{Tr} : \mathbb{R}^{n\times n} \to \mathbb{R}$ is the trace of A, i.e. the sum of diagonal elements of A. For the inner product we use the standard dot product, $\langle \mathbf{a}, \mathbf{b}\rangle = \mathbf{a}^\mathsf{T}\mathbf{b}$, when dealing with vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and the trace inner product, $\langle A, B\rangle = \mathrm{Tr}(AB)$, when dealing with matrices $A, B \in \mathbb{S}^n$. The accompanying norms will be the Euclidean and the Frobenius norms respectively, that is $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a}\rangle}$ for $\mathbf{a} \in \mathbb{R}^n$ and $\|A\| = \sqrt{\langle A, A\rangle}$ for $A \in \mathbb{S}^n$. For $\mathbf{a} \in \mathbb{Z}^n_+$ we define $|\mathbf{a}| = \sum_{i=1}^n |a_i|$. For any set $K \subseteq \mathbb{R}^{n\times m}$ we denote its convex hull by $\mathrm{conv}(K)$ and its conic hull by $\mathrm{cone}(K)$. The closure of a set K shall be denoted $\mathrm{cl}(K)$ whereas its interior will be written as $\mathrm{int}(K)$. Then by $\dim(K)$ we denote the dimension of the set K.

The operator $\mathrm{Diag} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ turns a vector $\mathbf{d}$ into a diagonal matrix D where $D_{ii} = d_i, i = 1, \ldots, n$. Conversely $\mathrm{diag} : \mathbb{R}^{n\times n} \to \mathbb{R}^n$ returns a vector $\mathbf{a}$ such that $a_i = A_{ii}$, $i = 1, \ldots, n$, for any (not necessarily diagonal) matrix A. By a *scaling* of a matrix $A \in \mathbb{S}^n$ we mean the multiplication DAD, where $D := \mathrm{Diag}(d)$, $d \in \mathbb{R}_{++}$. Similarly, a matrix PAP, for some $A \in \mathbb{S}^n$ and $P \in \mathcal{P}_n$, is referred to as a *permutation* of A, where $\mathcal{P}_n$ is the set of all $n \times n$ permutation matrices. The operator $\mathrm{Vec} : \mathbb{R}^{n\times m} \to \mathbb{R}^{nm}$ turns a matrix $A = [a_1, \ldots, a_m]$, where a_i is the i^{th} column of A, into one long vector by stacking the columns of A, that is $\mathrm{Vec}(A) = [a_1^\mathsf{T}, \ldots, a_m^\mathsf{T}]^\mathsf{T}$. The Hadamard product of two matrices $A, B \in \mathbb{R}^{n\times m}$ is denoted as $A \circ B$. By $A \otimes B$ we mean the Kronecker product of the matrices A and B, for $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times q}$.

A graph is a pair $G = (V, E)$, where V (the vertex set) is some finite nonempty set and E (the edge set) is a set of 2-element subsets of V. By $\bar{G} = (V, \bar{E})$ we denote the complement graph of $G = (V, E)$ where $\{i, j\} \in \bar{E}$ if and only if $\{i, j\} \notin E$, for all $i, j \in V$.

When we refer to a *polynomial (respectively exponential) time* algorithm we mean that the maximum or worst-case computation time of the algorithm is polynomial (respectively exponential) in the encoding lengths of the inputs. Alternatively we say a problem is tractable if a polynomial time (or efficient) algorithm for it exists. By $\mathcal{O}(\bullet)$ we denote the usual big-O notation concerning limit behavior of functions, i.e. for any two functions f and

$g : \mathbb{R} \to \mathbb{R}$ we have $f(x) = \mathcal{O}(g(x))$ if there exists a constant $M \in \mathbb{R}_{++}$ such that $\limsup_{x\to\infty} |\frac{f(x)}{g(x)}| \leq M$. For thorough introductions to complexity theory the reader is referred to Garey and Johnson [GJ79], and Grötschel, Lovász, and Schrijver [GLS88], and de Klerk [dK08]. During the remainder of this thesis we shall assume that the reader is familiar with basic complexity theory, including an understanding of encoding lengths.

1.2 The copositive cone and its dual

A symmetric matrix A is said to be *copositive* if the quadratic function $\mathbf{x}^\mathsf{T} A\mathbf{x}$ is nonnegative over the nonnegative orthant, that is for $\mathbf{x} \geq 0$. We will denote the cone of copositive matrices as follows.

Definition 1.1. The copositive cone is defined as

$$\mathcal{COP}^n = \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n_+\}. \tag{1.1}$$

The first ever mention of copositive forms, as well as the coinage of the term 'copositive' is credited to Motzkin [Mot52]. As mentioned before, the membership problem for the copositive cone, that is checking whether or not a given matrix is in $\mathcal{COP}^n$, is NP-hard. It turns out (see e.g. [BD08]) that in the definition of the copositive cone (1.1) we can limit $\mathbf{x}$ such that its Euclidean norm is equal to one, i.e.

$$\mathcal{COP}^n = \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x}\text{for all}\mathbf{x} \in \mathbb{R}^n_+,\ \|\mathbf{x}\| = 1\}.$$

We will use this fact later on in Chapter 2 to prove some complexity results. It should be noted that the definition of the copositive cone looks very similar to that of the well-studied *semidefinite cone*

$$\mathcal{S}^n_+ := \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n\}. \tag{1.2}$$

Yet copositive programming falls in the category of NP-hard problems whereas semidefinite programs are known to be solvable in polynomial time up to any desired accuracy. Note that this is currently the strongest known result with respect to the complexity of semidefinite programming and can be shown with the help of the so called *ellipsoid method* [YN76], [Sho77]. In particular it was the paper by Shor [Sho77] that gave the first explicit statement of the ellipsoid method as it is currently known. The ellipsoid method was also employed by Khachiyan [Kha79] to show that linear programming was in P. For a more thorough introduction to the ellipsoid method we refer the reader to Grötschel et al. [GLS88]. An important part of the ellipsoid method is the need for a so called *oracle* that decides whether or not a matrix is in (or within some guaranteed distance to) the feasible set. In Chapter 2 we will use a framework and machinery related to the ellipsoid method to prove, among other things, a complexity result for the dual of the copositive cone. For the semidefinite case an oracle is obtained with the help of the Cholesky factorization, which decomposes a matrix $A \in \mathcal{S}_+^n$ such that $A = BB^\mathsf{T}$, where $B \in \mathbb{R}^{n \times n}$ is a lower triangular matrix. The existence of such a decomposition, or in fact any decomposition $A = BB^\mathsf{T}$ with $B \in \mathbb{R}^{n \times k}$, provides an alternative definition of a positive semidefinite matrix, a fact easily seen from the definition of the semidefinite cone. In fact one of the convenient properties of the semidefinite cone is that it has several such equivalent definitions which we will list in the following proposition.

Proposition 1.2. *Let $A \in \mathbb{S}^n$, then the following statements are equivalent:*

i. the matrix A is positive semidefinite (resp. positive definite),

ii. all eigenvalues of A are nonnegative (positive),

iii. the determinant of every principal submatrix of A is nonnegative (positive),

iv. every principal submatrix of A is positive semidefinite (positive definite),

v. there exists a (nonsingular square) matrix B such that $A = BB^\mathsf{T}$.

For a proof of Proposition 1.2 we refer the reader to [BSM03, Theorem 1.10]. For the copositive cone, despite its similarity with the semidefinite cone in terms of its definition, very few conditions exist that are both necessary and sufficient. Moreover any such conditions, as we will see later on, are far less practical than their semidefinite counterparts. An exception to this observation is when $n \leq 4$, in which case we have that $\mathcal{COP}^n = \mathcal{S}_+^n + \mathcal{N}^n$. On the other hand, a couple of conditions similar to the conditions listed in Proposition 1.2 do exist, which instead are only necessary or only sufficient. We present these conditions concerning copositive matrices together with several others in the following proposition.

Proposition 1.3. *Let $A \in \mathcal{COP}^n$. Then we have:*

i. if P is a permutation matrix and D is a nonnegative diagonal matrix, then $PDADP^\mathsf{T} \in \mathcal{COP}^n$.

ii. every principal submatrix of A is also copositive.

iii. $(A)_{ii} \geq 0$ for all i.

iv. *if* $(A)_{ii} = 0$ *for some* i*, then* $(A)_{ij} \geq 0$ *for all* j*.*

v. $(A)_{ij} \geq -\sqrt{(A)_{ii}(A)_{jj}}$ *for all* i, j*.*

vi. *if there exists a strictly positive vector* $\mathbf{v}$ *such that* $\mathbf{v}^\mathsf{T} A\mathbf{v} = 0$*, then* A *is positive semidefinite,*

vii. *at least one eigenvalue, and moreover the sum of all eigenvalues, of* A *is nonnegative.*

For proofs of the properties (*i*) - (*vi*) we refer the reader to [Dia62] and [Dic13]. For a proof of (*vii*) see [HUS10], the latter paper moreover gives a condition in terms of so called Pareto eigenvalues that is both necessary and sufficient. It is a well known fact that the inequalities from (*v*) together with (*iii*) of Proposition 1.3 provide conditions for copositivity that are both necessary and sufficient for 2×2 matrices. In particular, this result follows immediately from the fact that $\mathcal{COP}^2 = \mathcal{S}^2_+ + \mathcal{N}^2$ combined with property (*iii*) of Proposition 1.2. Similar, allbeit more expensive, inequalities for checking copositivity can be obtained for the 3×3 [CS94] and 4×4 [PY93] case.

Note that although every principal submatrix of a copositive matrix is also copositive, the reverse does not hold. Sufficient, as well as necessary conditions in terms of principal submatrices, can however be constructed, as was shown by Gaddum [Gad58]. In this paper a recursive strategy for detecting copositivity is suggested based on a link to game theory, that is, the author obtains the following result.

Theorem 1.4. *Let* A *be a symmetric matrix and suppose that every principal submatrix of* A *is copositive. Then* A *is copositive if and only if*

$$\min_{\mathbf{x}\in\Delta_n} \max_{\mathbf{y}\in\Delta_n} \mathbf{y}^\mathsf{T} A\mathbf{x} = \max_{\mathbf{y}\in\Delta_n} \min_{\mathbf{x}\in\Delta_n} \mathbf{y}^\mathsf{T} A\mathbf{x} > 0. \tag{1.3}$$

The equality (1.3) can be verified via linear programming. Another recursive copositivity test requiring all principle submatrices to be copositive, is suggested by Cottle, Habetler, and Lemke [CHL70]. This result furthermore puts conditions on the so called adjugate matrix. For $A \in \mathbb{R}^{n\times n}$ the *adjugate matrix* of A is defined as an $n\times n$ matrix with $\operatorname{Adj}(A)_{ji} = (-1)^{i+j} \det(A(i, j))$, where $A(i, j)$ is a submatrix obtained from A by deleting row i and column j.

Theorem 1.5. *Let* A *be a symmetric matrix and suppose that every principle submatrix of* A *is copositive. Then* A *is copositive if and only if the determinant of* A *is nonnegative or the adjugate matrix of* A *contains a negative entry.*

Bomze [Bom87, Bom89] suggested yet another recursive way to check copositivity by introducing a criterion that looks similar to the Schur-complement for semidefinite matrices, albeit far more restricted.

Theorem 1.6. *Let* $\mathbf{b} \in \mathbb{R}^{n-1}$ *and* $C \in \mathbb{S}^{n-1}$. *Then* $\begin{bmatrix} a & \mathbf{b}^\mathsf{T} \\ \mathbf{b} & C \end{bmatrix}$ *is copositive if and only if*

i. $a \geq 0$ *and* C *is copositive, and*

ii. $\mathbf{x}^\mathsf{T}(aC - \mathbf{b}\mathbf{b}^\mathsf{T})\mathbf{x} \geq 0$ *for all* $\mathbf{x} \in \mathbb{R}^{n-1}_+$ *for which* $\mathbf{b}^\mathsf{T}\mathbf{x} \leq 0$.

An interesting remark regarding these conditions comes from [HUS10]. In that paper the authors observe that Theorem 1.6 boils down to checking copositivity of the 2×2 matrices

$$\begin{bmatrix} a & \mathbf{b}^\mathsf{T}\mathbf{x} \\ \mathbf{b}^\mathsf{T}\mathbf{x} & \mathbf{x}^\mathsf{T}C\mathbf{x} \end{bmatrix}$$

for all $\mathbf{x} \in \mathbb{R}^{n-1}_+$, which we can do by verifying a number of inequalities via properties (*iii*) and (*v*) of Proposition 1.3, as was mentioned before.

An algorithmic approach to check for copositivity was introduced in [BD08]. In this paper the authors introduce a formulation of the quadratic form $\mathbf{x}^\mathsf{T}A\mathbf{x}$ in barycentric coordinates with respect to the standard simplex. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be the vertices of a simplex. Then a point $\mathbf{x}$ in the simplex can be written as $\mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{v}_i$ with $\lambda_i \geq 0$ and $\sum_{i=1}^n \lambda_i = 1$. Then

$$\mathbf{x}^\mathsf{T}A\mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \mathbf{v}_i^\mathsf{T} A \mathbf{v}_j \tag{1.4}$$

and so, for A to be copositive, a sufficient condition is

$$\mathbf{v}_i^\mathsf{T} A \mathbf{v}_j \geq 0 \text{ for all } i, j = 1, \ldots, n. \tag{1.5}$$

Partitioning the standard simplex into smaller simplices gives increasingly more and stronger constraints of this type. As it turns out, in the limit these type of constraints describe the copositive cone. A more general algorithm that also allows for a more general partitioning strategy is suggested in [BE12]. Moreover this paper provides a number of sufficient conditions for copositivity via linear and quadratic programming techniques.

The idea of using the standard simplex and barycentric coordinates has been considered before in [ACE95] where criteria for copositivity of matrices in $\mathbb{S}^n$ were constructed under the assumption that all principle submatrices (of order $n-1$) are copositive.

Another property that sets the copositive cone apart from the semidefinite cone concerns the dual cone.

Definition 1.7. Given a set $K \subseteq \mathbb{S}^n$, the *dual set* is defined by

$$K^* := \{A \in \mathbb{S}^n \mid \langle A, B\rangle \geq 0 \text{ for all } B \in K\}.$$

A set K is called *self-dual* if $K^* = K$.

Note that the semidefinite cone is a self-dual cone, a fact that can easily be seen from the definition given in (1.2) combined with the equivalent definition provided by (*iv*) of Proposition 1.2. To show the inclusion $(\mathcal{S}^n_+)^* \subseteq \mathcal{S}^n_+$, take $A \in (\mathcal{S}^n_+)^*$. We can take an arbitrary $\mathbf{b} \in \mathbb{R}^n$ and obtain

$$0 \leq \langle A, \mathbf{b}\mathbf{b}^\mathsf{T}\rangle = \mathbf{b}^\mathsf{T} A\mathbf{b}, \tag{1.6}$$

so $A \in \mathcal{S}^n_+$. To see the reverse inclusion, take $A \in \mathcal{S}^n_+$. To show that $A \in (\mathcal{S}^n_+)^*$, take $B \in \mathcal{S}^n_+$ and decompose as $B = \sum_{i=1}^k \mathbf{b}_i\mathbf{b}_i^\mathsf{T}$. Then

$$\langle A, B\rangle = \sum_{i=1}^k \langle A, \mathbf{b}_i\mathbf{b}_i^\mathsf{T}\rangle = \sum_{i=1}^k \mathbf{b}_i^\mathsf{T} A\mathbf{b}_i \geq 0, \tag{1.7}$$

which shows $A \in \mathcal{S}^n_+$.

Observe that the inequalities (1.6) and (1.7) in the proof given above stay valid when we let $A \in \mathcal{COP}^n$ as long as we furthermore demand that all vectors are nonnegative, i.e. $\mathbf{l}_1, \dots, \mathbf{l}_k, \mathbf{b} \in \mathbb{R}^n_+$. This immediately gives us the definition of the dual of $\mathcal{COP}^n$ which is known as the *completely positive cone.* The completely positive cone was first mentioned roughly ten years after the introduction of the copositive cone in a paper by Hall Jr. [HJ62].

Definition 1.8. The completely positive cone is defined as

$$\mathcal{CP}^n = \{A \in \mathbb{S}^n \mid A = \sum_{i=1}^k \mathbf{b}_i\mathbf{b}_i^\mathsf{T},\, \mathbf{b}_i \geq 0 \text{ for all } i\}$$

and is the dual of the copositive cone, that is $(\mathcal{CP}^n)^* = \mathcal{COP}^n$ and $\mathcal{CP}^n = (\mathcal{COP}^n)^*$. A decomposition $A = \sum_{i=1}^k \mathbf{b}_i\mathbf{b}_i^\mathsf{T}$ with $\mathbf{b}_i \in \mathbb{R}^n_+$ for all i is referred to as a *rank 1 decomposition.* More generally, for any matrix $A \in \mathcal{CP}^n$, $A = \sum_{i=1}^k B_iB_i^\mathsf{T}$ with $B_i \in \mathbb{R}^{n\times k}$ is referred to as a rank k decomposition of A.

As we will see in Chapter 2 the problem of deciding whether a matrix is completely positive or not is an NP-hard problem as well. Moreover, in Chapter 2 we will even give slightly stronger complexity results for both the copositive and the completely positive cone than the NP-hardness for their respective membership problems. Whether or not the membership problem for the completely positive cone is in the class NP as well is still an open question at the moment of writing. This question is connected to the well-studied problem of finding the so called CP-rank of a matrix.

Definition 1.9. Let $A \in \mathcal{CP}^n$, that is $A = \sum_{i=1}^{k} \mathbf{b}_i \mathbf{b}_i^\mathsf{T}$, where $\mathbf{b}_i \geq 0$ for $i = 1, \dots, k$. Then the *CP-rank* of A is defined as the smallest possible such k.

For a detailed introduction to this topic the reader is referred to Chapter 3 of [BSM03]. It should be noted that bounds for the CP-rank exist that, moreover, are of polynomial nature with respect to n, the order of the matrices. This however does not solve the question whether or not membership for $\mathcal{CP}^n$ is in NP. The reason for this is that the entries of A and therefore also the entries of the vectors $\mathbf{b}_1, \dots \mathbf{b}_k$ in Definition 1.9 are allowed to be in $\mathbb{R}_+$ and hence their encoding lengths need not be bounded with respect to n. An interesting development in this area is that the Drew-Johnson-Loewy conjecture that is mentioned in [BSM03] has since been almost entirely disproven, at least for the general case, in recent papers by Bomze et al. [BSU14a, BSU14b]. The Drew-Johnson-Loewy conjecture states that the upper bound to the CP-rank is equal to the lower bound $\lfloor \frac{n}{4} \rfloor$ that was found by Drew, Johnson, and Loewy [DJL94]. Bomze et al. [BSU14a] managed to find counterexamples on the boundary of the completely positive cone for $7 \leq n \leq 11$. They then managed to construct counterexamples for $n \geq 12$ as well in [BSU14b] obtaining a new bound for the CP-rank, $\frac{n^2}{2} + \mathcal{O}(n^{3/2})$. Shortly before these two papers, disproving the conjecture for the general case, the conjecture was shown to hold for the 5×5 case by Shaked-Monderer et al. [SMBJS13]. Whether or not the conjecture holds for 6×6 completely positive matrices is still an open question.

As with the copositive cone, equivalent definitions for complete positivity in the form of practical conditions that are both necessary and sufficient, do not currently seem to exist. Several necessary conditions have been found, some of which we will list in the following proposition.

Proposition 1.10. *Let $A, B \in \mathcal{CP}^n$. Then*

i if P is a permutation matrix and D is a nonnegative diagonal matrix, then $PDADP^\mathsf{T} \in \mathcal{CP}^n$, *ii* *every principal submatrix of* A *is completely positive,*

iii $A \otimes B$ *is also completely positive.*

For proofs we refer the reader to [BSM03]. As with the copositive cone, sufficient conditions exist as well, for example conditions based on diagonal dominance and comparison matrices. We say that a matrix A is *diagonally dominant (dd)* if $|(A)_{ii}| \geq \sum_{j \neq i} |(A)_{ij}|$ for all i. Furthermore, a matrix A is called *scaled diagonally dominant (sdd)* if there exists a positive diagonal matrix D such that DAD is diagonally dominant. For dd and sdd matrices we have the following theorem.

Theorem 1.11. *Let* $A \in \mathbb{S}^n$ *be a scaled diagonally dominant matrix with nonnegative diagonal entries. Then* $A \in \mathcal{S}_+^n$.

Proof. If A is sdd then DAD is dd for some diagonal matrix D with positive diagonal entries. The result now follows from [BSM03, Proposition 1.8] in combination with (v) of Proposition 1.2. □

Finally, the *comparison matrix* of A is defined as

$$M(A)_{ij} = \begin{cases} |(A)_{ij}| & \text{if } i = j, \\ -|(A)_{ij}| & \text{if } i \neq j. \end{cases}$$

Now there is the following results on complete positivity.

Proposition 1.12. *Let* $A \in \mathcal{N}^n$. *Then*

i *if* A *is (scaled) diagonally dominant then* A *is completely positive,*

ii *if the comparison matrix* $M(A)$ *is positive semidefinite, then* A *is completely positive.*

iii *if* A *is positive semidefinite with rank equal to* r, *and if* A *has an* $r \times r$ *principle submatrix that is diagonal, then* A *is completely positive.*

For proofs of (i) and (ii) we again refer the reader to [BSM03]. It should be noted that (i) is only shown for diagonally dominant matrices in [BSM03]. The fact that this property also holds for scaled diagonally dominant matrices can easily be seen from the fact that if $DAD \in \mathcal{CP}^n$ for a diagonal matrix D with positive diagonal then $A \in \mathcal{CP}^n$. Moreover, several more technical and more specific sufficient conditions can be found in that same book, including several conditions based on graph theoretical concepts. Condition (iii) is a result by Shaked-Monderer [SM09]. This result has slightly been improved upon by

Kalofolias and Gallopoulos [KG12] who manage to construct an explicit rank 2 decomposition for this case.

Dong, Lin and Chu [DLC14] developed a heuristic method to check complete positivity. Their method works as long as the CP-rank of the matrix being evaluated is equal to the rank of that matrix. Another approach is suggested by Dickinson and Dür [DD12], who present an algorithm that deals with sparse matrices. They prove that this algorithm can be used to determine complete positivity, as well as give an explicit factorization, for tridiagonal and acyclic matrices in linear time.

Both the copositive and the completely positive cone are so called *proper cones*, i.e. both cones are full dimensional, closed, convex, and pointed. For proofs we refer the reader to [Dic13, Chapter 5]. The interior of the copositive cone is described, analogous to the interior of the positive semidefinite cone, as

$$\mathrm{int}(\mathcal{COP}^n) = \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A \mathbf{x} > 0 \text{ for all } \mathbf{x} \in \mathbb{R}_+ \setminus \{0\}\}. \tag{1.8}$$

The interior of the completely positive cone, however, is not quite as straight forward. It was investigated in [DS08] where the authors obtained the following description:

$$\mathrm{int}(\mathcal{CP}^n) = \{A \in \mathbb{S}^n \mid A = BB^\mathsf{T},\, B = [B_1|B_2] \text{ where } B_1 > 0 \text{ nonsingular, } B_2 \geq 0\}$$

Several alternative definitions of the interior of the completely positive cone are given in [Dic13, Chapter 7].

A full description of the copositive cone in terms of its extreme rays is currently unknown. For $n \leq 4$, however, we know that $\mathcal{COP}^n = \mathcal{S}_+^n + \mathcal{N}^n$, for which all extreme rays are known from [HN63]. Attempts to go further in this direction have been made by Baumert [Bau65, Bau66, Bau67], Baston [Bas69], and Ycart [Yca82]. The paper by Hall and Newman [HN63] gives a full description of the extreme rays of the completely positive cone. If we let $\mathrm{Ext}(K)$ be the set of extreme rays for some set $K \in \mathbb{R}^{n\times m}$, then Hall and Newman show that

$$\mathrm{Ext}(\mathcal{CP}^n) = \{\mathbf{a}\mathbf{a}^\mathsf{T} \mid \mathbf{a} \in \mathbb{R}_+^n \setminus \{\mathbf{0}\}\}. \tag{1.9}$$

Moreover it was shown in [Dic11] that all such extreme rays are also exposed rays. For the case of 5×5 matrices a full description of the extreme rays of the copositive cone is given by Hildebrand [Hil12]. Details of this description will be given in Chapter 4 of this thesis. For $n \geq 6$ a full characterization of the extreme rays of $\mathcal{COP}^n$ is not known. However a number of extreme rays of the copositive cone have been found and investigated by [HN63, Dic11]. In particular [Dic13] gives an overview of the currently known results in this area and also manages to give a full description of the maximal faces of the copositive cone. For further surveys on properties of the copositive and completely positive cone, the reader is referred to Hiriart-Urruty and Seeger [HUS10] and Berman and Shaked-Monderer [BSM03].

1.3 Approximation hierarchies

Over the last several years a lot of effort has been spent to obtain approximations for the copositive cone, as well as for the completely positive cone. The necessity for such approximations seems two-fold. On the one hand the complexity results regarding these cones imply that no efficient oracles exist unless P= NP. In fact, as we will see later in Chapter 2 even the problem of deciding whether a matrix is within a guaranteed distance of the copositive or completely positive cone is NP-hard. In practice this means that none of the general conic optimization frameworks can be used to solve copositive programs for even medium sized instances. On the other hand no non-efficient easy to use state-of-the-art solvers for copositive programs are currently available either. Hence, in order to deal with copositive programs, especially for medium to large instances, we are left to approximations which are generally more tractable and moreover, as we shall see below, can make use of existing frameworks and solvers.

Some obvious approximations include the cones $\mathcal{S}_+^n$, $\mathcal{S}_+^n + \mathcal{N}^n$ and $\mathcal{S}_+^n \cap \mathcal{N}^n$. The latter is referred to in the literature as the *doubly nonnegative cone*. From the definitions we get the following simple inclusions

$$\mathcal{CP}^n \subseteq \mathcal{S}_+^n \cap \mathcal{N}^n \subseteq \mathcal{S}_+^n + \mathcal{N}^n \subseteq \mathcal{COP}^n,$$

where we get equality for $\mathcal{CP}^n \subseteq \mathcal{S}_+^n \cap \mathcal{N}^n$ and $\mathbb{S}^n + \mathcal{N}^n \subseteq \mathcal{COP}^n$ if and only if $n \leq 4$. Examples showing that this is not true anymore for $n \geq 5$, are the so called *Horn-matrix* [HN63],

$$H = \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix} \in \mathcal{COP}^5 \setminus (\mathcal{S}_+^5 + \mathcal{N}^5), \tag{1.10}$$

and the matrix

$$H^* = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 & 6 \end{pmatrix} \in (\mathcal{S}_+^5 \cap \mathcal{N}^5) \setminus \mathcal{CP}^5, \qquad (1.11)$$

which is taken from [BSM03]. For a more thorough description of the difference between the cones $\mathcal{CP}^n$ and $\mathcal{S}_+^n \cap \mathcal{N}^n$, especially for the 5×5 case, we refer the reader to [BAD09] and [DA13]. Moreover [DA13] provides a construction that separates matrices in $(\mathcal{S}_+^5 \cap \mathcal{N}^5) \setminus \mathcal{CP}^5$ from $\mathcal{CP}^n$, as well as such matrices of arbitrary size having some block structure.

1.3.1 Approximations via simplicial partitioning

Note that the cones mentioned above can be handled using the well studied framework of semidefinite programming. In fact many of the existing approximations for copositive programs involve the construction of cones that make use of semidefinite programming techniques. An exception to this is the technique introduced in [BD08], which is based on the conditions in (1.5). The set described by these conditions gets refined in the approach described by Bundfuss and Dür resulting in a hierarchy of cones approximating $\mathcal{COP}^n$ from within, a result made explicit in [BD09]. Assume we partition the standard simplex $r \in \mathbb{Z}_+$ times into a number of smaller simplices, then denote by $\mathbb{P}_r \subseteq \Delta_n$ and $\mathbb{E}_r$ the sets of vertices and edges of these simplices respectively. The authors then introduce the hierarchy of inner approximations of $\mathcal{COP}^n$

$$\mathcal{I}_r^n := \left\{ A \in \mathbb{S}^n \;\middle|\; \begin{array}{l} \mathbf{v}^\mathsf{T} A \mathbf{v} \geq 0 \text{ for all } \mathbf{v} \in \mathbb{P}_r \\ \mathbf{u}^\mathsf{T} A \mathbf{v} \geq 0 \text{ for all } (\mathbf{u}, \mathbf{v}) \in \mathbb{E}_r \end{array} \right\}, \quad r \in \mathbb{Z}_+.$$

They furthermore show that these cones converge to the copositive cone, i.e. $\operatorname{cl}\left(\bigcup_{r \in \mathbb{Z}_+} \mathcal{I}_r^n\right) = \mathcal{COP}^n$. The authors suggest outer approximations for the copositive cone as well by relaxing the set of constraints (1.5) to $\mathbf{v}^\mathsf{T} A \mathbf{v} \geq 0$ instead. That is they introduce the hierarchy of outer approximations of $\mathcal{COP}^n$

$$\mathcal{O}_r^n := \left\{ A \in \mathbb{S}^n \mid \mathbf{v}^\mathsf{T} A \mathbf{v} \geq 0 \text{ for all } \mathbf{v} \in \mathbb{P}_r \right\}, \quad r \in \mathbb{Z}_+.$$

These cones are shown to converge to $\mathcal{COP}^n$ as well which means that $\bigcap_{r \in \mathbb{Z}_+} \mathcal{O}_r^n = \mathcal{COP}^n$. These approximations for the copositive cone can in turn be used to construct outer and inner approximations for the completely positive cone through duality. Explicit formulation of these cones approximating $\mathcal{CP}^n$ are provided in [BD09].

In general, nontrivial outer approximations for the copositive cone, and hence inner approximations of the completely positive cone, seem hard to find. They do exist, however their quality varies. Observe that we can trivially create outer approximations of the copositive cone, similar to the one described above, by simply defining a cone

$$\{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in U\}$$

where $U \subset \mathbb{R}^n_+$. For example we can define U as the set of binary vectors.

The idea of using simplicial partitioning is not unique to copostive programming though, and has been applied to other types of nonlinear optimization problems, see e.g. [Hor76, Kea78, TH88, Tuy88, Hor97]. For a more detailed description of such techniques the reader is referred to [Dic13, Chapter 9] and [HPT95].

1.3.2 The polyhedral approximation cones $\mathcal{C}_n^r$

The construction of so called approximation hierarchies like the ones we saw in the previous section, is an approach that is often employed to approximate the copositive or completely positive cone. A number of such hierarchies have been proposed in the literature, however unlike the hierarchies introduced in [BD09], most of them hinge on two basic concepts. First the condition $\mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n_+$ is replaced by some sufficient condition $\mathbf{x}^\mathsf{T} A\mathbf{x} \in \phi$ for nonnegativity that is closed under multiplication. Doing so automatically creates a cone

$$\Phi_n := \{A \in \mathbb{S}^n \mid \mathbf{x}^\mathsf{T} A\mathbf{x} \in \phi\} \subseteq \mathcal{COP}^n$$

that is a (not necessarily strict) subset of the copositive cone. This cone is then expanded upon by multiplying the quadratic term a number of times with some function f that is positive over $\mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ for which the condition $f \in \phi$ trivially holds, often chosen as $f(\mathbf{x}) = (\sum_{i=1}^n x_i)$ in existing hierarchies for $\mathcal{COP}^n$. This produces a hierarchy of cones

$$\Phi_n^r := \{A \in \mathbb{S}^n \mid f(\mathbf{x})^r \mathbf{x}^\mathsf{T} A\mathbf{x} \in \phi\} \subseteq \mathcal{COP}^n, \quad r \in \mathbb{Z}_+.$$

Note that $\Phi_n^r \subseteq \Phi_n^{r+1}$ for all $r \in \mathbb{Z}_+$ due to the fact that $f \in \phi$ by assumption while ϕ is closed under multiplication. Furthermore note that for any positive function f over $\mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ it holds that $f(\mathbf{x}) \cdot \mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n_+$ implies $\mathbf{x}^\mathsf{T} A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n_+$, ensuring that $\Phi_n^r \subseteq \mathcal{COP}^n$ for all $r \in \mathbb{Z}_+$.

An obvious sufficient condition for any polynomial over $\mathbb{R}^n_+$ to be nonnegative, is to demand that all its coefficients are nonnegative, which can be verified using linear programming techniques. This idea was employed by de Klerk and Pasechnick in [dKP02] to construct a hierarchy of polyhedral cones $\mathcal{C}_n^0, \mathcal{C}_n^1, \ldots$ that converge to $\mathcal{COP}^n$ which they defined as follows:

$$\mathcal{C}_n^r := \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^n x_i\Big)^r \mathbf{x}^\mathsf{T} A\mathbf{x} \text{ has nonnegative coefficients } \Big\}. \quad (1.12)$$

Note that if $\big(\sum_{i=1}^n x_i\big)^r \mathbf{x}^\mathsf{T} A\mathbf{x}$ has nonnegative coefficients for some $r \geq 1$ then $\big(\sum_{i=1}^n x_i\big)^{r+1} \mathbf{x}^\mathsf{T} A\mathbf{x} = \big(\sum_{i=1}^n x_i\big)\big(\sum_{i=1}^n x_i\big)^r \mathbf{x}^\mathsf{T} A\mathbf{x}$ must have nonnegative coefficients as well so that $\mathcal{C}_n^r \subseteq \mathcal{C}_n^{r+1} \subseteq \mathcal{COP}^n$. The convergence of the cones $\mathcal{C}_n^r$ that was claimed above is a direct result from a theorem by Pólya.

Theorem 1.13 (Pólya [Pól28])**.** *Let f be a homogeneous polynomial which is positive on the simplex $\Delta = \{\mathbf{z} \in \mathbb{R}^n_+ \mid \sum_{i=1}^n z_i = 1\}$. Then for sufficiently large N, all the coefficients of the polynomial $(\sum_{i=1}^n z_i)^N f(\mathbf{z})$ are nonnegative.*

This result implies that $\bigcup_{r \in \mathbb{Z}_+} \mathcal{C}_n^r = \operatorname{int}(\mathcal{COP}^n)$ so that $\operatorname{cl}(\bigcup_{r \in \mathbb{Z}_+} \mathcal{C}_n^r) = \mathcal{COP}^n$.

1.3.3 The Parrilo r-cones

The concept of creating hierarchies of cones approximating $\mathcal{COP}^n$ as described at the start of section 1.3.2, was first introduced by Parrilo in his thesis [Par00]. The sufficient condition he used was that of sum of squares, which can be dealt with via the framework of semidefinite programming. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is said to be *sum of squares (SOS)* if there exist a finite number of polynomials $g_1, \ldots, g_k \in \mathbb{R}[\mathbf{x}]$ such that $f = \sum_{i=1}^k g_i^2$. The connection between SOS and semidefinite programming is made explicit by the following theorem.

Theorem 1.14 (See e.g. [Lau08])**.** *Let $f \in \mathbb{R}[\mathbf{x}]$ be a polynomial of degree $2d$. Furthermore let $(\mathbf{x}^{[\leq d]})$ be the vector containing all monomials of degree at most d. Then*

$$f \text{ is SOS} \Leftrightarrow \text{ there exists a } B \in \mathcal{S}^k_+ \text{ such that } f = (\mathbf{x}^{[\leq d]})^\mathsf{T} B(\mathbf{x}^{[\leq d]}). \quad (1.13)$$

Note that for general polynomials of degree $2d$, the matrix B as in Theorem 1.14 is of order k, where

$$k = \sum_{i=1}^{d} \binom{n+i-1}{i}.$$

In order to apply Theorem 1.14 we require a polynomial to have an even degree. Hence, for $r \in \mathbb{Z}_+$ and taking into account the discussion above the Parrilo r-cone as introduced in [Par00] is defined as

$$\mathcal{K}_n^r := \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^{n} x_i^2\Big)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x}) \text{ is SOS } \Big\}. \tag{1.14}$$

The squares of the variables $x_1, \dots, x_n$ that appear in (1.14) can be seen as a replacement of the non-negativity of $\mathbf{x}$. More practically it means the resulting polynomial $(\sum_{i=1}^{n} x_i^2)^r (\mathbf{x} \circ \mathbf{x})^\mathsf{T} A (\mathbf{x} \circ \mathbf{x})$ has an even degree so that we can use Theorem 1.14 to verify membership of $\mathcal{K}_n^r$ using semidefinite programming.

As with the cones $\mathcal{C}_n^r$ it can immediately be seen that $\mathcal{K}_n^r \subseteq \mathcal{K}_n^{r+1} \subseteq \mathcal{COP}^n$, for every $r \in \mathbb{Z}_+$. Next, observe that any polynomial over $\mathbb{R}_+^n$ with nonnegative coefficients can trivially be written as a sum of squares so that $\mathcal{C}_n^r \subseteq \mathcal{K}_n^r$ which furthermore implies that $\text{cl}(\bigcup_{r \in \mathbb{Z}_+} \mathcal{K}_n^r) = \mathcal{COP}^n$.

As it turns out $\mathcal{K}_n^0 = \mathcal{S}_+^n + \mathcal{N}^n$. Moreover recall that for $n \leq 4$ we have $\mathcal{COP}^n = \mathcal{S}_+^n + \mathcal{N}^n$, so that in fact $\mathcal{K}_n^0 = \mathcal{COP}^n$ for $n \leq 4$. Clearly this also means that $\mathcal{K}_n^0 \neq \mathcal{COP}^n$ for $n \geq 5$ due to the example of the Horn matrix (1.10) mentioned before. In fact as we shall see later in Chapter 4 there exist no other cases where $\mathcal{K}_n^r = \mathcal{COP}^n$ other than for when $r = 0$ and $n \leq 4$. Note that determining whether or not a polynomial with degree $2d$ is sum of squares requires solving a rather large positive semidefinite program of order $\sum_{i=1}^{d} \binom{n+d-1}{n-1}$ together with a large number of linear equality constraints. Even in the case of checking membership for $\mathcal{K}_n^r$, where we have a homogeneous polynomial of degree $2r+4$, the resulting positive semidefinite matrix B is still of the order $\binom{n+r+1}{n-1}$. This means that for increasing values of r and n, the cones $\mathcal{K}_n^r$ become computationally expensive very quickly.

1.3.4 The SOS approximation cones $\mathcal{Q}_n^r$

Another approximation hierarchy closely related to the Parrilo r-cones, is the series of cones introduced by Peña et al. [PVZ07]. They again use sum of squares as a sufficient condition for non-negativity, as in the Parrilo cones, but rather in a more restrictive way. For $\mathbf{x} \in \mathbb{R}_+^n$ and $\boldsymbol{\beta} \in \mathbb{Z}_+^n$ denote $\mathbf{x}^{\boldsymbol{\beta}} = \prod_{i=1}^{n} x_i^{\beta_i}$

and recall that $|\boldsymbol{\beta}| = \sum_{i=1}^{n} \beta_i$, then for any $r \in \mathbb{Z}_+$ and $n \in \mathbb{N}$ Peña et al. define their hierarchy of cones as

$$\mathcal{Q}_n^r = \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^{n} x_i \Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} = \sum_{\boldsymbol{\beta} \in \mathbb{Z}_+^n, |\boldsymbol{\beta}| = r} \mathbf{x}^{\boldsymbol{\beta}} \mathbf{x}^\mathsf{T} B_\beta \mathbf{x},\ B_{\boldsymbol{\beta}} \in \mathcal{S}_+^n + \mathcal{N}^n \Big\}. \tag{1.15}$$

As with the Parrilo cones we again have the inclusions $\mathcal{Q}_n^r \subseteq \mathcal{Q}_n^{r+1} \subseteq \mathcal{COP}^n$ as well as the property that $\mathrm{cl}(\bigcup_{r \in \mathbb{Z}_+} \mathcal{K}_n^r) = \mathcal{COP}^n$, due to the fact that obviously $\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r$ for all $r \in \mathbb{Z}_+$. Again, recall that $\mathcal{S}_+^n + \mathcal{N}^n = \mathcal{K}_n^0$. Then immediately from Theorem 1.14 it follows that $\mathbf{x}^\mathsf{T} B_\beta \mathbf{x}$, and hence $\mathbf{x}^\beta \mathbf{x}^\mathsf{T} B_\beta \mathbf{x}$, is sum of squares for every $\boldsymbol{\beta} \in \mathbb{Z}_+^n$, $|\boldsymbol{\beta}| = r$. In other words, the cones $\mathcal{Q}_n^r$ break the polynomial $\big(\sum_{i=1}^{n} x_i\big)^r \mathbf{x}^\mathsf{T} A \mathbf{x}$ into a number of smaller bits which they then demand to be sum of squares representable. This way the authors restrict the more general SOS condition from the Parrilo cones $\mathcal{K}_n^r$ to a number of smaller SOS conditions. Furthemore note that the terms $\mathbf{x}^{\boldsymbol{\beta}}$ in the summation on the right-hand side of the equality in (1.15), are simply the monomials of degree r produced by $(\sum_{i=1}^{n} x_i)^r$. From this it can be seen that we can equivalently express membership of the cones $\mathcal{Q}_n^r$ via a set of *linear matrix inequalities (LMIs)*. In particular, it was noted in [PVZ07] that in order to decide whether $A \in \mathcal{Q}_n^r$, we can instead check whether there exist matrices $B_{\boldsymbol{\beta}} \in \mathcal{S}_+^n$, $\boldsymbol{\beta} \in \mathbb{Z}_+^n$ where $|\boldsymbol{\beta}| = r$ such that

$$\Big(\sum_{i=1}^{n} x_i \Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} \trianglerighteq \sum_{\boldsymbol{\beta} \in \mathbb{Z}_+^n, |\boldsymbol{\beta}| = r} \mathbf{x}^\beta \mathbf{x}^\mathsf{T} B_{\boldsymbol{\beta}} \mathbf{x}, \tag{1.16}$$

where "$\trianglerighteq$" indicates that the coefficients on the left-hand side are greater than or equal to those on the right-hand side for any given monomial. For example, to verify membership of $\mathcal{Q}_n^2$ for some matrix $A \in \mathbb{S}^n$ we obtain the LMIs

$$\Big(\sum_{i=1}^{n} x_i \Big)^2 \mathbf{x}^\mathsf{T} A \mathbf{x} \trianglerighteq \sum_{1 \le i \le j \le n} x_i x_j \mathbf{x}^\mathsf{T} B_{ij} \mathbf{x}, \quad B_{ij} \in \mathcal{S}_+^n. \tag{1.17}$$

Furthermore it is noted in [PVZ07] that we can use Proposition 9 from [PVZ06] to equivalently write the Parrilo cones as

$$\mathcal{K}_n^r = \Big\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^{n} x_i \Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} = \sum_{\boldsymbol{\beta} \in \mathbb{Z}_+^n, |\boldsymbol{\beta}| \le r+2} \mathbf{x}^{\boldsymbol{\beta}} \mathbf{x}^\mathsf{T} B_{\boldsymbol{\beta}} \mathbf{x},\ B_{\boldsymbol{\beta}} \in \mathcal{S}_+^n + \mathcal{N}^n \Big\}. \tag{1.18}$$

From this we immediately see that $\mathcal{Q}_n^r \subseteq \mathcal{K}_n^r$ for every $r \in \mathbb{Z}_+$ and $n \in \mathbb{N}$. Moreover, from (1.18) it becomes clear that the number of variables and LMIs for the $\mathcal{K}_n^r$ cones grows at a much faster rate, with both n and r, than for the cones $\mathcal{Q}_n^r$. Comparing to the example given above where $r = 2$, i.e. (1.17), we get a whole extra matrix variable B of size $\binom{n+1}{2} \times \binom{n+1}{2}$ for $\mathcal{K}_n^2$, i.e. denoting by $(\mathbf{x}^{[d]})$ the vector containing all monomials of degree exactly d we obtain the LMIs

$$\Big(\sum_{i=1}^{n} x_i\Big)^2 \mathbf{x}^\mathsf{T} A \mathbf{x} \trianglerighteq \sum_{1\leq i\leq j\leq n} x_i x_j \mathbf{x}^\mathsf{T} B_{ij} \mathbf{x} + (\mathbf{x}^{[2]})^\mathsf{T} B (\mathbf{x}^{[2]}).$$

1.3.5 Approximating cones via DSOS and SDSOS

As we have seen so far, particularly in Theorem 1.14, a polynomial can be determined to be sum of squares using semidefinite programming. In theory this means that for a given value of r we can determine membership for both $\mathcal{K}_n^r$ and $\mathcal{Q}_n^r$ for any matrix (up to any desired accuracy) in polynomial time with respect to the size n of this matrix. In practice however, especially for large scale optimization, semidefinite programming is not nearly as efficient as linear and second order cone programming. Moreover we can see from (1.15) and (1.18) that the size of the semidefinite programs increases quite rapidly with both n and r for these sum of squares problems. As a consequence of these two issues alternatives are needed to deal with large scale optimization. Several approximations have been suggested for this purpose based on linear or second order cone programming, an example are the cones $\mathcal{C}_n^r$ described above. Other examples include the hierarchies of cones introduced by Ahmadi and Majumdar [AM14]. Their approximations are based on relaxations of the SOS constraints. Ahmadi and Majumdar introduce two concepts called diagonally dominant sum of squares and scaled diagonaly dominant sum of squares. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is said to be *diagonally dominant sum of squares (DSOS)* if there exist monomials $m_i \in \mathbb{R}[\mathbf{x}]$ and nonnegative constants α_i and $\beta_{i,j}$, $i, j = 1, \ldots, k$ for some $k \in \mathbb{N}$, such that

$$f(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i m_i^2 + \sum_{i=1}^{k}\sum_{i=1}^{k} \beta_{i,j}(m_i \pm m_j)^2.$$

Similarly, a polynomial $f \in \mathbb{R}[\mathbf{x}]$ is said to be *scaled diagonaly dominant sum of squares (SDSOS)* if there exist monomials $m_i \in \mathbb{R}[\mathbf{x}]$ and constants $\alpha_i \in \mathbb{R}_+$ and $\beta_i, \gamma_i \in \mathbb{R}$, $i = 1, \ldots, k$ for some $k \in \mathbb{N}$ such that

$$f(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i m_i^2 + \sum_{i=1}^{k}\sum_{i=1}^{k} (\beta_i m_i \pm \gamma_j m_j)^2.$$

Analogue to the approximation hierarchies presented above the authors then introduce a hierarchy of sets *r-DSOS* and *r-SDSOS*. That is,

$$\text{r-DSOS} := \Big\{ f \in \mathbb{R}[\mathbf{x}] \mid \big(\sum_{i=1}^{n} x_i^2\big)^r f \text{ is DSOS } \Big\},$$

and

$$\text{r-SDSOS} := \Big\{ f \in \mathbb{R}[\mathbf{x}] \mid \big(\sum_{i=1}^{n} x_i^2\big)^r f \text{ is SDSOS } \Big\}.$$

Restricting the polynomials to those of the form $f = \mathbf{x}^\mathsf{T} A\mathbf{x}$ for $A \in \mathbb{S}^n$, combined with Theorem 2.1 of [AM14] the authors propose the following hierarchies of cones approximating $\mathcal{COP}^n$:

$$\mathrm{DD}_n^r := \Big\{ A \in \mathbb{S}^n \mid \big(\sum_{i=1}^{n} x_i^2\big)^r (\mathbf{x}\circ\mathbf{x})^\mathsf{T} A(\mathbf{x}\circ\mathbf{x}) = (\mathbf{x}^{[r+1]})^\mathsf{T} B(\mathbf{x}^{[r+1]}),\ B \text{ is dd}\Big\}, \tag{1.19}$$

and

$$\mathrm{SDD}_n^r := \Big\{ A \in \mathbb{S}^n \mid \big(\sum_{i=1}^{n} x_i^2\big)^r (\mathbf{x}\circ\mathbf{x})^\mathsf{T} A(\mathbf{x}\circ\mathbf{x}) = (\mathbf{x}^{[r+1]})^\mathsf{T} B(\mathbf{x}^{[r+1]}),\ B \text{ is sdd}\Big\}, \tag{1.20}$$

where we recall that $(x^{[r+1]})$ is the vector containing all monomials of degree $r+1$. Note the relationship of (1.19) and (1.20) with Theorem 1.11 and Theorem 1.14, making the concept that these cones are a relaxation of the sum of squares conditions, as presented above, explicit. Moreover from this observation we immediately get that

$$\mathrm{DD}_n^r \subseteq \mathrm{SDD}_n^r \subseteq \mathcal{K}_n^r$$

From [AM14, Theorem 2.2] we know that DD_n^r is polyhedral while SDD_n^r has a second order cone representation. Comparing the cones (1.19) and (1.20) to the alternative formulation (1.18) of $\mathcal{K}_n^r$, we furthermore see that apart from having an easier matrix constraint (i.e. diagonally dominant and scaled diagonally dominant versus positive semidefinite) the number of matrix variables is significantly smaller as well. Then, as with the other hierarchies it can easily be seen that $\mathrm{DD}_n^r \subseteq \mathrm{DD}_n^{r+1}$ and $\mathrm{SDD}_n^r \subseteq \mathrm{SDD}_n^{r+1}$ for all $r \in \mathbb{Z}_+$.

1.3.6 Other approximations for $\mathcal{COP}^n$ and $\mathcal{CP}^n$

A generalization that contains the hierarchies $\mathcal{C}_n^r$, $\mathcal{Q}_n^r$, and $\mathcal{K}_n^r$ has been suggested by Dobre and Vera [DV13]. They define the hierarchy of cones

$$\mathbb{H}_{k,n}^r := \left\{ A \in \mathbb{S}^n \mid \Big(\sum_{i=1}^n x_i\Big)^r \mathbf{x}^\mathsf{T} A \mathbf{x} = \sum_{j=r+2-k}^{r+2} \sum_{\boldsymbol{\beta}\in\mathbb{N}^n, |\boldsymbol{\beta}|=j} \mathbf{x}^{\boldsymbol{\beta}} \sigma_{\boldsymbol{\beta}}(\mathbf{x}),\ \sigma_{\boldsymbol{\beta}}(\mathbf{x}) \in \Sigma_{r+2-|\boldsymbol{\beta}|} \right\}$$

where for any $d \in \mathbb{N}$, Σ_d denotes the set of sum of squares polynomials of degree d. Given $n \in \mathbb{N}$, for any nondecreasing sequence k_r, $r = 0, 1, 2, \ldots$ they obtain

$$\mathbb{H}_{k_r,n}^r \subseteq \mathbb{H}_{k_{r+1},n}^{r+1} \subseteq \mathcal{COP}^n$$

as well as $\mathrm{cl}(\bigcup_{r\in\mathbb{Z}_+} \mathbb{H}_{k_r,n}^r) = \mathcal{COP}^n$. Moreover they get that $\mathcal{C}_n^r = \mathbb{H}_{0,n}^r$ and $\mathcal{Q}_n^r = \mathbb{H}_{2,n}^r$ while $\mathcal{K}_n^r = \mathbb{H}_{k,n}^r$ for any $k \geq r+2$.

Another method to construct approximations for the copositive cone strongly related to the concept of sum of squares is that based on the theory of moments. Although most of this approximation theory is set up in a more general framework of possibly restricted nonnegative polynomials over some set K, it can be tailored to the copositive cone. That is, we can reduce these techniques to the copositive setting by restricting the polynomials to have degree at most two and by taking K as the nonnegative orthant. One of the most well known results in this area is the so called Lasserre hierarchy which was introduced in its general form in [Las11]. Approximation hierarchies, specifically for the copositive and completely positive cone, were then presented in [Las14]. These result were later extended by Dickinson [Dic13, Corollary 11.3]. In fact, this extension explicitly shows the link between the theory of moments and sum of squares. Another way this link becomes apparent is via the dual of the cones $\mathcal{K}_n^r$, see e.g. [Dic13, Section 11.6]. As the theory of moments is outside of the scope of this text we refer the interested reader to [GL07] and [Dic13, Chapter 10] for an introduction to this topic.

Finally, two more hierarchies of outer approximations of the completely positive cone were suggested by Dong [Don13] and are constructed using symmetric tensors. A *tensor* of order r and dimension n is an object for which each entry is sepcified using r indices each of which takes values in $\{1, \ldots, n\}$. Let $\sigma : \{1, \ldots, r\} \to \{1, \ldots, r\}$ be a permutation. Then T is a *symmetric tensor* of order r if $T(\sigma(i_1), \ldots, \sigma(i_r)) = T(i_1, \ldots, i_r)$ for all permutations σ. Observe that a symmetric tensor for $r = 2$ is simply a symmetric matrix. Denoting the set of symmetric tensors of order r and dimension n by $\mathbb{S}_n^r$, Dong furthermore defines the operators **Slices**$(\bullet)$ and **Collapse**$(\bullet)$. In particular **Slices**(Z) is the set of $n \times n$ symmetric matrices that can be produced from $Z \in \mathbb{S}_n^{r+2}$ by fixing exactly r indices of Z. Then **Collapse**(Z) is the sum of all such matrices produced by **Slices**(Z). The approximation hierarchies are then obtained by first rewriting the completely positive cone as

$$\mathcal{CP}^n = \{A \in \mathbb{S}^n \mid \exists Z \in \mathbb{S}_n^{r+2}, \mathbf{Slices}(Z) \subseteq \mathcal{CP}^n, A = \mathbf{Collapse}(Z)\}.$$

The condition **Slices**$(Z) \subseteq \mathcal{CP}^n$ in the equation above is then relaxed to the weaker conditions **Slices**$(Z) \subseteq \mathcal{N}^n$ and **Slices**$(Z) \subseteq \mathcal{S}_+^n \cap \mathcal{N}^n$ respectively. Doing so produces the approximation hierarchies

$$\mathcal{T}_n^r = \{A \in \mathbb{S}^n \mid \exists Z \in \mathbb{S}_n^{r+2}, \mathbf{Slices}(Z) \subseteq \mathcal{N}^n, A = \mathbf{Collapse}(Z)\}, \tag{1.21}$$

and

$$\mathcal{TD}_n^r = \{A \in \mathbb{S}^n \mid \exists Z \in \mathbb{S}_n^{r+2}, \mathbf{Slices}(Z) \subseteq \mathcal{S}_+^n \cap \mathcal{N}^n, A = \mathbf{Collapse}(Z)\}. \tag{1.22}$$

Furthermore, it is shown in [Don13] that the cones $\mathcal{T}_n^r$ and $\mathcal{C}_n^r$ are mutually dual, as are the cones $\mathcal{TD}_n^r$ and $\mathcal{Q}_n^r$. A result following directly from these duality statements combined with Theorem 1.13 is that the cones $\mathcal{T}_n^r$ and $\mathcal{TD}_n^r$ converge to $\mathcal{CP}^n$, i.e. $\bigcap_{r \in \mathbb{Z}_+} \mathcal{T}_n^r = \mathcal{CP}^n$ as well as $\bigcap_{r \in \mathbb{Z}_+} \mathcal{TD}_n^r = \mathcal{CP}^n$.
Finally, in light of all these approximation hierarchies we introduce the lifting rank of a matrix $A \in \mathbb{S}^n$ that we will particularly make use of in Chapter 3 in an effort to shorten and simplify notation.

Definition 1.15. Let K be some cone in $\mathbb{R}^{n \times n}$. For any matrix $A \in K$ and any hierarchy of sets $(\mathbb{Y}_n^r)_{r \in \mathbb{Z}_+}$ approximating K we define the *lifting rank* of A as,

$$r^*_{\mathbb{Y}_n^r}(A) = \min\{r \in \mathbb{Z}_+ \mid A \in \mathbb{Y}_n^r\}. \tag{1.23}$$

When there does not exist an r for which $A \in \mathbb{Y}_n^r$ we define $r^*_{\mathbb{Y}_n^r}(A) = \infty$.

As an example, for any matrix $A \in \mathcal{COP}^4$ we have $r^*_{\mathcal{K}_4^r}(A) = 0$.

1.4 Copositive (Completely Positive) Programming

When we speak of a copositive, or alternatively a completely positive program, we are referring to a linear program over either the copositive or completely positive cone respectively. That is, a *copositive program* is of the form

$$\begin{aligned} &\min \langle C, X\rangle \\ &\text{s.t.}\ \langle A_i, X\rangle = b_i \quad i = 1, \dots, m \\ &\quad X \in \mathcal{COP}^n, \end{aligned} \tag{1.24}$$

where the dual problem, a *completely positive program*, is

$$\begin{aligned} &\max \mathbf{b}^{\mathsf{T}}\mathbf{y} \\ &\text{s.t.}\ C - \sum_{i=1}^{m} y_i A_i \in \mathcal{CP}^n,\ \mathbf{y} \in \mathbb{R}^n, \end{aligned} \tag{1.25}$$

for given matrices A_i, $C \in \mathbb{R}^{n\times n}$ and vector $\mathbf{b} \in \mathbb{R}^n$. Note that both of these programs are convex due to the fact that both $\mathcal{COP}^n$ and $\mathcal{CP}^n$ are convex cones. It was then shown by Eichfelder and Jahn [EJ08] that if Slater's condition is satisfied, the Karush-Kuhn-Tucker optimality conditions hold, and moreover we get strong duality as long as both problems are feasible and at least one of them is strictly feasible, i.e. allows for a feasible solution in the interior of either the copositive or completely positive cone. At the moment of writing no exact algorithms exist to solve copositive or completely positive programs to optimality. We can however solve strictly feasible copositive programs up to any desired accuracy using the algorithm introduced by Bundfuss and Dür [BD09]. This algorithm extends the idea presented in [BD08] that we described earlier in this chapter, see Section 1.3.1. In particular the authors propose an adaptive approach where moreover the partitioning of the simplex no longer happens uniformly, but is instead guided in an adaptive manner via the objective function. A heuristic approach that approximates completely positive problems from within via a descent method was introduced by Bomze et al. [BJR11].

Another way of dealing with copositive programs is via approximations, for example via any of the hierarchies that were mentioned in Section 1.3. In particular, substituting $\mathcal{COP}^n$ in (1.24) by any subset of $\mathcal{COP}^n$ one immediately obtains upper bounds to this minimization problem. The dual of any such subset of $\mathcal{COP}^n$ furthermore provides us with an outer approximation to the completely positive cone, allowing us to obtain upper bounds for the completely positive maximization problem (1.25) as well. Similarly, lower bounds for (1.24) and upper bounds for (1.25) can be obtained via outer approximations of $\mathcal{COP}^n$ and inner approximations of $\mathcal{CP}^n$ respectively.

As mentioned before, the motivation for copositive programming comes from the fact that a number of NP-hard problems can be written as copositive or completely positive programs. Combined with the approximation techniques described in Section 1.3, a number of new bounds and other interesting results have been obtained in recent years. The remainder of this section will be dedicated to reviewing several such results.

One of the most straightforward examples comes from a paper by Bomze et al. [BDdK$^+$00], where the authors rewrite the standard quadratic problem as a completely positive program. The standard quadratic problem is defined as

$$\min\{\mathbf{x}^\mathsf{T} Q\mathbf{x} \mid \mathbf{e}^\mathsf{T}\mathbf{x} = 1,\, \mathbf{x} \in \mathbb{R}^n_+\}. \tag{1.26}$$

By writing $\mathbf{x}^\mathsf{T} Q\mathbf{x} = \langle Q, \mathbf{x}\mathbf{x}^\mathsf{T}\rangle$ one immediately sees from the definition of $\mathcal{CP}^n$ that $\mathbf{x}\mathbf{x}^\mathsf{T}$ is a completely positive rank 1 matrix. Hence one can define the following completely positive relaxation of (1.26)

$$\min\{\langle Q, X\rangle \mid \langle X, E\rangle = 1,\, X \in \mathcal{CP}^n\}. \tag{1.27}$$

Note that (1.27) is a linear program over the cone $\mathcal{CP}^n$. In particular, the linearity of the objective function implies that the optimal solution to (1.27) is attained at an extreme point of $\mathcal{CP}^n$. As mentioned before, see (1.9), the extreme rays of the completely positive cone are exactly its rank 1 matrices, and so the completely positive formulation (1.27) is in fact exact. Bomze and de Klerk [BdK02] later used this completely positive formulation, together with the approximation hierarchies $\mathcal{C}^r_n$ from (1.12) and $\mathcal{K}^r_n$ from (1.14), to construct a polynomial time approximation scheme for the standard quadratic problem.

One of the most well known results in copositive programming is Burer's result concerning binary quadratic optimization from [Bur09]. In this paper it was shown that the binary quadratic program

$$\begin{aligned} \min\, & \mathbf{x}^\mathsf{T} Q\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.}\, & \mathbf{a}_i^\mathsf{T}\mathbf{x} = b_i \qquad (i = 1, \ldots, m) \\ & \mathbf{x} \geq 0 \\ & x_j \in \{0, 1\} \qquad\quad (j \in B) \end{aligned} \tag{1.28}$$

can be written equivalently as

$$
\begin{aligned}
\min\ & \langle Q, X\rangle + \mathbf{c}^{\mathsf{T}}\mathbf{x} \\
\text{s.t.}\ & \mathbf{a}_i^{\mathsf{T}}\mathbf{x} = b_i && (i = 1, \dots, m) \\
& \mathbf{a}_i^{\mathsf{T}} X \mathbf{a}_i = b_i^2 && (i = 1, \dots, m) \\
& x_j = X_{jj} && (j \in B) \\
& \begin{bmatrix} 1 & \mathbf{x}^{\mathsf{T}} \\ \mathbf{x} & X \end{bmatrix} \in \mathcal{CP}^{n+1}
\end{aligned}
\tag{1.29}
$$

under the assumption that the conditions $\mathbf{x} \geq 0$ and $\mathbf{a}_i^{\mathsf{T}}\mathbf{x} = b_i$, $i = 1, \dots, m$, imply that $x_j \leq 1$ for all $j \in B$. This condition is referred to by Burer as *the key condition* and can, without loss of generality, always be achieved by introducing slack variables s_i and adding the constraints $x_i + s_i = 1$, $i = 1, \dots, n$ to the completely positive program (1.29). Moreover, Burer showed that under the additional assumption that there exists a $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{h} := \sum_{i=1}^m y_i a_i \geq 0$ and $\sum_{i=1}^m y_i b_i = 1$, one can eliminate the variables $\mathbf{x}$ from (1.29) altogether. These extra assumptions, in other words, reduce (1.29) from an optimization problem over $\mathcal{CP}^{n+1}$ to an optimization problem over $\mathcal{CP}^n$, that is (1.29) is equivalent to

$$
\begin{aligned}
\min\ & \langle Q, X\rangle + \mathbf{c}^{\mathsf{T}} X\mathbf{h} \\
\text{s.t.}\ & \mathbf{a}_i^{\mathsf{T}} X\mathbf{h} = b_i && (i = 1, \dots, m) \\
& \mathbf{a}_i^{\mathsf{T}} X \mathbf{a}_i = b_i^2 && (i = 1, \dots, m) \\
& [X\mathbf{h}]_j = X_{jj} && (j \in B) \\
& X \in \mathcal{CP}^n \\
& \mathbf{h}^{\mathsf{T}} X \mathbf{h} = 1.
\end{aligned}
$$

Another interesting note on the feasible set of the completely positive program (1.29) was published by Bomze and Jarre [BJ10]. In particular they provide a direct way to prove equivalence of (1.28) and (1.29) which is significantly shorter than that in [Bur09]. The linear case, that is that of mixed binary linear programs, was treated by Natarajan et al [NTZ11] in a robust optimization setting. For this type of robust optimization problem the authors managed to obtain a completely positive formulation, and moreover provide an implicit connection between their formulation and the theory of moments.

Another well known result for copositive programming is that regarding the stability number. Given a graph $G = (V, E)$, the *stability number* is the maximum cardinality of a subset of vertices such that no two of them are connected by an edge. A copositive formulation for this problem, containing only one variable and one matrix constraint, was proposed by de Klerk and

Pasechnick [dKP02]. Let α_G denote the stability number of a graph $G = (V, E)$, and let A_G be its adjacency matrix, then the authors show that

$$\alpha_G = \min_{\lambda \in \mathbb{R}} \{\lambda \mid \lambda(I + A_G) - E \in \mathcal{COP}^n\}. \tag{1.30}$$

Replacing $\mathcal{COP}^n$ with $\mathcal{C}_n^r$ they then obtain a hierarchy of polyhedral approximations for the stability number, which they denote by $\zeta^{(r)}(G)$, $r \in \mathbb{Z}_+$. The authors show that $\lfloor \zeta^{(\alpha_G^2)} \rfloor = \alpha_G$. Moreover, it was conjectured by De Klerk and Pasechnik that this polyhedral approximation hierarchy would be able to find the stability number for $r = \alpha_G$, i.e. $r^*_{\mathcal{C}_n^r}(\alpha_G(I + A_G) - E) = \alpha_G$. This conjecture was shown to be true by Laurent and Gvozdenović [GL07] for graphs with stability number at most 8. The dual of the copositive formulation (1.30) is given by

$$\alpha_G = \max\{\langle E, X\rangle \mid \langle A_G + I, X\rangle = 1,\ X \in \mathcal{CP}^n\}. \tag{1.31}$$

By replacing the completely positive constraint $X \in \mathcal{CP}^n$ in (1.31) with the doubly nonnegative constraint $X \in \mathcal{S}_+^n \cap \mathcal{N}^n$, we obtain the famous upper bound known as the Lovász-Schrijver ϑ' number. A copositive formulation for the stability number of infinite graphs was suggested by Dobre et al. [DDV14].

Closely related to the stability number is the so called chromatic number of a graph. The *chromatic number*, χ_G, for a graph G is defined as the smallest number of colors needed to color the vertices of G such that any two vertices with the same color are not connected by an edge. The relation with the stability number comes from the fact that in any coloring of a graph each color defines a stable set. A copositive formulation for this problem was found by Gvozdenović and Laurent [GL08]. For a graph G, let G_k be the graph obtained by taking the Cartesian product of the complete graph on k vertices and G. Then the following is a copositive formulation of the chromatic number:

$$\chi_G = \max_{y,z \in \mathbb{R}} \{y \mid \tfrac{1}{n^2}(k - y)E_{nk} + z(n(I_{nk} + A_{G_k}) - E_{nk}) \in \mathcal{COP}^{nk},\ k = 1, \ldots, n\}.$$

Improvements on the Lovász ϑ number in the direction of the chromatic number were obtained by Dukanovic and Rendl [DR10]. Given a graph $G = (V, E)$ with $|E| = m$, the authors define for $\mathbf{y} \in \mathbb{R}^{|E|}$ the operator $A_G^\mathsf{T}\mathbf{y} := \sum_{\{i,j\} \in E} y_{ij}(\mathbf{e}_i\mathbf{e}_j^\mathsf{T} + \mathbf{e}_j\mathbf{e}_i^\mathsf{T})$, where $\mathbf{e}_i$ denotes the i^{th} unit vector. Then they present the following completely positive program

$$\min_{z \in \mathbb{R}_+} \{z \mid zI_m + A_G^\mathsf{T}\mathbf{y} - E_m \in \mathcal{S}_+^m,\ zI_m + A_G^\mathsf{T}\mathbf{y} \in \mathcal{CP}^m\}.$$

The authors furthermore show that the optimal value of this completely positive formulation is equal to the so called fractional chromatic number. The *fractional chromatic number* is similar to the chromatic number, except that vertices are assigned sets of colors rather than a single color. Moreover the authors use the Parrilo cones $\mathcal{K}_n^r$ from (1.14) to obtain approximations, which they then simplify for vertex transitive graphs. They also show that the Lovász ϑ number can be improved upon significantly for specific Hamming graphs.

The quadratic assignment problem was reformulated as a copositive program by Povh and Rendl [PR09]. Let $A, B \in \mathbb{S}^n$ and $C \in \mathbb{R}^{n\times n}$. Then the quadratic assignment problem is defined as

$$\min\Big\{\langle X, AXB + C\rangle \mid X \in \mathcal{P}_n\Big\}. \tag{1.32}$$

Next, Povh and Rendl define $c := \mathrm{Vec}(C)$ and show that (1.32) can be formulated as the following copositive program

$$\max_{S,T\in\mathbb{S}^n, v\in\mathbb{R}}\Big\{\mathrm{Tr}(S) + \mathrm{Tr}(T) + n^2 v \mid A\otimes B + \mathrm{Diag}(c) - I\otimes S - T\otimes I - vE_{n^2} \in \mathcal{COP}^{n^2}\Big\}.$$

This is done by adding redundant constraints, and with the help of Lagrangian duality. A more detailed description of this method is given in Chapter 5, where we will use this technique to find a copositive formulation for the graph isomorphism problem.

Another problem studied by Povh and Rendl [PR07] is the so called graph partitioning problem. The *graph partitioning problem* is the problem of partitioning the vertices of an edge weighted graph, $G_w = (V, E, W)$, into sets V_1,V_2,V_3 of prescribed cardinalities $m_1, m_2, m_3 \in \mathbb{Z}_+$ respectively, such that the total weight of all edges between V_1 and V_2 is minimized. For their completely positive formulation of this problem the authors define matrices $E_n^{ij} := \mathbf{e}_i\mathbf{e}_j^\mathsf{T}$. Furthermore, let A_{G_w} be the weighted adjacency matrix of the weighted graph G_w. Then for any such graph G_w the graph partitioning problem can equivalently be written as

$$
\begin{aligned}
\min\ & \tfrac{1}{2}\Big\langle E_3^{12} + E_3^{21} \otimes A_{G_w}, Y \Big\rangle & \\
\text{s.t. }\ & \tfrac{1}{2}\Big\langle \big(E_3^{ij} + E_3^{ji}\big) \otimes I_n, Y \Big\rangle = m_i \delta_{ij} & 1 \le i \le j \le 3 \\
& \Big\langle E_3 \otimes E_n^{ii} \Big\rangle = 1 & 1 \le i \le n \\
& \Big\langle \Big(\sum_{j=1}^{3} E_3^{ij} \Big) \otimes \Big(\sum_{i=1}^{n} E_n^{ij} \Big), Y \Big\rangle = m_i & 1 \le i \le 3,\ 1 \le j \le n \\
& \tfrac{1}{2}\Big\langle \big(E_3^{ij} + E_3^{ji}\big) \otimes E_n, Y \Big\rangle = m_i m_j & 1 \le i \le j \le n \\
& Y \in \mathcal{CP}^{3n}.
\end{aligned}
\tag{1.33}
$$

The completely positive formulation (1.33) is then relaxed to obtain a semidefinite program, whose optimal value turns out to be equal to the spectral bound that was found by Helmberg et al. [HRMP95]. Contrary to that bound however, the semidefinite formulation obtained by Povh and Rendl allows for a natural way to add extra constraints without losing tractability of the problem. The authors use this observation to obtain a tightened version of the semidefinite program and hence the spectral bound for the graph partitioning problem. For small instances the authors show that this tightened version can in fact give significant improvements over the spectral bound.

For other surveys on the field of copositive programming we refer the reader to surveys by Dür (2010) [Dür10], Burer (2012) [Bur12], Bomze (2012) [Bom12], Dickinson (2013) [Dic13], and Bomze, Schachinger and Uchida (2012) [BSU12].