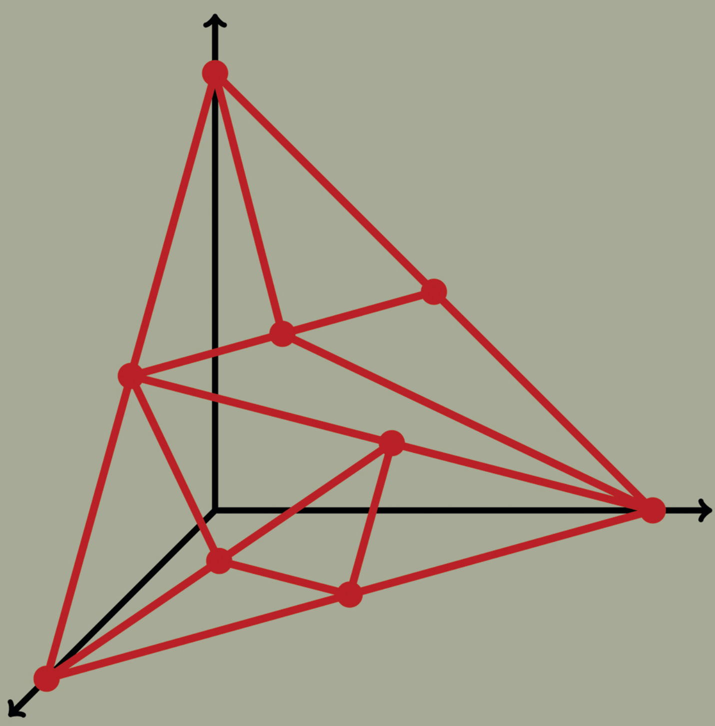
# COPOSITIVE PROGRAMMING and RELATED PROBLEMS



Faizan Ahmed

# COPOSITIVE PROGRAMMING

# and

# RELATED PROBLEMS

**Faizan Ahmed**

Graduation Committee:

| | |
|---|---|
| Prof.dr.ir. B.J. Geurts (Chairman) | University of Twente |
| Prof. Dr. Marc Uetz (Promotor) | University of Twente |
| Dr. Georg Still (Assistant Promotor) | University of Twente |
| Prof. Dr. Mirjam Duer | University of Trier, Germany |
| Prof. Dr. Frank Vallentin | University of Cologne, Germany |
| Prof. Dr. Anton A. Stoorvogel | University of Twente |
| Dr. Walter Kern | University of Twente |

The research presented in this dissertation was (partially) funded by Higher Education Commission of Pakistan under the project `` Overseas Scholarship Program for MS/MPhil Leading to PhD, Phase-II Batch-II'' and carried out at the group of Discrete Mathematics and Mathematical Programming (DMMP), Department of Applied Mathematics, Faculty of Electrical Engineering, Mathematics and Computer Science of the University of Twente, the Netherlands.

CTIT Ph.D. Thesis Series No. 14-312 Centre for Telematics and Information Technology P.O. Box 217, 7500 AE Enschede, The Netherlands.

Typeset with LaTeX
Printed by CPI-Wöhrman Print Service-Zutphen

Cover Designed by: Faizan Ahmed, Noor-Ul-Sabah and Zaharah Allah Bukhsh



**ISBN:** 978-90-365-3672-1
**ISSN:** 1381-3617 (CTIT Ph.D. thesis Series No. 14-312)
http://dx.doi.org/10.3990/1.9789036536721

# COPOSITIVE PROGRAMMING
# and
# RELATED PROBLEMS

**DISSERTATION**

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on wednesday 28, May 2014 at 1645 hours

by

**Faizan Ahmed**

Born on March 4, 1984
in Rawalpindi, Pakistan

This dissertation has been approved by:

| | |
|---|---|
| Prof. Dr. Marc Uetz | Promotor |
| Dr. Georg Still | Assistant Promotor |

# Contents

# 1

# Preliminaries

FOUNDATION of mathematical optimization relies on the urge to utilize available resources to their optimum. This leads to mathematical programs where an objective function is optimized over a set of constraints. The set of constraints can represent different structures, for example, a polyhedron, a box or a cone. Mathematical programs with cone constraints are called cone programs. A sub area of mathematical optimization is the one where the number of variables is finite while the number of constraints is infinite, known as semi-infinite programming. In this chapter we will start with a general introduction into the thesis. In the second section some basic definitions are given which are used throughout the thesis. The third and the fourth section provide a brief review of results on cone programming and semi-infinite programming, respectively. In section five we will briefly discuss cone programming relaxations. In the last section we shall give an overview over results presented in the thesis.

## 1.1 Introduction

In mathematical optimization an *objective function* is required to be optimized over a set of side conditions called *constraints*. More precisely, mathematical optimization, refers to the following problem:

$$\max \quad f(\mathbf{x}) \quad s.t. \quad g_j(\mathbf{x}) \leq 0, \quad j \in J, \mathbf{x} \in S$$

where $S \subseteq \mathbb{R}^n$, $J$ an index set (possibly infinite) and $f : \mathbb{R}^n \to \mathbb{R}$, $g_j : \mathbb{R}^n \to \mathbb{R}$. The function $f$ is called objective function while the functions $g_j$ represent constraints. A point $\mathbf{x} \in S$ is called feasible, if it satisfies all constraints $g_j(\mathbf{x}) \leq 0$, $j \in J$. The optimization problem is called feasible if there exists at least one point $\mathbf{x} \in S$ satisfying all constraints.

If a point $\overline{\mathbf{x}} \in S$ satisfies all constraints and the value of the objective function, $f(\overline{\mathbf{x}})$, is optimal, then this point, $\overline{\mathbf{x}}$, is called a *solution*. An optimization problem can have more than one solution, or no solution at all.

Mathematical programming emerged as an independent area of mathematics in the second half of the previous century. Its root can be traced back to the work of ancient Chinese mathematicians, to the work of Euler, Leibniz, Lagrange and Newton (for a history of optimization see [77]). Mathematical optimization is a rich field of mathematics with numerous applications. In order to give a flavour of applicability of mathematical optimization to real world problems, we quote: `` In many of their approaches to understand nature, physicists, chemists, biologists, and others assume that the systems they try to comprehend tend to reach a state that is characterized by the optimality of some function'' [77] and ``To make decisions optimally is a basic human desire. Whenever the situation and the objectives can be described quantitatively, this desire can be satisfied, to some extent, by using mathematical tools, specifically those provided by optimization theory and algorithms'' [11].

Mathematical optimization is a vast area of mathematics. It can be classified in various ways. A fundamental classification is linear optimization and nonlinear optimization. Nonlinear optimization contains both ``hard'' and ``easy'' problems. Nonlinear optimization can be further classified as convex optimization and non-convex optimization. A sub-area of mathematical optimization is the one where the number of variables are finite while the number of constraints are infinite, known as semi-infinite programming.

In mathematical optimization the constraint set may represent a geometrical structure. If the variables are restricted to take values from a so-called cone, then we have a cone program. Cone programming not only contains convex programming as a special case, but some nonconvex optimization problems can also be reformulated as a cone program.

Cone programs over the copositive cone or the completely positive cone are referred to as copositive programming. In the last decade copositive programming has caught much attention due to the fact that many hard optimization problems can be exactly reformulated as a copositive program. In this thesis we shall deal with copositive programming and problems related to copositive programming. As we shall see, feasibility in copositive programming amounts to solving a so-called standard quadratic optimization problem. Optimality conditions and solution methods for copositive programming are also discussed from a viewpoint of linear semi-infinite programming. We will also look at the sharpness of copositive programming relaxations of quadratically constrained quadratic programs.

## 1.2 Basic Definitions

In this section we will give the basic notations and definitions used throughout the thesis. Following the usual convention the set of all real numbers will be denoted by $\mathbb{R}$ while $\mathbb{R}_+$ denotes the set of all nonnegative real numbers. Similarly for given positive integers $m, n$, $\mathbb{R}^m$ and $\mathbb{R}^{m\times n}$ denote the set of all real vectors of size $m$ and the set of all $m \times n$ real matrices, respectively. Moreover, the vectors will be denoted by bold small letters while the elements of the vectors will be denoted by small letters with subscripts. For example, the $i^{th}$ element of $\mathbf{v} \in \mathbb{R}^m$ will be written as $v_i$. For the complete list of notation the interested reader is referred to the List of Notations given at page 121.

This thesis is mainly concerned with cone programming or specifically copositive programming. First we will define what is meant by a convex set and a convex cone,

**Definition 1.1** (Convex Set). A set $S \subseteq \mathbb{R}^m$ is called convex if for each $\mathbf{v}, \mathbf{u} \in S$ and $0 \leq \lambda \leq 1$ we have $\lambda\mathbf{v} + (1-\lambda)\mathbf{u} \in S$.

**Definition 1.2** (Convex Hull). Let $S \subset \mathbb{R}^m$ be an arbitrary set. The set,

$$\text{conv}\,(S) := \left\{ \mathbf{v} : \mathbf{v} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i, \mathbf{v}_i \in S; \lambda_i \geq 0 \text{ for } i = 1, \cdots, n, \sum_{i=1}^{n} \lambda_i = 1, n \geq 1 \right\}$$

is called the *convex hull* of $S$.

**Definition 1.3** (Convex Cone). A set $K \subseteq \mathbb{R}^{m \times n}$ which is closed under nonnegative multiplication and addition, i.e., $U, V \in K \Rightarrow \lambda\,(U + V) \in K$ for all $\lambda \geq 0$, is called a *convex cone*. A cone is *pointed* if $K \cap -K = \{0\}$. The *dual* of a cone $K$ is defined as:

$$K^* = \left\{ U \in \mathbb{R}^{m \times n} : \langle U, V \rangle \geq 0, \;\; \forall\, V \in K \right\}$$

where $\langle ., . \rangle$ stands for the standard inner product, i.e.,

$$\langle U, V \rangle = \text{tr}(U^T V) = \sum_{i,j} u_{ij} v_{ij} \text{ for } U, V \in \mathbb{R}^{m \times n},$$

with $u_{ij}$ denoting the $ij^{th}$ element of the matrix $U$.

In the above definition tr denotes the trace of the matrix and $U^T$ denotes the transpose of $U$.

There are three special cases of convex cones which are important with respect to the material presented in this thesis. These cones are formed by certain subsets of symmetric matrices. We will define these matrices and the associated cones. In the definitions below and throughout the thesis $\mathcal{S}_m$ denotes the cone of all symmetric $m \times m$ matrices.

**Definition 1.4** (Positive Semidefinite Matrix). A matrix $Q \in \mathcal{S}_m$ is called positive semidefinite if $\mathbf{v}^T Q \mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^m$. The set of all $m \times m$ positive semidefinite matrices defines a cone called the *positive semidefinite cone*. We will denote this cone by $\mathcal{S}_m^+$.
Similarly, $Q \in \mathcal{S}_m$ is called positive definite if $Q \in \mathcal{S}_m^+$ and $\mathbf{v}^T Q \mathbf{v} = 0$ holds if and only if $\mathbf{v} = \mathbf{o}$, where $\mathbf{o} \in \mathbb{R}^m$ is the zero vector. The set of all positive definite matrices is denoted by $\mathcal{S}_m^{++}$.

**Definition 1.5** (Copositive Matrix). A matrix $Q \in \mathcal{S}_m$ is called copositive if $\mathbf{v}^T Q \mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}_+^m$. The set of all $m \times m$ copositive matrices defines a cone called the *copositive cone*. We will denote this cone by $\mathcal{C}_m$.

**Definition 1.6** (Completely Positive Matrix)**.** A matrix $Q \in \mathcal{S}_m$ is called completely positive if there exist a matrix $B \in \mathbb{R}_+^{m\times n}$, for some $n \in \mathbb{N}$, such that $Q = BB^T$. The set of all $m \times m$ completely positive matrices defines a cone called the *completely positive cone*. We will denote this cone by $\mathcal{C}_m^*$.

## 1.3 Cone Programming

In this section we will briefly discuss some results on cone programming. Cone programming is an important class of mathematical programming. Cone programming refers to the following pair of primal dual programs,

$$(Cone_P) \qquad \max_{\mathbf{x}\in\mathbb{R}^n} \quad \mathbf{c}^T\mathbf{x} \quad s.t. \qquad B - \sum_{i=1}^{n} x_i A_i \in \mathcal{K}$$

where $A_i, B \in \mathcal{S}_m$, $\mathbf{c} \in \mathbb{R}^n$ and $\mathcal{K}$ is a given cone of symmetric $m \times m$ matrices. The dual of the above program can be written as follows:

$$(Cone_D) \qquad \min \quad \langle B, Y\rangle \quad s.t. \quad \langle A_i, Y\rangle = c_i, \quad \forall\, i = 1, ..., n, \quad Y \in \mathcal{K}^*$$

In mathematical programming, duality theory plays a crucial role in formulating optimality conditions and devising solution algorithms. Duality theory can be further classified into two categories: weak duality and strong duality. In weak duality we investigate, if the optimal value of the primal problem is upper bounded by the value of the dual problem. Strong duality investigates the conditions under which equality holds for optimal values of the primal problem and the dual problem. $(Cone_P)$ and $(Cone_D)$ satisfy weak duality.

**Lemma 1.7** (Weak Duality)**.** *Let* $\mathbf{x}$ *and* $Y$ *be feasible solutions for* $(Cone_P)$ *and* $(Cone_D)$ *respectively, then* $\mathbf{c}^T\mathbf{x} \leq \langle Y, B\rangle$*.*

*Proof.* We have

$$\begin{aligned}
\mathbf{c}^T\mathbf{x} &= \sum_{i=1}^{n} c_i x_i = \sum_{i=1}^{n} x_i\langle A_i, Y\rangle = \sum_{i=1}^{n} \langle x_i A_i, Y\rangle = \left\langle \sum_{i=1}^{n} x_i A_i, Y \right\rangle \\
&= \langle B, Y\rangle - \left\langle B - \sum_{i=1}^{n} x_i A_i, Y \right\rangle \\
&\leq \langle B, Y\rangle
\end{aligned}$$

□

In the case of linear programming, i.e. the case, when $\mathcal{K} = \mathcal{N}_m$, where $\mathcal{N}_m$ denotes the cone of all $m \times m$ symmetric nonnegative matrices, then, whenever $(Cone_P)$ or $(Cone_D)$ are feasible, we have equality in the optimal values, i.e., we have a zero duality gap. Moreover, if both $(Cone_P)$ and $(Cone_D)$ are feasible then both optimal values are attained. Strong duality does not hold for cone programming, in general. In the example below and throughout the thesis for a mathematical program $(P)$, $val(P)$ and $\mathcal{F}(P)$ will denote the value and the set of feasible points for the program $(P)$.

**Example 1.8** (Strong Duality May Fail). *Consider,*

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \; \mathbf{c} = \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

*then $(Cone_P)$ and $(Cone_D)$ takes the following form,*

$$(Cone_P) \quad \max_{\mathbf{x} \in \mathbb{R}^2} \quad 2x_2 \quad s.t. \quad \begin{pmatrix} -x_1 & -x_2 & 0 \\ -x_2 & 0 & 0 \\ 0 & 0 & 1-2x_2 \end{pmatrix} \in \mathcal{K}$$

$$(Cone_D) \quad \min \; y_{33} \quad s.t. \quad y_{11} = 0, y_{12} + y_{33} = 1, \; Y := \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{12} & y_{22} & y_{23} \\ y_{13} & y_{23} & y_{33} \end{pmatrix} \in \mathcal{K}^*$$

*It is clear that for the case $\mathcal{K} = \mathcal{K}^* = \mathcal{N}_3$ we have,*

$$val(Cone_P) = val(Cone_D) = 0.$$

*It is not difficult to verify that for the case $\mathcal{K} = \mathcal{K}^* = \mathcal{S}_3^+$ we have $val(Cone_P) = 0$ and $val(Cone_D) = 1$ even though both problems are feasible.*

*For the case $\mathcal{K} = \mathcal{C}_3$ we have (cf. Lemma 2.10),*

$$\mathcal{F}(Cone_P) = \{\mathbf{x} \in \mathbb{R}^3 : x_1 \leq 0, x_2 \leq 0, x_1(2x_2 - 1) \geq 0, 1 - 2x_2 \geq 0\}$$

*From this we get $val(Cone_P) = 0$. Now take $\mathcal{K}^* = \mathcal{C}_3^*$ and note that the necessary and sufficient condition for $Y \in \mathcal{C}_3^*$ is that $Y \in \mathcal{S}_3^+ \cap \mathcal{N}_m$ (see (2.8) on page 31).*

*Then we get $val(Cone_D) = 1$ attained by*

$$\mathcal{C}_3^* \ni Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}.$$

For strong duality, in conic programming we need extra conditions on the constraints. These conditions are normally called *constraint qualifications*. The most well-known constraint qualification is the so-called *Slater condition*. In the case of $(Cone_P)$ the Slater condition reads:

**Definition 1.9** (Primal Slater Condition)**.** We say that $(Cone_P)$ satisfies the Slater condition if there exists $\mathbf{x} \in \mathbb{R}^n$ such that $B - \sum_{i=1}^n x_i A_i \in \text{int}(\mathcal{K})$.

Here $\text{int}(\mathcal{K})$ denotes the interior of the cone $\mathcal{K}$. The Slater condition for the dual $(Cone_D)$ can be defined in a similar manner. Note that in the above example both the primal and the dual do not satisfy the Slater condition. By assuming that the Slater conditions holds, one can derive a strong duality result for cone programming.

**Theorem 1.10** (Strong Duality)**.** *For the primal dual cone programs $(Cone_P)$ and $(Cone_D)$ the following holds.*

*i. If the primal problem $(Cone_P)$ satisfies the Slater condition and $\mathcal{F}(Cone_D)$ is nonempty, then the dual problem $(Cone_D)$ attains its optimal values and $val(Cone_P) = val(Cone_D)$.*

*ii. If the dual problem $(Cone_D)$ satisfies the Slater condition and $\mathcal{F}(Cone_P)$ is nonempty then the primal problem $(Cone_P)$ attains its optimal values and $val(Cone_P) = val(Cone_D)$.*

*Proof.* See e.g. [11]. □

### 1.3.1 Linear Programming

As mentioned earlier for the case $\mathcal{K} = \mathcal{N}_m$, $(Cone_P)$ and $(Cone_D)$ becomes a linear program (LP). Linear programming is an intensively studied sub-area of mathematical optimization. There exists a plethora of real world problems which can be formulated as a linear programming problem (see e.g. [61, Chapter 2], [68]).

Duality plays an important role in developing algorithms for solving mathematical optimization problems. Since linear programming has nice duality properties, it is no surprise that there exist many state of the art algorithms for solving linear programs.

The most well known and widely used method is the *simplex method* originally developed by Dantzig. Although the simplex method is adopted widely for solving linear programs, it is well known that the method can take exponential time in a worst case scenario [104]. This drawback led to the search for new algorithms for linear programming with polynomial time complexity. The real breakthrough in this area came when Khachiyan [101] published his polynomial time *ellipsoidal algorithm*. In spite of the promising polynomial time running time of the ellipsoidal method, it is not suitable for most applications due to its slow convergence. Another breakthrough came with the work of Karmarkar [99] on *interior point methods*, which were proved to be polynomial with faster convergence guarantees. For details on interior point methods for solving linear optimization problems the interested reader is referred to [132].

### 1.3.2 Semidefinite Programming

The cone program for the special case when $\mathcal{K} = \mathcal{S}_m^+$ is referred to as semidefinite program(SDP). Semidefinite programming can be seen as a natural generalization of linear programming where linear inequalities are replaced by semidefinitness conditions.

In contrast to linear programming even if all data in the SDP are rational we can end up in an irrational solution.

**Example 1.11.** *Consider,*

$$(Cone_P) \quad \max_{x\in\mathbb{R}} \quad x \quad s.t. \qquad \begin{pmatrix} 2 & -x \\ -x & 1 \end{pmatrix} \in \mathcal{K}$$

*then for $\mathcal{K} = \mathcal{S}_2^+$ it can be easily verified that the solution is $\overline{x} = val(Cone_P) = \sqrt{2}$ while for $\mathcal{K} = \mathcal{N}_2$ we have $\overline{x} = 0$.*

Since a rational SDP (when all input data in SDP are rational) can have an irrational solution, we cannot hope for an exact polynomial solution method. However, there exist algorithms which can approximate the solution of SDP to any fixed precision in polynomial time. The interior point methods of Karmarkar are generalized to SDP in [6, 5]. The ellipsoidal method of

Khachiyan is also generalized to SDP, but as in the case of linear programming, the ellipsoidal method suffers from slow convergence.

SDP has become a very attractive area of research among the optimization community due to its large applications. The most appealing and useful application of SDP is the SDP relaxation, which has numerous applications in combinatorial optimization. Although strong duality does not hold in general for SDP, in most SDP relaxations of combinatorial optimization problems strong duality is satisfied (see e.g. [33, 127, 128]).

The most popular SDP relaxation is for the *Max-Cut* problem. Using a SDP relaxation along with randomization, Goemans and Williamson [74] has obtained a $0.878$- approximation algorithm for the Max-Cut problem. This was a major breakthrough for SDP. It has opened a way for the application of SDP in combinatorial optimization problems. This problem is further discussed in [130]. The SDP relaxation of the *stability number* of a graph resulted in the so-called Lovasz theta number. The theta number has not only provided a bound on the stability number of the graph but also provided a polynomial time algorithm for finding the stability number in a so-called perfect graph, for details see [110, 120]. The well known spectral bundle methods for the *eigenvalue optimization* problem are based on the concept of SDP, for details see [152]. SDP has been proved very useful for approximating nonlinear problems. Specifically quadratically constrained quadratic programs(QCQP) are approximated by the use of SDP relaxations (for details see [4, 9, 148]). There are many other complex problems for which SDP has provided promising results, this list of problems includes the satisfiability problem [8, 83], maximum clique and graph colouring [26, 57, 56], non-convex quadratic programs [65], graph partitioning [69, 122, 153, 155], nonlinear 0-1 programming [106, 107], the knapsack problem [86], the travelling salesman problem [49], the quadratic assignment problem [122, 155], subgraph matching [136], statistics [152, Chapter 16 and 17], control theory [148], structural design [152, Chapter 15] and many other areas of science and engineering. In [152], a lot of material on theory, methods and applications of SDP is presented.

### 1.3.3 Copositive Programming

The cones of positive semidefinite matrices and of nonnegative matrices have the nice property that both are self dual. In this subsection we will briefly discuss cone programs over the cone of copositive matrices which is not self dual. Here we rewrite the cone program for the special case when $\mathcal{K} = \mathcal{C}_m$, since it will be widely discussed throughout the thesis.

$$(COP_P) \qquad \max_{x\in\mathbb{R}^n} \; c^T x \qquad \textit{s.t.} \quad B - \sum_{i=1}^{n} x_i A_i \in \mathcal{C}_m$$

$$(COP_D) \qquad \min_{Y\in\mathcal{S}_m} \; \langle Y, B\rangle \qquad \textit{s.t.} \quad \langle Y, A_i\rangle = c_i \; \forall\, i = 1,\dots,n, \;\; Y \in \mathcal{C}_m^*,$$

with $\mathbf{c} \in \mathbb{R}^n$ and $A_i, B \in \mathcal{S}_m$. We assume throughout that the matrices $A_i$, $i = 1,\dots,n$ are linearly independent.

During the last years, copositive programming has attracted much attention due to the fact that many difficult (NP-hard) quadratic and integer programs can be reformulated equivalently as copositive programs (COP) (see e.g. [28, 39, 47, 124, 123]). This reformulation clearly does not make these intractable problems tractable, but this reformulation can lead to new approximation guarantees for NP-hard problems as is the case for the standard quadratic optimization problem (see [28] and Remark 5.15).

From Example 1.8, it is clear that strong duality does not hold for copositive programming in general. In Chapter 5, we will briefly discuss duality in copositive programming from the viewpoint of linear semi-infinite programming. In [30], examples of COP are given where either attainability of a solution fails or there exists a nonzero duality gap.

It is well known that copositive programming is NP-hard. A main problem lies in checking the membership of a matrix in the cone of copositive matrices. Note, that it has been established that checking if a matrix is copositive is co-NP-hard [117]. Since there cannot exists a polynomial algorithm for solving copositive programming (assuming P $\neq$ NP), one has to rely on approximation methods. There exist roughly three method/algorithms for solving/approximating copositive programs namely the $\epsilon$-approximation algorithm [38] and its variations [143, 158, 157], approximation hierarchy based methods [28, 47, 120, 154] and feasible descent methods [19]. The $\epsilon$-approximation algorithm approximates $(COP_P)$ while approximation hierarchy based methods exist for both $(COP_P)$ [28, 47, 120] and $(COP_D)$ [154]. The feasible descent method in [19] approximates $(COP_D)$. The $\epsilon$- approximation algorithm of Bundfuss and Dür is reanalysed as a special case of a discretization method for semi-infinite programming (see Section 1.4) in Chapter 5. For surveys on results and methods for copositive programming the interested reader is referred to [18, 30, 58].

## 1.4 Semi-infinite Programming

In semi-infinite programming, as mentioned before, the objective function is optimized under an infinite set of constraints. In this section we shall restrict ourself to linear semi-infinite programming problems (LSIP). One can write LSIP in the following form,

$$(SIP_P) \qquad \max_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \qquad s.t. \quad b(\mathbf{z}) - a(\mathbf{z})^T\mathbf{x} \geq 0 \quad \forall\, \mathbf{z} \in Z,$$

with an infinite compact index set $Z \subseteq \mathbb{R}^m$ and continuous functions $a : Z \to \mathbb{R}^n$ and $b : Z \to \mathbb{R}$. It is not difficult to show that $\mathcal{F}(SIP_P)$ is closed.

One can associate different dual problems with $(SIP_P)$. Here we shall use the so-called Haar dual, which reads as follows,

$$(SIP_D) \qquad \min_{y_\mathbf{z}} \quad \sum_{\mathbf{z}\in Z} y_\mathbf{z} b(\mathbf{z}) \quad s.t. \quad \sum_{\mathbf{z}\in Z} y_\mathbf{z} a(\mathbf{z}) = \mathbf{c},\ y_\mathbf{z} \geq 0,$$

where only a finite number of dual variables $y_\mathbf{z}, \mathbf{z} \in Z$ (are allowed to) attain positive values. For the formulation of the Haar dual the interested reader is referred to [44], while the properties of the Haar dual are discussed in [71].

Note that $(SIP_D)$ is feasible if and only if **c** belongs to the cone generated by vectors $a(\mathbf{z}),\ \mathbf{z} \in Z$, that is

$$(SIP_D) \text{ is feasible if and only if } \mathbf{c} \in \text{cone}\{a(\mathbf{z}) : \mathbf{z} \in Z\} \tag{1.1}$$

LSIP has been widely applied in many areas of engineering including, but not limited to: the pattern recognition problem , the maximum likelihood regression and robust optimization (see [88, 73, 108, 149]).

The duality theory for LSIP is very well studied. In contrast to linear programming, again, strong duality does not hold in general for LSIP. In order to ensure strong duality, as before, we need Slater conditions for LSIP. The primal and the dual Slater conditions for LSIP are given below.

**Definition 1.12** (Slater Condition(LSIP)). The primal Slater condition holds

$$\text{if there exists} \quad \mathbf{x} \in \mathbb{R}^n \text{ with } b(\mathbf{z}) - a(\mathbf{z})^T \mathbf{x} > 0 \ \ \forall\, \mathbf{z} \in Z \tag{1.2}$$

We say that the dual Slater condition holds if

$$c \in \operatorname{int}(\operatorname{cone}\{a(\mathbf{z}) : \mathbf{z} \in Z\}) \tag{1.3}$$

We introduce the *upper level sets* for LSIP,

$$\mathcal{F}_\alpha(SIP_P) = \{\mathbf{x} \in \mathcal{F}(SIP_P) : \mathbf{c}^T\mathbf{x} \geq \alpha\},\ \alpha \in \mathbb{R}.$$

Let $\mathcal{S}(SIP_P)$ denote the set of maximizers of $(SIP_P)$. Recall that, in general, for LSIP strong duality need not hold and solutions of $(SIP_P)$ and/or $(SIP_D)$ need not exist. However, the following is true for linear SIP (see Theorem 1.10 for a corresponding result in cone programming).

**Theorem 1.13.** *We have:*

*i. If either* (1.2) *or* (1.3) *holds, then* $val(SIP_P) = val(SIP_D)$.

*ii. Let* $\mathcal{F}(SIP_P)$ *be non-empty. Then*

(1.3) *holds* $\Leftrightarrow$ $\forall \alpha \in \mathbb{R}$: $\mathcal{F}_\alpha(SIP_P)$ *is compact* $\Leftrightarrow$ $\emptyset \neq \mathcal{S}(SIP_P)$*is compact.*

*Thus, if one of these equivalent conditions holds, then a solution of* $(SIP_P)$ *exists.*

*A result as in ii. also holds for the dual problem.*

*Proof.* See, *e.g.*, [88, Theorems 6.9, 6.11] and [108, Theorem 4] for the second equivalence in *ii.*. □

In the theorem below we will give optimality conditions for LSIP. This requires the so-called KKT conditions.

**Definition 1.14** (Active Index Set). Let $\overline{\mathbf{x}} \in \mathcal{F}(SIP_P)$. Then the active index set for $\overline{\mathbf{x}}$ denoted by $Z(\overline{\mathbf{x}})$ is given by,

$$Z(\overline{\mathbf{x}}) = \{\overline{\mathbf{z}} \in Z : a(\overline{\mathbf{z}})^T\overline{\mathbf{x}} = b(\overline{\mathbf{z}})\} \tag{1.4}$$

The set $Z(\overline{\mathbf{x}})$ is a closed and compact subset of $Z$.

**Definition 1.15** (KKT Condition). A feasible point $\overline{\mathbf{x}} \in \mathcal{F}(SIP_P)$ is said to satisfy the KKT condition if there exist multipliers $\mu_1, ..., \mu_k \geq 0$ and indices $\overline{\mathbf{z}}_j \in Z(\overline{\mathbf{x}})$, $j = 1, \cdots, k$ such that,

$$\nabla_{\mathbf{x}} \mathbf{c}^T \overline{\mathbf{x}} - \sum_{j=1}^{k} \mu_j \nabla_{\mathbf{x}} (a(\overline{\mathbf{z}}_j)^T \overline{\mathbf{x}} - b(\overline{\mathbf{z}}_j))) = \mathbf{o}$$

or equivalently,

$$\sum_{j=1}^{k} \mu_j a(\overline{\mathbf{z}}_j) = \mathbf{c} \tag{1.5}$$

The optimality conditions for LSIP are given below,

**Theorem 1.16.** *If a point $\overline{\mathbf{x}} \in \mathcal{F}(SIP_D)$ satisfies the KKT condition* (1.5) *then $\overline{\mathbf{x}}$ is a (global) maximizer of $(SIP_P)$. On the other hand under the conditions* (1.2) *a maximizer $\overline{\mathbf{x}}$ of $(SIP_P)$ must satisfy the KKT conditions.*

*Proof.* See [108, Theorem 3] and [88, Theorem 2(b)]. □

Although LSIP is a convex program, the existence of a polynomial time algorithm is not possible for LSIP. The main difficulty lies in checking the constraint $a(\mathbf{z})^T \overline{\mathbf{x}} \leq b(\mathbf{z})$ for all $\mathbf{z} \in Z$. The numerical methods available can be classified into five main categories: discretization methods, local reduction method, exchange methods, simplex-like methods and descent methods.

*Discretization methods* are based on solving a sequence of finite programs. The sequence of finite programs are solved according to some pre-defined grid generation scheme or some cutting plane scheme. The method boost for their global convergence guarantees. Beside the global convergence guarantee, discretization methods are known to be very slow in practice. Interestingly the $\epsilon$- approximation algorithm [38] for solving copositive programs can be seen as a special case of a discretization method. We will discuss this relation in detail in Chapter 5.

In the *local reduction method* the original problem is replaced by a locally equivalent problem with finitely many inequality constraints. The problem can also be replaced by a system of nonlinear equations with finitely many unknowns. This system can be solved by Newton's method and hence these methods may have good convergence results. Reduction based SQP-methods are one example of these kind of methods.

The *exchange methods* can be seen as a compromise between discretization methods and reduction methods. Hence they are more efficient than discretization methods. For details see [87, 88, 129].

The *simplex-like* methods for solving LSIP problems, as the name suggests, are modifications of the simplex method for linear programming (for details see [7]).

For more details on theory algorithms and applications of LSIP the interested reader is referred to [72].

## 1.5 Cone Programming Relaxations of Quadratic Problems

In this section a brief introduction into cone programming relaxations for quadratic programs is presented.

We consider the following quadratic program,

$$(QCQP) \quad \min_{\mathbf{u}} \ \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \mathbf{u}^T A_j \mathbf{u} + 2\mathbf{c}_j^T\mathbf{u} + \mathbf{b}_j \leq 0, \ \ \forall\, j \in J \quad \mathbf{u} \in K$$

where $J := \{1, 2, \cdots, k\}$, $K \subseteq \mathbb{R}^m$ is a closed convex cone, $A_j \in \mathcal{S}_m, \mathbf{c}_j \in \mathbb{R}^m$ and $b_j \in \mathbb{R}$. If $A_j \notin \mathcal{S}_m^+$ then $(QCQP)$ is not convex. A standard way to make this program convex is to gather all nonlinearities in one constraint. To do so, we introduce a matrix $U$, such that $U = \mathbf{u}\mathbf{u}^T$ and consider,

$$\mathbf{u}^T A_i \mathbf{u} = \left\langle A_i, \mathbf{u}\mathbf{u}^T \right\rangle = \langle A_i, U \rangle\,.$$

Then $(QCQP)$ can be equivalently written as,

$$(QCQP) \quad \min_{\mathbf{u},U} \ \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \begin{array}{l} \langle A_j, U\rangle + 2\mathbf{c}_j^T\mathbf{u} + b_j \leq 0, \ \ \forall\, j \in J, \\ U = \mathbf{u}\mathbf{u}^T, \ \ \mathbf{u} \in K \end{array}$$

The cone programming relaxation, relaxes the constraint $U = \mathbf{u}\mathbf{u}^T$ into cone constraints. To do so, we define the cone of matrices,

$$\mathcal{K}^* := \left\{ Y \in \mathcal{S}_{m+1} : Y = \sum_{j=1}^{r} \mu_j \begin{pmatrix} 1 \\ \mathbf{u}_j \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{u}_j \end{pmatrix}^T, \mathbf{u}_j \in K, \mu_j \geq 0, r \in \mathbb{N} \right\}$$

Note that $U = \mathbf{u}\mathbf{u}^T$ can be equivalently written as $\left(\begin{smallmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ \mathbf{u} \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ \mathbf{u} \end{smallmatrix}\right)^T$ and then use the relaxation $\left(\begin{smallmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{smallmatrix}\right) \in \mathcal{K}^*$. Note also that $\langle A_j, U\rangle + 2\mathbf{c}_j^T\mathbf{u} + b_j$ can be written as $\left\langle Q_j, \left(\begin{smallmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{smallmatrix}\right)\right\rangle$ where $Q_j := \left(\begin{smallmatrix} b_j & \mathbf{c}_j^T \\ \mathbf{c}_j & A_j \end{smallmatrix}\right)$.

For the cases when $K = \mathbb{R}^m$ and $K = \mathbb{R}^m_+$, we obtain the following SDP and COP relaxations for $(QCQP)$:

$$(SDP) \qquad \min\ \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \begin{array}{l} \left\langle Q_j, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \right\rangle \leq 0,\ j \in J \\ \text{and} \quad \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \in \mathcal{S}^+_{m+1} \end{array}$$

$$(COP) \qquad \min\ \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \begin{array}{l} \left\langle Q_j, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \right\rangle \leq 0,\ j \in J \\ \text{and} \quad \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \in \mathcal{C}^*_{m+1} \end{array}$$

A natural question is to ask how sharp these relaxations can be? We analyse this question in Chapter 4.

## 1.6 Thesis Outline

The main focus of this thesis is copositive programming and related problems. In this section an outline of the thesis, with an indication of the main results, is given.

Chapter 2, is a review of results related to set-semidefinite cones. Results on the copositive cone and its dual, the completely positive cone, are also discussed. The following are the main (new) results presented in this chapter:

- With the help of an example, it is shown that the well-known Schur complement for semidefinite matrices cannot be extended to the case of general set-semidefinite matrices.
- Some (known) characterizations of copositivity and complete positivity are provided.
- It is shown that positive diagonally dominant matrices belong to the interior of the completely positive cone.

The results of Chapter 4 and Chapter 5 have appeared in [2] and [1] respectively, while Chapter 3 is based on the working paper [3]. The main results of these chapters are listed below.

Chapter 3 mainly deals with the standard quadratic programming problem (StQP). The following are the main (new) results discussed in this chapter:

- A characterization of strict local maximizers is provided. In the literature, the characterizations for strict local maximizers are given under the condition that the candidate maximizer satisfies strict complementarity. Our characterization does not require this condition.

- We show that standard quadratic programming problem involving nonsingular matrix for which all principle submatrices are nonsingular has at least one strict local maximizer.

- Results on Lipschitz stability and strong stability of strict local maximizers with respect to perturbations in the matrix involved are studied. These results are obtained by applying (known) results of parametric optimization to the special case of standard quadratic programming.

- It is shown that generically every local maximizer is a strict local maximizer.

- A review of evolutionarily stable strategies is given with an emphasis on the maximum number of ESS and the relation of ESS with strict local maximizers of StQP

In Chapter 4, we look at the extension of a result which compares the feasible set of a nonconvex quadratic program and the feasible set of its semidefinite relaxation. We give an extension of this result for the case of set-semidefinite relaxations.

In Chapter 5, we reformulate a copositive program as a linear semi-infinite program. The main contributions in this chapter are:

- We study COP from the viewpoint of LSIP and rediscuss optimality and duality results for COP.

- We interpret different approximation schemes for solving COP as a special case of the discretization method for LSIP. This interpretation leads to sharper error bounds for the values and solutions of the approximate programs in dependence on the mesh size. With the help of examples we illustrate the structure of the original problem and the approximation schemes.

- The question of order of maximizers for COP is also analysed. It is shown with the help of examples that for COP maximizers of an arbitrarily high order can exist.

**Publications Underlying This Thesis**

- F. Ahmed and G. J. Still, *Quadratic maximization on the unit simplex: structure, stability, genericity and application in biology*, Memorandum 2034, Department of Applied Mathematics, University of Twente, Enschede, February 2014. **(Chapter 3)**
- F. Ahmed and G. Still, *A note on set-semidefinite relaxations of nonconvex quadratic programs*, Journal of Global Optimization, 57 (2013), pp. 1139--1146. **(Chapter 4 and Section 2.1)**
- F. Ahmed, M. Dür, and G. Still, *Copositive programming via semi-infinite optimization*, Journal of Optimization Theory and Applications, 159 (2013), pp. 322--340. **(Chapter 5)**

# 2

# Cones of Matrices

A quadratic form is said to be set-semidefinite if it is nonnegative over some closed cone. It is interesting to study the cone of matrices associated with such quadratic forms due to their applicability in many areas including mathematical programming. In this chapter we will briefly describe some results on set-semidefinite matrices. We will give particular emphasis to a special set-semidefinite cone namely the copositive cone. We shall describe the cone properties and characterizations for checking the membership in these cones and their dual cones.

## 2.1 Set-Semidefinite Cone

The notion of a set-semidefinite cone is a generalization of the positive semidefinite cone. We will study set-semidefinite relaxations of nonconvex quadratic programs in Chapter 4. Most of the results presented in this section have appeared in [2].

We start by defining the set-semidefinite cone,

**Definition 2.1.** For a given closed cone $K \subseteq \mathbb{R}^m$ we define the set $\mathcal{C}_m(K)$ of $K$-semidefinite $m \times m$-matrices and its dual cone $\mathcal{C}^*_m(K)$ of $K$-positive $m \times m$-

matrices as:

$$\mathcal{C}_m(K) = \{Q \in \mathcal{S}_m : \mathbf{v}^T Q\mathbf{v} \geq 0 \;\forall \mathbf{v} \in K\} \tag{2.1}$$

$$\mathcal{C}_m^*(K) = \left\{ U = \sum_j \alpha_j \mathbf{u}_j \mathbf{u}_j^T : \alpha_j \geq 0,\ \mathbf{u}_j \in K \right\} \tag{2.2}$$

For $K = \mathbb{R}^m$ we obtain the (self-dual) cone $\mathcal{S}_m^+$ of positive semidefinite matrices and for $K = \mathbb{R}_+^m$ the cones of copositive respectively completely positive matrices.

The study of nonnegativity of a quadratic form over a convex cone can be traced back to the work of Cottle et al [46]. Sturm and Zhang have studied the properties of such cones in detail [145] while algebraic properties of these cones is the topic of Gowda et al [76].

The cones $\mathcal{C}_m(K)$ and $\mathcal{C}_m^*(K)$ are closed and convex [76]. In the following lemma we will show that indeed the dual of $\mathcal{C}_m(K)$ is given by (2.2).

**Lemma 2.2.** *For any closed set $K \subseteq \mathbb{R}^m$ the dual of $\mathcal{C}_m(K)$ is $\mathcal{C}_m^*(K)$ as given in Definition 2.1.*

*Proof.* We show that with

$$\mathcal{C} := \left\{ U = \sum_j \alpha_j \mathbf{u}_j \mathbf{u}_j^T : \alpha_j \geq 0, \mathbf{u}_j \in K \right\},$$

we have $\mathcal{C}_m(K) = \mathcal{C}^*$. By using $\mathcal{C}^{**} = \mathcal{C}$ (for closed convex cones $\mathcal{C}$, see e.g. [67, Lemma 4.4.1]) we find the identity claimed in the lemma.
"$\subset$": If $Q \in \mathcal{C}_m(K)$ then for all $U \in \mathcal{C}$ we obviously have,

$$\langle Q, U\rangle = \sum_j \alpha_j \langle Q, \mathbf{u}_j \mathbf{u}_j^T\rangle \geq 0$$

implying $Q \in \mathcal{C}^*$.
"$\supset$": Suppose $Q \notin \mathcal{C}_m(K)$, *i.e.*, $\mathbf{u}^T Q\mathbf{u} < 0$ for some $\mathbf{u} \in K$. Then for $U = \mathbf{u}\mathbf{u}^T \in \mathcal{C}$ it follows $\langle U, Q\rangle < 0$, so that $Q \notin \mathcal{C}^*$. □

In linear algebra, the *Schur complement* plays an important role for developing properties and characterizations of matrices. For example in developing copositivity criteria, Väliaho [146] has made use of the Schur's complement. In

Lemma 2.3 a generalization of the Schur complement is given. Let in the following $K \subseteq \mathbb{R}^m$ be a closed cone.

**Lemma 2.3.** *It holds*

$$\begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} \in \mathcal{C}_{m+1}(\mathbb{R}_+ \times K) \quad \Leftrightarrow \quad \begin{array}{l} \gamma \geq 0,\ C \in \mathcal{C}_m(K) \quad \textit{and} \\ \mathbf{v}^T(\gamma C - \mathbf{c}\mathbf{c}^T)\mathbf{v} \geq 0 \ \ \forall \mathbf{v} \in K \ \textit{with}\ \mathbf{c}^T\mathbf{v} \leq 0 \ . \end{array} \tag{2.3}$$

*Proof.* The left-hand side means:

$$(\alpha\ \mathbf{v})^T \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} \begin{pmatrix} \alpha \\ \mathbf{v} \end{pmatrix} = \gamma\alpha^2 + 2\alpha\mathbf{c}^T\mathbf{v} + \mathbf{v}^T C\mathbf{v} \geq 0 \ \ \forall \alpha \geq 0,\ \mathbf{v} \in K \ .$$

``$\Rightarrow$'': The above inequality implies $\gamma \geq 0$, $\mathbf{v}^T C\mathbf{v} \geq 0$ for all $\mathbf{v} \in K$ and in the case $\mathbf{c}^T\mathbf{v} \geq 0$ we are done. In the case $\mathbf{c}^T\mathbf{v} \leq 0, \gamma = 0$ we also obtain $\mathbf{c}^T\mathbf{v} = 0$. For the remaining case $\mathbf{c}^T\mathbf{v} \leq 0, \gamma > 0$ we write

$$0 \leq \gamma\alpha^2 + 2\alpha\mathbf{c}^T\mathbf{v} + \mathbf{v}^T C\mathbf{v} = \frac{1}{\gamma}(\gamma\alpha + \mathbf{c}^T\mathbf{v})^2 + \frac{1}{\gamma}\mathbf{v}^T(\gamma C - \mathbf{c}\mathbf{c}^T)\mathbf{v} \ .$$

Then the assumption $\mathbf{v}^T(\gamma C - \mathbf{c}\mathbf{c}^T)\mathbf{v} < 0$ for some $\mathbf{v} \in K, \mathbf{c}^T\mathbf{v} \leq 0$ leads to a contradiction (with a choice $\gamma\alpha = -\mathbf{c}^T\mathbf{v} \geq 0$). The direction ``$\Leftarrow$'' is easy. $\square$

It is interesting to note that in the special case of positive semidefinite matrices, the above lemma coincides with the well known Schur complement result,

$$\begin{pmatrix} 1 & \mathbf{v}^T \\ \mathbf{v} & V \end{pmatrix} \in \mathcal{S}^+_{m+1} \quad \Leftrightarrow \quad V - \mathbf{v}\mathbf{v}^T \in \mathcal{S}^+_m \ .$$

Unfortunately such a relation is no more true for $\mathcal{C}^*_m(K)$. We only have,

**Lemma 2.4.** *Let $V \in \mathcal{S}_m, \mathbf{v} \in K$ be such that $V - \mathbf{v}\mathbf{v}^T \in \mathcal{C}^*_m(K)$. Then also*

$$\begin{pmatrix} 1 & \mathbf{v}^T \\ \mathbf{v} & V \end{pmatrix} \in \mathcal{C}^*_{m+1}(\mathbb{R}_+ \times K).$$

*Proof.* By definition, the matrix $V - \mathbf{v}\mathbf{v}^T \in \mathcal{C}^*_m(K)$ can be written in the form

$$V - \mathbf{v}\mathbf{v}^T = \sum_{j=1}^{k} \lambda_j \mathbf{u}_j \mathbf{u}_j^T \quad \text{with } \lambda_j \geq 0,\ \mathbf{u}_j \in K,\ j = 1, \ldots, k \ .$$

So, the decomposition

$$\begin{pmatrix} 1 & \mathbf{v}^T \\ \mathbf{v} & V \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{v} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{v} \end{pmatrix}^T + \sum_{j=1}^{k} \lambda_j \begin{pmatrix} 0 \\ \mathbf{u}_j \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{u}_j \end{pmatrix}^T$$

holds and recalling $\mathbf{v} \in K$, this matrix is an element of $\mathcal{C}^*_{m+1}(\mathbb{R}_+ \times K)$. □

The converse of Lemma 2.4 is not true in general (if $K \neq \mathbb{R}^m$). Consider the following example,

**Example 2.5.** *Take the copositive case, i.e. , $K = \mathbb{R}^m_+, m = 2$, and choose,*

$$V = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \mathbf{v} = (1, 1)^T$$

*Then,*

$$\begin{pmatrix} 1 & \mathbf{v}^T \\ \mathbf{v} & V \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}^T + \frac{1}{2} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}^T \in \mathcal{C}^*_{m+1}(\mathbb{R}_+ \times \mathbb{R}^m_+)$$

*but $V - \mathbf{v}\mathbf{v}^T = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right) \notin \mathcal{C}^*_m(\mathbb{R}^m_+)$, since a necessary condition for $Q \in \mathcal{C}^*_m$ (as is clear from* (2.2)*) is that $Q \in \mathcal{N}_m$.*

Now we will consider a generalization of the set-semidefinite cone. For a closed convex cone $K$ and a fixed $\alpha \in \mathbb{R}$ we consider:

$$\mathcal{C}_m(K, \alpha) := \left\{ Q \in \mathcal{S}_m : \mathbf{v}^T Q \mathbf{v} - \alpha \mathbf{v}^T \operatorname{diag}(Q) \geq 0, \forall \mathbf{v} \in K \right\} \tag{2.4}$$

where $\operatorname{diag}(A) \in \mathbb{R}^m$ is the vector of the diagonal elements of the matrix $A \in \mathcal{S}_m$, i.e., $\operatorname{diag}(A) = (a_{11}, \ldots, a_{mm})^T$. In the sequel $\operatorname{Diag}(\mathbf{u})$ denotes the matrix with $\mathbf{u} \in \mathbb{R}^m$ on the main diagonal while all other elements are zero. In order to construct the dual of $\mathcal{C}_m(K, \alpha)$, notice the relation $\mathbf{v}^T \operatorname{diag}(Q) = \langle Q, \operatorname{Diag}(\mathbf{v}) \rangle$, and find:

$$\mathbf{v}^T Q \mathbf{v} - \alpha \mathbf{v}^T \operatorname{diag}(Q) = \langle Q, \mathbf{v}\mathbf{v}^T - \alpha \operatorname{Diag}(\mathbf{v}) \rangle$$

Then by construction the dual of $\mathcal{C}_m(K, \alpha)$ will be:

$$\mathcal{C}^*_m(K, \alpha) := \left\{ U = \sum_j \lambda_j (\mathbf{u}_j \mathbf{u}_j^T - \alpha \operatorname{Diag}(\mathbf{u}_j)) : \lambda_j \geq 0, \mathbf{u}_j \in K \right\}$$

It is easily shown, as in the proof of Lemma 2.2, that the cones $\mathcal{C}_m(K,\alpha), \mathcal{C}^*_m(K,\alpha)$ are dual to each other.

Note that the set-semidefinite cones $\mathcal{C}_m(K)$ and $\mathcal{C}^*_m(K)$ are the special instances of $\mathcal{C}_m(K,\alpha)$ and $\mathcal{C}^*_m(K,\alpha)$ respectively for the case when $\alpha = 0$.

## 2.2 Copositive Cone

In this section we will describe a special type of a set-semidefinite cone namely the copositive cone. Recall from Definition 1.5 that a matrix $Q \in \mathcal{S}_m$ is copositive if and only if $\mathbf{v}^T Q\mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^m_+$. The set of all $m \times m$ copositive matrices forms a closed, convex, full dimensional and non polyhedral cone [37]. In this section we will confine ourself to the relation between copositivity and positive semidefiniteness, characterizations of copositivity and finally some words on the interior and extreme rays of the copositive cone.

The copositive matrices were introduced in 1952 by Motzkin [114]. Since then these matrices caught attention of researchers. Much work has been done on extending results on positive semidefinite matrices to copositive matrices. Copositivity has vast applications in different areas of science and engineering. For an overview of these applications the interested reader is referred to [18] and the references therein.

### 2.2.1 Copositivity and Positive Semidefiniteness

In this subsection we will discuss the relations between copositivity and positive semidefiniteness. From the definition of copositive matrices, it is clear that every positive semidefinite matrix is also copositive, but the converse is not true in general. For example, the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ is clearly copositive but not positive semidefinite. We will describe special cases where the two classes coincide. We start with the following lemma which says that every matrix with non-positive off-diagonal entries is copositive if and only if it is positive semidefinite. In the following $\mathbb{R}_{++}$ denote the set of positive real numbers.

**Lemma 2.6** ([96]). *Let $Q \in \mathcal{S}_m$ and all off-diagonal entries of $Q$ are non-positive ($q_{ij} \leq 0$ for all $i \neq j$) then $Q$ is copositive if and only if it is positive semidefinite.*

*Proof.* If $Q \in \mathcal{S}^+_m$, then the lemma is obvious. For the converse suppose that $Q \in \mathcal{C}_m$, then for all $\mathbf{v} \in \mathbb{R}^m_+, \mathbf{v}^T Q\mathbf{v} \geq 0$ , also for $\mathbf{u} = -\mathbf{v}, \mathbf{u}^T Q\mathbf{u} \geq 0$. Now suppose that $\mathbf{v} \in \mathbb{R}^m$ has at least one zero, one positive and one negative component, then consider $\mathbf{v} = \begin{pmatrix} \mathbf{o} & \mathbf{u} & \mathbf{w} \end{pmatrix}^T$, where $\mathbf{o}$ is a zero vector of dimension $t$, $\mathbf{u} \in \mathbb{R}^s_{++}$ and $-\mathbf{w} \in \mathbb{R}^{m-t-s}_{++}$. Partition the matrix $Q$ such that

$$Q = \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12}^T & Q_{22} & Q_{23} \\ Q_{13}^T & Q_{23}^T & Q_{33} \end{pmatrix}$$

where $Q_{11}$ is the $t \times t$ matrix, $Q_{12}$ is the $t \times s$ matrix, $Q_{13}$ is the $t \times (m-s-t)$ matrix, $Q_{22}$ is the $s \times s$ matrix, $Q_{23}$ is the $s \times (m-s-t)$ matrix and $Q_{33}$ is the $(m-s-t) \times (m-s-t)$ matrix. Note that $Q_{23}\mathbf{w} \geq \mathbf{o}$ since both $Q_{23}$ is non-positive and $\mathbf{w}$ is negative. Hence we have,

$$\mathbf{v}^T Q \mathbf{v} = \underbrace{\mathbf{u}^T Q_{22} \mathbf{u}}_{\geq 0} + 2\mathbf{u}^T Q_{23} \mathbf{w} + \underbrace{\mathbf{w}^T Q_{33} \mathbf{w}}_{\geq 0} \geq 0$$

The case when $\mathbf{v}$ does not contain a zero entry can be proved similarly. So for all $\mathbf{v} \in \mathbb{R}^m$, $\mathbf{v}^T Q \mathbf{v} \geq 0$. Hence the matrix is positive semidefinite. □

Semidefinite matrices are normally characterized by their eigenvalues since it is well known that a matrix is positive semidefinite if and only if all its eigenvalues are nonnegative. As one can already see from the above discussion, copositive matrices may have negative eigenvalues. Now the question arises how many negative eigenvalues a copositive matrix can have? The following example provides an answer to the question,

**Example 2.7.** *Let $Q := (1 + \frac{\varepsilon}{m})E - \varepsilon I \in \mathcal{S}_m$ for some $\varepsilon > 0$, small, where $E \in \mathcal{S}_m$ is the matrix of ones while $I \in \mathcal{S}_m$ is the identity matrix. Clearly for $0 < \varepsilon \leq \frac{m}{m-1}$, Q is copositive since it is nonnegative. Moreover, $m$ is an eigenvalue of Q since $Q\mathbf{e} = m\mathbf{e}$, where $\mathbf{e}$ is the vector of ones. Also $-\varepsilon$ is an eigenvalue with multiplicity $m-1$ since $Q\left(\begin{smallmatrix} -1 \\ e_i \end{smallmatrix}\right) = -\varepsilon\left(\begin{smallmatrix} -1 \\ e_i \end{smallmatrix}\right)$ for $i = 1, \cdots, m-1$ and the set $\left\{\left(\begin{smallmatrix} -1 \\ e_i \end{smallmatrix}\right), i = 1, \cdots, m-1\right\}$ is linearly independent, where $e_i$ are the unit vectors of length $m-1$.*

### 2.2.2 Characterization of Copositivity

In the literature, there exist several characterizations of copositivity. These characterizations are based on determinants of submatrices, on a solution of an associated system of equations or on exploiting the structure of the matrix. In this subsection we will start with a simple necessary condition for copositivity. Here and throughout the thesis we shall take $\mathcal{U} := \{1, \cdots, m\}$.

**Lemma 2.8.** *Let $Q \in \mathcal{C}_m$ then $q_{ii} \geq 0$ for all $i \in \mathcal{U}$.*

*Proof.* Let $Q \in \mathcal{C}_m$. Then for $e_i \in \mathbb{R}^m_+, i \in \mathcal{U}$ we have $q_{ii} = e_i^T Q e_i \geq 0$. □

It is clear from Example 2.7 that copositivity cannot be completely characterized with the help of nonnegative eigenvalues. But a partial characterization can be obtained by relating the number of positive eigenvalues with the copositivity of principal submatrices of certain order.

**Theorem 2.9.** *Suppose that a matrix $Q \in \mathcal{S}_m$ has $p$ positive eigenvalues, $p < m$. Then $Q$ is copositive if and only if all the principal submatrix of order $p+1$ and less are copositive.*

*Proof.* See [96, Theorem 4.16]. □

In the following lemma, we will provide conditions, for copositivity, for matrices of order two and three and refer the interested reader to [121], for the case of order four matrices.

**Lemma 2.10.** *The following holds,*

*i. $Q \in \mathcal{S}_2$ is copositive if and only if,*

$$q_{11} \geq 0, q_{22} \geq 0, q_{12} + \sqrt{q_{11}q_{22}} \geq 0$$

*ii. $Q \in \mathcal{S}_3$ is copositive if and only if,*

$$\begin{aligned} q_{11} \geq 0, q_{22} \geq 0, q_{33} \geq 0 \\ \bar{A} := q_{12} + \sqrt{q_{11}q_{22}} \geq 0, \bar{B} := q_{13} + \sqrt{q_{11}q_{33}} \geq 0, \bar{C} := q_{23} + \sqrt{q_{22}q_{33}} \geq 0 \\ \sqrt{q_{11}q_{22}q_{33}} + q_{12}\sqrt{q_{33}} + q_{13}\sqrt{q_{22}} + q_{23}\sqrt{q_{11}} + \sqrt{2}\sqrt{\bar{A}\bar{B}\bar{C}} \geq 0 \end{aligned}$$

*Proof.* See [78, 92]. □

A criterion for determining copositivity based on the structure of the principal submatrices is developed by Keller and appeared in [45]. This criterion uses the cofactors of the matrix.

**Definition 2.11** (Cofactor and Adjoint of the Matrix)**.** Let $Q^{ij}$ denote the matrix obtained from $Q$ after deleting the $i^{th}$ row of $Q$ and the $j^{th}$ column of $Q$, then the $ij^{th}$ *cofactor* of $Q$ ,denoted by $C_{ij}$, is given by,

$$C_{ij} = (-1)^{i+j} \det(Q^{ij}) .$$

The transpose of the matrix of all cofactors, denoted by adj$(Q)$, is called *adjoint* of $Q$, i.e., $(\text{adj}(Q))_{ij} = C_{ji}$. The adjoint of a matrix $Q$ is related to the determinant and the inverse of the matrix by the following identity,

$$Q^{-1} = \frac{1}{\det(Q)} \text{adj}(Q) \quad \text{ or } \quad \det(Q)I = Q\, \text{adj}(Q)\,.$$

First we consider two simple lemmas.

**Lemma 2.12** ([93]). *Let $Q \in \mathcal{C}_m$ and let $\mathbf{v} \in \mathbb{R}^m_+$. Then $\mathbf{v}^T Q\mathbf{v} = 0$ implies $Q\mathbf{v} \geq 0$.*

*Proof.* Let $Q \in \mathcal{C}_m$. Then for $\lambda > 0$ we have $\mathbf{v} + \lambda e_i \in \mathbb{R}^m_+$, and thus,

$$0 \leq (\mathbf{v} + \lambda e_i)^T Q(\mathbf{v} + \lambda e_i) = 2\lambda e_i^T Q\mathbf{v} + \lambda^2 e_i^T Qe_i = 2\lambda(Q\mathbf{v})_i + \lambda^2 q_{ii}$$

By dividing by $\lambda > 0$ and letting $\lambda \to 0$, we obtain

$$e_i^T Q\mathbf{v} = (Q\mathbf{v})_i \geq 0.$$

This holds for every $i \in \mathcal{U}$. $\square$

**Lemma 2.13** ([146]). *Let $Q \in \mathcal{C}_m$ and $\det(Q) \neq 0$, then the inverse of $Q$ cannot contain a non-positive column.*

*Proof.* Let $B = Q^{-1}$ and let some column say $\mathbf{b}_i$ be non-positive. Take $\mathbf{v} = -\mathbf{b}_i$, i.e., $\mathbf{v} \in \mathbb{R}^m_+$, which implies $Q\mathbf{v} = -Q\mathbf{b}_i = -e_i$ (since $\mathbf{b}_i$ is the $i^{th}$ column of $Q^{-1}$). Hence we get

$$\mathbf{v}^T Q\mathbf{v} = -\mathbf{b}_i(-Q\mathbf{b}_i) = -\mathbf{b}_i(-e_i) = b_{ii} \leq 0.$$

Since $Q$ is copositive equality holds in the above relation, i.e., $\mathbf{v}^T Q\mathbf{v} = 0$, which contradicts the results given in Lemma 2.12 (since $(Q\mathbf{v})_i < 0$). $\square$

Note that if $Q \in \mathcal{C}_m$ then all principal submatrices are copositive. The next result enables us to determine when a matrix is not copositive given that certain principal submatrices are copositive. Here and in the rest of the thesis, for a matrix $Q \in \mathcal{S}_m$ and an index set $J \subseteq \{1, 2, \ldots, m\}$, $Q_J$ will denote the principal submatrix obtained after deleting the rows and the columns of the matrix $Q$ not corresponding to the elements of the index set $J$, i.e., $Q_J \in \mathbb{R}^{|J|\times|J|}$ and $(Q_J)_{ij} = q_{ij}$ for all $i, j \in J$ where $(Q_J)_{ij}$ is the $ij^{th}$ element of the matrix $Q_J$.

**Theorem 2.14** ([45, Theorem 3.1]). *Let $Q \in \mathcal{S}_m$ and let all principal submatrices of $Q$ of order up to $m - 1$ be copositive. Then $Q \notin \mathcal{C}_m$ if and only if $\text{adj}(Q) \in \mathcal{N}_m$ and $\det(Q) < 0$.*

*Proof.* For a proof see [45, Theorem 3.1]. □

The following theorem is stated in [45]. Here we will include a proof for the sake of completeness.

**Theorem 2.15** (Keller [45]). *A matrix $Q \in \mathcal{S}_m$ is copositive if and only if each principal submatrix $Q_J$ for which all cofactors of the last row are nonnegative has nonnegative determinant. This includes for $|J| = 1$ the condition $q_{ii} \geq 0, i \in \mathcal{U}$.*

*Proof.* Suppose that $Q \in \mathcal{C}_m$, then each principal submatrix of $Q$ is also copositive. So it is sufficient to show that if the cofactors of the last row of $Q$ are nonnegative then the determinant is also nonnegative. It is not difficult to verify that $\mathbf{v} = \text{adj}(Q)e_m$ gives the cofactors of the last row. Since the cofactors of the last row are nonnegative **v** is nonnegative. We find,

$$\begin{aligned}
\mathbf{v}^T Q\mathbf{v} &= (\text{adj}(Q)e_m)^T\, Q\, (\text{adj}(Q)e_m) \\
&= e_m^T\, \text{adj}(Q)Q\, \text{adj}(Q)e_m \\
&= \det(Q)e_m^T\, \text{adj}(Q)e_m = \det(Q)\{\text{adj}(Q)\}_{mm} \geq 0\,.
\end{aligned}$$

Since $\{\text{adj}(Q)\}_{mm}$ is nonnegative the only possibility when $\det(Q)$ can be negative is when $\{\text{adj}(Q)\}_{mm} = 0$. Since the cofactors of the last row are nonnegative this implies that the last column in $\text{adj}(Q)$ is nonnegative. Hence if $\det(Q) < 0$, then we will get a non-positive column in $Q^{-1}$ which is a contradiction to Lemma 2.13.

For the converse suppose that each principle submatrix $Q_J$ for which all cofactors of last row are nonnegative have nonnegative determinant. In order to show that $Q \in \mathcal{C}_m$ holds we will use induction with respect to $m$.

We start the induction with $m = 1$, where the assumption yields $q_{11} \geq 0$.

For the induction step we suppose that all principal submatrices of order $k \leq m - 1$ are copositive. Now for $k = m$ we have two conditions:

i. each principal submatrix $Q_J$ for which all cofactors of the last row are nonnegative have nonnegative determinant.

ii. all the principal submatrices of order $m - 1$ are copositive.

Suppose now that ii. holds and the matrix $Q$ is not copositive, i.e., $Q \notin \mathcal{C}_m$. Then from Theorem 2.14, we have $\text{adj}(Q) \in \mathcal{N}_m$ and $\det(Q) < 0$. But this is a clear contradiction to $i.$ above. This concludes the proof. □

The characterization above suggests to check copositivity with the help of the computation of the determinants of all $2^m - 1$ principal submatrices, which is not computationally efficient. For the special case of tridiagonal matrices however this characterization led to a polynomial time algorithm for testing copositivity, see [126, Corollary 1].

The following theorem gives an alternative characterization for copositivity which relies on the solution of a system of inequalities for each principal submatrix instead of calculating the determinants of each submatrix.

**Theorem 2.16** (Gaddum [66]). *Let $Q \in \mathcal{S}_m$. Then $Q$ is copositive if and only if for all $J \subseteq \{1, 2, \ldots, m\}$, the following system has a solution,*

$$Q_J \mathbf{v}_J \geq 0 \quad \mathbf{v}_J \geq 0 \quad \mathbf{e}_{|J|}^T \mathbf{v}_J = 1 \,. \tag{2.5}$$

*Here $\mathbf{v}_J$ is the subvector such that $\mathbf{v}_J := (v_j : j \in J)$.*

*Proof.* For a simple proof see [48, Theorem 1]. □

### 2.2.3 Interior and Extreme Rays

The notions of interior and extreme rays of a cone helps to understand the geometry of the cone which in turn is useful for characterizations. Before proceeding further, we define what is meant by an extreme ray of a cone.

**Definition 2.17** (Extreme Ray). Let $K$ be a closed, pointed and full dimensional convex cone. Then the *ray* generated by $U \in K \backslash \{O\}$ is defined to be the set $\{\alpha U : \alpha \geq 0\}$. Moreover, $U \in K \backslash \{O\}$ defines an *extreme ray* of $K$ if

$$U_1, U_2 \in K, \;\; U = U_1 + U_2 \quad \Rightarrow \quad U_1, U_2 \in \{\alpha U : \alpha \geq 0\}$$

Ext$(K)$ will denote the set of elements of $K$ which generate extreme rays.

In the above definition and in the rest of the thesis $O$ denotes the zero matrix of appropriate dimension. A general characterization of the extreme rays of the copositive cone is unknown. But there exists partial results. These results are summarized below.

**Theorem 2.18.** *For $m \geq 2$ the following holds,*

*i.* $\alpha(e_i e_j^T + e_j e_i^T) \in$ Ext$(\mathcal{C}_m)$, *where $i, j = 1, \cdots m, \alpha > 0$*

*ii.* $\mathbf{c}\mathbf{c}^T \in$ Ext$(\mathcal{C}_m)$ *where $\mathbf{c} \in \mathbb{R}^m \backslash (\mathbb{R}_+^m \cup (-\mathbb{R}_+^m))$*

*iii.* $PDQDP \in$ Ext$(\mathcal{C}_m)$ *if and only if $Q \in$ Ext$(\mathcal{C}_m)$, where $P$ is a permutation matrix and $D$ is a diagonal matrix with $d_{ii} > 0$ for all $i$.*

*Proof.* For a proof see [52, 82]. □

Moreover the extreme rays of the set of copositive matrices $\{Q = (q_{ij}) \in \mathcal{C}_m : q_{ij} \in \{-1, 0, 1\}, q_{ii} = 1, \ \forall \ i, j\}$ are discussed in [93]. In the case of $5 \times 5$ matrices a complete characterization of extreme rays of $\mathcal{C}_5$ is provided by [89]. But it is still an open question whether there is an explicit characterization of the extreme rays of the copositive cone in general.

For the copositive cone it is well known that the interior consists of the set $\mathcal{C}_m^+$ of so-called *strictly copositive matrices* (see e.g. [37, Lemma 2.3],[12, Chapter 1, Section 2]) defined by,

$$\mathcal{C}_m^+ := \left\{ Q \in \mathcal{C}_m : \mathbf{v}^T Q \mathbf{v} = 0 \text{ implies } \mathbf{v} = \mathbf{o} \right\}, \tag{2.6}$$

that is $\mathcal{C}_m^+ = \text{int}(\mathcal{C}_m)$.

As mentioned earlier the set of all positive semidefinite matrices forms a cone which is contained in the cone of copositive matrices, i.e., $\mathcal{S}_m^+ \subseteq \mathcal{C}_m$. The set of all nonnegative matrices, denoted by $\mathcal{N}_m$, is also contained in $\mathcal{C}_m$. So clearly also $\mathcal{N}_m + \mathcal{S}_m^+ \subseteq \mathcal{C}_m$ holds . For $m \leq 4$ this inclusion turns into an equality [50], but for $m \geq 5$ the inclusion is strict. The following is the well known counter example.

**Example 2.19** ([50, 58]). *Consider the so-called Horn-matrix [50],*

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

*Let* $\mathbf{v} \in \mathbb{R}_+^5$. *We can write,*

$$\begin{aligned} \mathbf{v}^T H \mathbf{v} &= (v_1 - v_2 + v_3 + v_4 - v_5)^2 + 4v_2v_4 + 4v_3(v_5 - v_4) \\ &= (v_1 - v_2 + v_3 - v_4 + v_5)^2 + 4v_2v_5 + 4v_1(v_4 - v_5) \end{aligned}$$

*If* $v_5 \geq v_4$ *then* $\mathbf{v}^T H \mathbf{v} \geq 0$ *follows from the first expression. If* $v_5 \leq v_4$ *then* $\mathbf{v}^T H \mathbf{v} \geq 0$ *is obtained from the second expression. Note that* $H \notin \mathcal{S}_m^+$ *and* $H \notin \mathcal{N}_m$. *Moreover, the matrix* $H$ *cannot be decomposed as the sum of a nonnegative and a positive semidefinite matrix. This follows from* $\mathcal{S}_m^+ \subseteq \mathcal{C}_m, \mathcal{N}_m \subseteq \mathcal{C}_m$ *and the fact that the matrix* $H$ *is in* $\text{Ext}(\mathcal{C}_5)$*(cf. [93]).*

In view of $\mathcal{N}_m + \mathcal{S}_m^+ \subseteq \mathcal{C}_m$, it is interesting to know the relationship between

$\mathcal{N}_m + \mathcal{S}_m^+$ and the interior of the copositive cone. It is well known that neither $\text{int}(\mathcal{C}_m) \subseteq \mathcal{N}_m + \mathcal{S}_m^+$ nor $\text{int}(\mathcal{C}_m) \supseteq \mathcal{N}_m + \mathcal{S}_m^+$ holds true.

**Example 2.20.**

$$Q := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{\in \mathcal{S}_2} + \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\in \mathcal{N}_2}$$

*but the matrix is not in the interior of the copositive cone since $e_2^T Q e_2 = 0$.*

For recent results and a discussion on the geometry of the copositive cone we refer the interested reader to [52] and the references therein.

## 2.3 Completely Positive Cone

In this section we will briefly consider the completely positive cone. Here, we will confine ourself to a characterization of complete positivity of a matrix and known results on the cp-rank. The last subsection will describe some results on the extreme rays and the interior of the completely positive cone. The set of all $m \times m$ completely positive matrices generate a closed, convex, non polyhedral and full dimensional cone. Recall that it is called the cone of completely positive matrices and denoted by $\mathcal{C}_m^*$ (cf. Definition 1.6). The matrices in $\mathcal{C}_m^*$ can also be written as a sum of diadic products of rank one matrices,

$$\mathcal{C}_m^* = \left\{ A \in \mathcal{S}_m : A = \sum_{k=1}^{N} \mathbf{b}_k \mathbf{b}_k^T \text{ with } \mathbf{b}_k \in \mathbb{R}_+^m, N \in \mathbb{N} \right\} \tag{2.7}$$

It is interesting to note that the span of the columns of the matrix $A$ coincides with the span of the decomposition vectors $\mathbf{b}_i$.

**Lemma 2.21** ([13]). *Let $A \in \mathcal{C}_m^*$ and $A = BB^T = \sum_{i=1}^{k} \mathbf{b}_i \mathbf{b}_i^T$ then*

$$\text{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_m\} = \text{Span}\{\mathbf{b}_1, \cdots, \mathbf{b}_k\}$$

*where $\mathbf{a}_1, \cdots, \mathbf{a}_m$ and $\mathbf{b}_1, \cdots, \mathbf{b}_k$ are the columns of $A$ and $B$ respectively.*

Recall that the copositive cone and the completely positive cone are dual to each other (in Lemma 2.2, put $K = \mathbb{R}_+^m$).

A necessary condition for a matrix to be completely positive is that the matrix should be nonnegative and positive semidefinite. The set of all nonnegative positive semidefinite matrices is known as the set of *doubly nonnegative matrices* and denoted by $DNN_m := \mathcal{S}_m^+ \cap \mathcal{N}_m$. For $m \leq 4$ it is well known that (see [13]),

$$A \in \mathcal{C}_m^* \quad \text{if and only if} \quad A \in DNN_m \, . \tag{2.8}$$

Hence checking if a matrix of order four or less is completely positive amounts to checking if the matrix is nonnegative and positive semidefinite. But for the matrices of order greater than four this is not true in general,

**Example 2.22.**

$$A = \begin{bmatrix} 1 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & \frac{3}{4} & 0 \\ 0 & 0 & \frac{3}{4} & 1 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

*It is clear that $A \in \mathcal{N}_m$, also $A \in S_m^+$ since,*

$$\mathbf{v}^T A \mathbf{v} = \left(\frac{1}{2}v_1 + v_2 + \frac{1}{2}v_3\right)^2 + \left(\frac{1}{2}v_1 + \frac{1}{2}v_4 + v_5\right)^2 + \frac{1}{2}\left(v_1 - \frac{1}{2}v_3 - \frac{1}{2}v_4\right)^2 + \frac{5}{8}\left(v_3 + v_4\right)^2$$

*But $A$ is not completely positive, since $\langle A, H \rangle = -\frac{1}{2}$, where $H \in \mathcal{C}_5$ is the Horn matrix given in Example 2.19 (cf. Definition 1.3).*

Testing if a matrix is completely positive is an NP-hard problem [51]. But for some classes of matrices checking complete positivity is easy. For example every diagonally dominant matrix (see Definition 2.30) is well known to be completely positive [13, Theorem 2.5](see also (2.9)). Another example is the class of *binary matrices* which are completely positive if and only if they are positive semidefinite [107, Corollary 1]. For certain specially structured sparse matrices Dickinson and Dür have been able to formulate a linear time algorithm for testing complete positivity [54].

The following characterization of completely positive matrices is recursive, in the sense that it depends on the complete positivity of smaller matrices along with some other conditions (see Lemma 2.3 for the corresponding result for copositive matrices).

**Theorem 2.23.** *Let $A \in S_m$ be written in block form,*

$$A = \begin{pmatrix} a & \mathbf{v}^T \\ \mathbf{v} & V \end{pmatrix}$$

*then $A$ is completely positive if and only if $V = CC^T$ for some $C \in \mathbb{R}_+^{(m-1)\times n}$ (i.e. $V$ is completely positive) and there exists a nonnegative vector $\mathbf{w}$ such that $\mathbf{v} = C\mathbf{w}$ and $a = \mathbf{w}^T\mathbf{w}$.*

*Proof.* See [13, Theorem 2.16]. □

The smallest value of $N$ for which the factorization (2.7) of the matrix $A$ is possible is called the *CP-rank* of the matrix and denoted by $CP\text{-}rank(A)$.

By Lemma 2.21 the CP-rank of a completely positive matrix is always greater than or equal to the rank of the matrix. For the case of matrices of order three or less the CP-rank is exactly equal to the rank of the matrix [13, Theorem 3.2]. For general $m \times m$ matrices the following is known about the CP-rank.

**Theorem 2.24.** *Let $A \in \mathcal{C}_m^*$ and $r := rank(A)$,*

*i. if $r \geq 2$ then it holds:*

$$CP\text{-}rank(A) \leq \frac{r(r+1)}{2} - 1$$

*ii. if $r \geq 1$ and there exists a nonsingular $r \times r$ principal submatrix of $A$ with $N$ zeros above the diagonal, then*

$$CP\text{-}rank(A) \leq \frac{r(r+1)}{2} - N$$

*Proof.* For a proof of $i.$ see, [13, Theorem 3.4] or [84, 138], for a proof of $ii.$ see [138] or [13, Theorem 3.5]. □

For recent results on CP-rank of a completely positive matrix the interested reader is referred to [139, Corollary 5.1].

In [55], the following bound on the CP-rank of completely positive matrices is conjectured.

**Conjecture 2.25.** *If $A \in \mathcal{C}_m^*,\ m \geq 4$ then $CP\text{-}rank(A) \leq \left\lfloor \frac{m^2}{4} \right\rfloor$.*

**Definition 2.26** (M-Matrix). Let $A \in \mathbb{R}^{m \times m}$ then $A$ is called an M-matrix if $A$ can be expressed in the form $A = sI - B$, where $B = (b_{ij})$ with $b_{ij} \geq 0$, for all $1 \leq i, j \leq m$, and $s \geq \rho(B)$, the maximum of the moduli of the eigenvalues of $B$.

**Definition 2.27** (Comparison Matrix). Let $A \in \mathbb{R}^{m \times m}$ and,

$$(M(A))_{ij} := \begin{cases} |a_{ij}| & if\ i = j \\ -|a_{ij}| & otherwise \end{cases}$$

then $M(A)$ is called the *comparison matrix* of $A$.

If the comparison matrix of an $A \in \mathcal{N}_m$ is an M-matrix then $CP\text{-}rank(A) \leq \left\lfloor \frac{m^2}{4} \right\rfloor$ (for details see [55]). The Conjecture 2.25 is also proved for the matrices associated with the so-called cycle free completely positive graphs [13]. A proof of the Conjecture 2.25 for the case of $5 \times 5$ matrices is given in [140]. Moreover note that, for every even $m$ there exists a matrix with CP-rank $\frac{m^2}{4}$ as the following proposition, taken from [97], suggests,

**Proposition 2.28** ([97]). *For any even $m = 2n$ there exists an $m \times m$ matrix with CP-rank $\left\lfloor \frac{m^2}{4} \right\rfloor$.*

*Proof.* Let $E$ be the $n \times n$ all-one matrix, $I$ is the $n \times n$ identity matrix and $e_i$ is its $i^{th}$ column. Then consider the matrix,

$$A = \begin{pmatrix} nI & E \\ E & nI \end{pmatrix}$$

The matrix $A$ has a factorization,

$$A = \sum_{i,j=1}^{n} \begin{pmatrix} e_i \\ e_j \end{pmatrix} \begin{pmatrix} e_i \\ e_j \end{pmatrix}^T$$

Clearly the above decomposition contains $n^2 = \frac{m^2}{4}$ matrices. Note that any vector in a decomposition $A = \sum_{i,j=1}^{n} \begin{pmatrix} \mathbf{b}_i \\ \mathbf{c}_j \end{pmatrix} \begin{pmatrix} \mathbf{b}_i \\ \mathbf{c}_j \end{pmatrix}^T$ of the matrix $A$ will be of the form $\begin{pmatrix} \mathbf{b} \\ \mathbf{c} \end{pmatrix}$, $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n_+$ such that at most one element of $\mathbf{b}$ and of $\mathbf{c}$ is positive. Otherwise $A$ would have a nonzero element $a_{ij}, i \neq j, i, j \in \{1, \ldots, n\}$ or $i, j \in \{n+1, \ldots, m\}$. So the decomposition given above is minimal, that is $CP\text{-}rank(A) = n^2$. □

Note that for the matrix $A$ in the above proposition $M(A)$ is an M-matrix since,

$$M(A) = \begin{pmatrix} nI & -E \\ -E & nI \end{pmatrix} = n \begin{pmatrix} I & O \\ O & I \end{pmatrix} - \begin{pmatrix} O & E \\ E & O \end{pmatrix}.$$

Moreover it can be easily verified that $n$ is the largest eigenvalue of $\left(\begin{smallmatrix} O & E \\ E & O \end{smallmatrix}\right)$. It has been recently proven that Conjecture 2.25 is false for matrices of order from seven to eleven, for details see [31].

### 2.3.1 Interior and Extreme Rays

In this subsection we will briefly survey results on the interior and extreme rays of the completely positive cone. We will provide characterizations of the interior of the completely positive cone. We will also prove that every positive diagonally dominant matrix belongs to the interior of the completely positive cone.

In contrast to the copositive cone, an easy characterization for the interior of the completely positive cone $\mathcal{C}_m^*$ is not known. However, from $\mathcal{C}_m^* \subseteq \mathcal{S}_m^+ \cap \mathcal{N}_m$ we have,

$$\text{int}(\mathcal{C}_m^*) \subseteq \text{int}(\mathcal{S}_m^+) \cap \text{int}(\mathcal{N}_m)$$

So a necessary condition for a matrix $A \in \mathcal{C}_m^*$ to be in the interior of the completely positive cone is that the matrix $A$ is positive definite, i.e., $A \in \mathcal{S}_m^{++}$.

Dür and Still [59] have given a characterization of the interior of the completely positive cone. Dickinson [53] has added other characterizations. Here are these results.

**Theorem 2.29.** *The interior of the completely positive cone is given by,*

$$\begin{aligned} \text{int}(\mathcal{C}_m^*) &= \left\{ \sum_{k=1}^{N} \mathbf{b}_k \mathbf{b}_k^T : \begin{array}{l} \mathbf{b}_k \in \mathbb{R}_+^m \ \forall\, k = 1, \cdots, N \\ \mathbf{b}_k \in \mathbb{R}_{++}^m \ \forall\, k = 1, \cdots, m \\ \text{Span}\{\mathbf{b}_1, \cdots, \mathbf{b}_N\} = \mathbb{R}^m \end{array} \right\} \\ &= \left\{ \sum_{k=1}^{N} \mathbf{b}_k \mathbf{b}_k^T : \begin{array}{l} \mathbf{b}_1 \in \mathbb{R}_{++}^m, \mathbf{b}_k \in \mathbb{R}_+^m \ \forall\, k = 2, \cdots, N \\ \text{Span}\{\mathbf{b}_1, \cdots, \mathbf{b}_N\} = \mathbb{R}^m \end{array} \right\} \end{aligned}$$

*where* $\mathbb{R}_{++}^m := \{\mathbf{b} \in \mathbb{R}^m : b_i > 0,\ \forall\, i = 1, \ldots, m\}$.

*Proof.* For a proof see [59] and [53]. ☐

The characterization given in Theorem 2.29 provides a way to check if a matrix is in the interior of the completely positive cone or not. However this requires a completely positive matrix to be decomposed in a certain way. Recently Zhou and Fan [156] have presented an algorithm which when applied to a completely positive matrix $A \in \mathcal{C}_m^*$ returns the decomposition of the matrix in the form given in Theorem 2.29, if the matrix $A \in \text{int}(\mathcal{C}_m^*)$.

Here we will consider some matrices which belong to the interior of the completely positive cone. As a first example consider the matrix $A = I + \mathbf{b}\mathbf{b}^T$, where **b** is a positive vector. Then clearly $A$ is completely positive. Take, $t = \frac{\sqrt{b^T b+1}-1}{b^T b}$, then

$$A = (I + t\mathbf{b}\mathbf{b}^T)(I + t\mathbf{b}\mathbf{b}^T)^T = I + (2t + t^2\mathbf{b}^T\mathbf{b})\mathbf{b}\mathbf{b}^T = I + \mathbf{b}\mathbf{b}^T$$

Since $rank(A) = m$, and there are $m$ columns in the matrix $(I + t\mathbf{b}\mathbf{b}^T)$, it follows $CP\text{-}rank(A) = m$ [138] and by Theorem 2.29 $A \in \text{int}(\mathcal{C}_m^*)$.

Before presenting our next example we shall define the set of *diagonally dominant* matrices.

**Definition 2.30** (Diagonally Dominant)**.** Let $A = (a_{ij}) \in \mathcal{S}_m$ and

$$u_i := |a_{ii}| - \sum_{j=1, i\neq j}^{m} |a_{ij}|$$

then if $u_i \geq 0$ for all $i \in \mathcal{U}$ the matrix is called *diagonally dominant.*

It is well known that every nonnegative diagonally dominant matrix is completely positive. Indeed it is not difficult to verify that such a matrix can be decomposed as follows (see also [100]),

$$A = \sum_{i=1}^{m} u_i e_i e_i^T + \sum_{i=1, j=i+1}^{m} a_{ij}(e_i + e_j)(e_i + e_j)^T \,. \tag{2.9}$$

In the next theorem we will show that every positive diagonally dominant matrix belongs to the interior of the completely positive cone. Note that the positive diagonally dominant matrices of order two may be singular. Take for example the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. But for matrices of order $m > 2$, the positive diagonally dominant matrices are always nonsingular.

**Lemma 2.31.** *Let $A \in \mathcal{S}_m, m > 2$, be diagonally dominant and positive, then $A$ is nonsingular.*

*Proof.* It is well known that a diagonally dominant matrix with at least one $i \in \mathcal{U}$ such that $u_i > 0$, is nonsingular (see e.g. [94, Corollary 7.2.3]). Now consider the case when $u_i = 0$ for all $i \in \mathcal{U}$ and let $s \le a_{ij} \le l$. Then the following bounds on the eigenvalues of $A$ are provided in [90, Lemma 7.1]:

$$(m-2)s \le \lambda_i \le (m-2)l \text{ for } 1 \le i \le m-1, \text{ and } 2(m-1)s \le \lambda_m \le 2(m-1)l.$$

Hence $A$ is nonsingular. □

**Theorem 2.32.** *Let $A \in \mathcal{S}_m, m > 2$, be diagonally dominant and positive, then $A \in \text{int}(\mathcal{C}_m^*)$.*

*Proof.* In view of Lemma 2.31 it is clear that $A$ is of full rank. Let us consider $l := \min\{1, \min_{i,j=1..m}\{a_{ij}\}\}$, and define $\mathbf{b} = [l, \cdots, l]^T$, then it is not difficult to see that $B := A - \mathbf{b}\mathbf{b}^T$ is nonnegative. Now we will show that $B$ is also diagonally dominant.

$$\begin{aligned} b_{ii} - \sum_{i \neq j} b_{ij} &= a_{ii} - l^2 - \sum_{i \neq j} (a_{ij} - l^2) \\ &= a_{ii} - \sum_{i \neq j} a_{ij} + (m-2)l^2 \ge 0 \, . \end{aligned}$$

Hence $B$ is diagonally dominant and a decomposition of $A$ is given by

$$A = \mathbf{b}\mathbf{b}^T + \sum_{i=1}^{m} (b_{ii} - \sum_{i \neq j} b_{ij}) e_i e_i^T + \sum_{j>i}^{m} b_{ij}(e_i + e_j)(e_i + e_j)^T \, .$$

Since $\mathbf{b} \in \mathbb{R}^m_{++}$, by Theorem 2.29 $A \in \text{int}(\mathcal{C}_m^*)$. □

Here we would like to mention that the proof of Theorem 2.32 is independently obtained by [139, Theorem 2.2]. The following is an immediate corollary of the above theorem.

**Corollary 2.33.** *Let $A \in \mathcal{N}_m$ be positive and let its comparison matrix, $M(A)$, be an M-matrix then, $A \in \text{int}(\mathcal{C}_m^*)$.*

*Proof.* It is well known that if $M(A)$ is an M-matrix then there exists a positive diagonal matrix $D$ such that $DAD$ is diagonally dominant (see e.g. [55, page 305]). Since $D$ is positive, so is $DAD$. Let $DAD = BB^T$ such that $B \in \mathbb{R}^{m \times n}_+$ has at least one positive column. Note that such a decomposition exists owing to Theorem 2.32. Then clearly $A = D^{-1}B(D^{-1}B)^T$ will have at least one column positive, and by Theorem 2.29 $A \in \text{int}(\mathcal{C}_m^*)$. ☐

The extreme rays of the completely positive cone are very well known and described below.

**Lemma 2.34.** *The extreme rays of the cone $\mathcal{C}_m^*$ are the rank one matrices $U = \mathbf{u}\mathbf{u}^T$, where $\mathbf{u} \in \mathbb{R}_{++}^m$.*

*Proof.* See e.g. [82, Theorem 3.1 ]. ☐

# 3

# The Standard Quadratic Programming Problem

In the optimization literature quadratic programming (QP) normally refers to the set of problems with linear constraint(s) and quadratic objective function. In this chapter we will focus on a special instance of QP namely the standard quadratic programming problem (StQP). Our special interest in this program stems from the fact that it just represents the feasibility test in copositive programming. In sections two and three we will provide optimality conditions and a stability analysis for StQP, respectively. The notion of strict local maximizer of StQP is related, as we will see, to the concept of evolutionarily stable strategies (ESS) from population genetics. In section four we will provide a brief survey of evolutionarily stable strategies while focusing on the maximum number of ESS which can coexist for a certain matrix. The fifth section deals with vector iterations which are related to StQP and a similar program. In the last section we will look at some genericity results for the strict local maximizers of StQP.

## 3.1 Introduction

The standard quadratic programming problem can be written in the following form,

$$(StQP) \quad \max \quad q(\mathbf{v}) := \frac{1}{2}\mathbf{v}^T Q\mathbf{v} \quad s.t. \quad \mathbf{v} \in \Delta_m := \left\{ \mathbf{v} \in \mathbb{R}^m_+ : \sum_{i=1}^{m} v_i = 1 \right\}.$$

where $Q \in \mathcal{S}_m$. Standard quadratic programming is very well studied due to its vast applications in the areas of resource allocation problems [95], portfolio optimization problems [111], maximum weight clique problem [70, 115] and population genetics [35]. StQP is known to be NP-hard [115, 150]. In Bomze et al [29] a good survey of algorithms/methods for solving/approximating StQP is provided. There has been much work done on theory and algorithms for StQP (see e.g. [25]). However, the question of stability of StQP is not answered satisfactorily in the literature. In this chapter we will provide a stability analysis for StQP along with a characterization for a point to be a strict local maximizer, a review on evolutionarily stable strategy, vector iterations related to StQP and some genericity results.

Before proceeding further, note that $(StQP)$ can be used to formulate the feasibility criteria for programs over the cone of copositive matrices. In order to see this consider the copositive program in subsection 1.3.3 where it is required to check if the matrix $F(\mathbf{x}) \in \mathcal{S}_m$ is copositive. This can be done by solving $(StQP)$ with $Q := -F(\mathbf{x})$ and checking if $val(StQP)$ is non-positive.

## 3.2 Optimality Conditions

In this section the second order necessary and sufficient conditions for strict local maximizers of $(StQP)$ are formulated. For this purpose first we define the Lagrange function associated with $(StQP)$,

$$L(\mathbf{v}, \lambda, \mu) = q(\mathbf{v}) - \lambda(\mathbf{e}^T\mathbf{v} - 1) + \mu^T\mathbf{v}$$

where $\lambda \in \mathbb{R}, \mu \in \mathbb{R}^m_+$. Then the KKT conditions for $\mathbf{v} \in \Delta_m$ read,

$$\nabla_{\mathbf{v}} L(\mathbf{v}, \lambda, \mu) = Q\mathbf{v} - \lambda\mathbf{e} + \sum_{i \notin R(\mathbf{v})} \mu_i e_i = 0, \ \mu_i \geq 0, \tag{3.1}$$

where $R(\mathbf{v})$ is the support of the vector **v** defined below,

**Definition 3.1** (Support of a Vector). Let $\mathbf{v} \in \mathbb{R}^m_+$, then

$$R(\mathbf{v}) := \{i : v_i > 0\}. \tag{3.2}$$

Note that the constraints in $(StQP)$ are linear and the so called *linear independence constraint qualification* (see e.g. [63, page 280]) is satisfied implying that the Lagrange multipliers are uniquely determined. Let $\overline{\mathbf{v}}, \lambda, \mu$ satisfy the KKT conditions (3.1) then as usual $\overline{\mathbf{v}}, \mu$ is said to satisfy *strict complementarity* if $\overline{v}_i = 0$ implies $\mu_i > 0$. Define also for $\overline{\mathbf{v}}, \mu$ satisfying the KKT conditions,

$$\widetilde{S}(\overline{\mathbf{v}}) := \{i : \mu_i = 0\} \tag{3.3}$$

then clearly we have $R(\overline{\mathbf{v}}) \subseteq \widetilde{S}(\overline{\mathbf{v}})$. Moreover, strict complementarity is equivalent to $R(\overline{\mathbf{v}}) = \widetilde{S}(\overline{\mathbf{v}})$.

A vector $\mathbf{d} \in \mathbb{R}^m$ is said to be a *feasible direction* with respect to a feasible point $\overline{\mathbf{v}}$ if there exists $\alpha > 0$ such that $\overline{\mathbf{v}} + \alpha\mathbf{d}$ is also feasible. The conditions for optimality in nonlinear programming are generally described in terms of feasible directions or more precisely with the help of the *cone of critical directions* [63, Theorem 12.6]. For $(StQP)$ the cone of critical directions can be written as follows,

$$C(\overline{\mathbf{v}}) := \{\mathbf{d} \in \mathbb{R}^m : (Q\overline{\mathbf{v}})^T\mathbf{d} \geq 0, \mathbf{e}^T\mathbf{d} = 0, d_i \geq 0 \ \forall i \notin R(\overline{\mathbf{v}})\} \tag{3.4}$$

**Remark 3.2.** *Let $\overline{\mathbf{v}} \in \Delta_m, \mu \in \mathbb{R}^m_+$ satisfy the KKT conditions* (3.1) *and let $\mathbf{d} \in C(\overline{\mathbf{v}})$, then*

$$\begin{aligned} 0 \leq (Q\overline{\mathbf{v}})^T\mathbf{d} &= \left(\lambda\mathbf{e} - \sum_{i \notin R(\overline{\mathbf{v}})} \mu_i e_i\right)^T \mathbf{d} = - \sum_{i \notin R(\overline{\mathbf{v}})} \mu_i d_i \\ &= - \sum_{i \notin \widetilde{S}(\overline{\mathbf{v}})} \mu_i d_i - \sum_{i \in \widetilde{S}(\overline{\mathbf{v}}) \setminus R(\overline{\mathbf{v}})} \mu_i d_i \\ &= - \sum_{i \notin \widetilde{S}(\overline{\mathbf{v}})} \mu_i d_i \leq 0 \end{aligned}$$

*Since $\mu_i > 0$ for all $i \notin \widetilde{S}(\overline{\mathbf{v}})$ and $d_i \geq 0$ for all $i \notin \widetilde{S}(\overline{\mathbf{v}})$ ($d_i \geq 0$ for all $i \notin R(\overline{\mathbf{v}})$ implies $d_i \geq 0$ for $i \notin \widetilde{S}(\overline{\mathbf{v}})$ ) ,hence from the above we conclude that $d_i = 0$ for all*

$i \notin \widetilde{S}(\overline{\mathbf{v}})$. *So for a KKT point* $\overline{\mathbf{v}} \in \Delta_m$ *the cone of critical direction* (3.4) *reduces to,*

$$C(\overline{\mathbf{v}}) := \{\mathbf{d} \in \mathbb{R}^m : \mathbf{e}^T\mathbf{d} = 0, d_i = 0\ \forall i \notin \widetilde{S}(\overline{\mathbf{v}}), d_i \geq 0\ \forall i \in \widetilde{S}(\overline{\mathbf{v}})\backslash R(\overline{\mathbf{v}})\} \quad (3.5)$$

We also define the order of maximizer for $(StQP)$,

**Definition 3.3** (Order of Maximizer). A feasible point $\overline{\mathbf{v}} \in \Delta_m$ is a maximizer of $(StQP)$ of order $p > 0$, if with some $\gamma > 0, \varepsilon > 0$ the following holds,

$$q(\overline{\mathbf{v}}) \geq q(\mathbf{v}) + \gamma\|\mathbf{v} - \overline{\mathbf{v}}\|^p \quad \forall\ \mathbf{v} \in \Delta_m,\ \|\mathbf{v} - \overline{\mathbf{v}}\| < \varepsilon\ . \quad (3.6)$$

In the above definition $\|.\|$ denotes the Euclidean norm. Define also for $\varepsilon > 0$, the $\varepsilon$-neighbourhood of the point $\overline{\mathbf{v}} \in \mathbb{R}^m$ by $N_\epsilon(\overline{\mathbf{v}}) := \{\mathbf{v} \in \mathbb{R}^m : ||\mathbf{v} - \overline{\mathbf{v}}|| \leq \varepsilon\}$.

Although in the general case of nonlinear programming there is a gap between necessary and sufficient optimality conditions, for the case of $(StQP)$ there is no gap,

**Theorem 3.4.** *Let* $\overline{\mathbf{v}} \in \Delta_m$ *then,*

*i.* $\overline{\mathbf{v}}$ *is a strict local maximizer of* $(StQP)$ *if and only if* $\overline{\mathbf{v}}$ *satisfies the KKT conditions* (3.1) *and* $\mathbf{d}^TQ\mathbf{d} < 0$ *for all* $\mathbf{d} \in C(\overline{\mathbf{v}})\backslash\{\mathbf{o}\}$.

*ii.* $\overline{\mathbf{v}}$ *is a local maximizer of* $(StQP)$ *if and only if* $\overline{\mathbf{v}}$ *satisfies the KKT conditions* (3.1) *and* $\mathbf{d}^TQ\mathbf{d} \leq 0$ *for all* $\mathbf{d} \in C(\overline{\mathbf{v}})$.

*iii. If* $\overline{\mathbf{v}}$ *is a strict local maximizer of* $(StQP)$ *then with some* $\gamma, \varepsilon > 0$ *we have,*

$$q(\overline{\mathbf{v}}) - q(\mathbf{v}) \geq \gamma\|\mathbf{v} - \overline{\mathbf{v}}\|^2 \quad \forall\ \mathbf{v} \in N_\varepsilon(\overline{\mathbf{v}}) \cap \Delta_m$$

*that is,* $\overline{\mathbf{v}}$ *is also a strict local maximizer of order two.*

*Proof. i.* $\Rightarrow$ Let $\overline{\mathbf{v}}$ be a strict local maximizer. For the sake of contradiction we assume that there exists a $\mathbf{o} \neq \mathbf{d} \in C(\overline{\mathbf{v}})$ such that $\mathbf{d}^TQ\mathbf{d} \geq 0$. Now for a small $\alpha > 0$, it is clear that $\overline{\mathbf{v}} + \alpha\mathbf{d} \in \Delta_m$ then

$$\begin{aligned}(\overline{\mathbf{v}} + \alpha\mathbf{d})^TQ(\overline{\mathbf{v}} + \alpha\mathbf{d}) - \overline{\mathbf{v}}^TQ\overline{\mathbf{v}} &= 2\alpha\overline{\mathbf{v}}^TQ\mathbf{d} + \alpha^2\mathbf{d}^TQ\mathbf{d} \\ &\geq \alpha^2\mathbf{d}^TQ\mathbf{d} \geq 0\ .\end{aligned}$$

This holds since $\mathbf{d} \in C(\overline{\mathbf{v}})$ implies $\overline{\mathbf{v}}^TQ\mathbf{d} \geq 0$. Since $\alpha > 0$, we arrive at a contradiction that $\overline{\mathbf{v}}$ is not a strict local maximizer.
$\Leftarrow$ We directly prove that the condition $\mathbf{d}^TQ\mathbf{d} < 0$, for all $\mathbf{o} \neq \mathbf{d} \in C(\overline{\mathbf{v}})$ implies that $\overline{\mathbf{v}}$ is a strict local maximizers of order two (see *iii.*). So let $\mathbf{d}^TQ\mathbf{d} < 0$ for all $\mathbf{o} \neq \mathbf{d} \in C(\overline{\mathbf{v}})$ but assume $\overline{\mathbf{v}} \in \Delta_m$ is not a strict local maximizer of order two.

Then there exists an infinite sequence of feasible points $\mathbf{v}_k \to \overline{\mathbf{v}}$ satisfying $\mathbf{v}_k \neq \overline{\mathbf{v}}$ and,

$$q(\mathbf{v}_k) - q(\overline{\mathbf{v}}) \geq o\left(\|\mathbf{v}_k - \overline{\mathbf{v}}\|^2\right).$$

Now we write $\mathbf{v}_k = \overline{\mathbf{v}} + t_k \mathbf{d}_k$ where $\mathbf{d}_k \in \mathbb{R}^m, ||\mathbf{d}_k|| = 1, t_k > 0, t_k \to 0$ as $\mathbf{v}_k \to \overline{\mathbf{v}}$. The sequence $\mathbf{d}_k$ has a subsequence converging to some vector $\mathbf{d}$, $||\mathbf{d}|| = 1$, i.e., without loss of generality $\mathbf{d}_k \to \mathbf{d}$. The existence of such a subsequence is evident since $\mathbf{d}_k$ forms a sequence in the compact set of all vectors with unit norm. First we will show that $\mathbf{d} \in C(\overline{\mathbf{v}})$. For this we consider

$$o\left(t_k^2\right) \leq q(\mathbf{v}_k) - q(\overline{\mathbf{v}}) = t_k (Q\overline{\mathbf{v}})^T \mathbf{d}_k + \frac{1}{2} t_k^2 \mathbf{d}_k^T Q \mathbf{d}_k . \tag{3.7}$$

Divide (3.7) by $t_k$ and take $k \to \infty$ to arrive at $(Q\overline{\mathbf{v}})^T \mathbf{d} \geq 0$. Note that from $\mathbf{v}_k \in \Delta_m$ and $0 \leq (\mathbf{v}_k)_i = \overline{v}_i + (\mathbf{d}_k)_i$, we can conclude that $(\mathbf{d}_k)_i \geq 0$ for all $i \notin R(\overline{\mathbf{v}})$. Note also that $\mathbf{v}_k \in \Delta_m$ implies $\mathbf{e}^T \mathbf{v}_k = \mathbf{e}^T(\overline{\mathbf{v}} + t_k \mathbf{d}_k) = 1$ and $\mathbf{e}^T \mathbf{d}_k = 0$. For $k \to \infty$ this leads to $\mathbf{e}^T \mathbf{d} = 0$. Hence we can conclude that $\mathbf{d} \in C(\overline{\mathbf{v}})$.

From (3.7), the KKT conditions and the observation that $(\mathbf{d}_k)_i \geq 0$ for all $i \notin R(\overline{\mathbf{v}})$ and $\mu \in \mathbb{R}^m_+$ we obtain, $\sum_{i \notin R(\overline{\mathbf{v}})} \mu_i (\mathbf{d}_k)_i \geq 0$. So we have,

$$\begin{aligned}
o\left(t_k^2\right) &\leq t_k (Q\overline{\mathbf{v}})^T \mathbf{d}_k + \frac{1}{2} t_k^2 \mathbf{d}_k^T Q \mathbf{d}_k \\
&= t_k \Big(\lambda \mathbf{e} - \sum_{i \notin R(\overline{\mathbf{v}})} \mu_i e_i\Big)^T \mathbf{d}_k + \frac{1}{2} t_k^2 \mathbf{d}_k^T Q \mathbf{d}_k \\
&= \underbrace{-t_k \sum_{i \notin R(\overline{\mathbf{v}})} \mu_i (\mathbf{d}_k)_i}_{\geq 0} + \frac{1}{2} t_k^2 \mathbf{d}_k^T Q \mathbf{d}_k \\
&\leq \frac{1}{2} t_k^2 \mathbf{d}_k^T Q \mathbf{d}_k
\end{aligned}$$

Dividing by $t_k^2$ and letting $k \to \infty$ gives,

$$\mathbf{d}^T Q \mathbf{d} \geq 0$$

leading to a contradiction to the hypothesis and this also proves $iii$.
*ii.* Can be proven in a similar way. □

Here we would like to mention that, in the literature ( see e.g. [20, 22, 23, 24, 27]), the characterization for a strict local maximizer of StQP requires that strict complementarity holds while the characterization given in Theorem 3.4 does not impose such a condition.

In the literature the cone of critical directions is used for second order conditions. Second order necessary and sufficient conditions are also formulated in terms of so called Tangent spaces. Let $\bar{\mathbf{v}} \in \Delta_m$ satisfy the KKT conditions (3.1), then for the case of $(StQP)$ the tangent spaces are given by,

$$T(\bar{\mathbf{v}}) := \{\mathbf{d} \in \mathbb{R}^m : \mathbf{e}^T\mathbf{d} = 0, d_i = 0 \quad \forall i \notin R(\bar{\mathbf{v}})\} \tag{3.8}$$

$$T^+(\bar{\mathbf{v}}) := \{\mathbf{d} \in \mathbb{R}^m : \mathbf{e}^T\mathbf{d} = 0, d_i = 0 \quad \forall i \notin \widetilde{S}(\bar{\mathbf{v}})\}. \tag{3.9}$$

It can be readily verified that $T(\bar{\mathbf{v}}) \subseteq C(\bar{\mathbf{v}}) \subseteq T^+(\bar{\mathbf{v}})$. Moreover if strict complementarity holds at the KKT point $\bar{\mathbf{v}}$ then the three cones are equal.

**Corollary 3.5** ([24]). *Let $\bar{\mathbf{v}} \in \Delta_m$ satisfy the KKT conditions* (3.1)*. Let strict complementarity hold at $\bar{\mathbf{v}}$, i.e., $R(\bar{\mathbf{v}}) = \widetilde{S}(\bar{\mathbf{v}})$. Then $\bar{\mathbf{v}}$ is a strict local maximizer if and only if $\mathbf{d}^TQ\mathbf{d} < 0$ for all $\mathbf{d} \in T(\bar{\mathbf{v}})\setminus\{\mathbf{o}\}$.*

*Proof.* Follows immediately from Theorem 3.4. □

**Remark 3.6.** *As mentioned before, in general nonlinear programming, there is a gap between second order necessary and sufficient optimality conditions. But for the special case of $(StQP)$ it is shown that there is no gap. As a matter of fact our result can be seen as a special instance of a more general result which says that if the feasible set of the quadratic program is convex then there is no gap between second order necessary and sufficient conditions for optimality, for details see [32, Theorem 4].*

We have presented optimality conditions, now we will focus on another interesting result which in turn leads to conditions on the matrix such that the existence of a strict local maximizer is guaranteed. First we will provide some auxiliary results required in the proof.

**Definition 3.7** (Affine Subspace)**.** An *affine subspace* $W \subseteq \mathbb{R}^m$ is the translation of a subspace $V \subseteq \mathbb{R}^m$ by a vector $\mathbf{u}$, i.e.,

$$W = \{\mathbf{w} \in \mathbb{R}^m : \mathbf{w} = \mathbf{u} + \mathbf{v},\ \mathbf{v} \in V\}.$$

Moreover dim$(W)$ := dim$(V)$, where dim$(V)$ denotes the dimension of the space $V$.

**Definition 3.8** (Affine Hull)**.** For $S \subseteq \mathbb{R}^m$, the set of all affine combinations aff$(S)$ of $S$ is called *affine hull* of $S$, i.e.,

$$\text{aff}(S) = \left\{ \sum_{i=1}^{k} \alpha_i \mathbf{v}_i : \mathbf{v}_i \in S, \alpha_i \in \mathbb{R}, i = 1...k, \sum_{i=1}^{k} \alpha_i = 1, k \in \mathbb{N} \right\} .$$

It is not difficult to show that aff$(S)$ is the smallest affine space containing $S$.

**Definition 3.9** (Relative Interior)**.** Let $S$ be a convex set. A point $\mathbf{v} \in S$ is in the relative interior of $S$, denoted by rint$(S)$, if for any $\overline{\mathbf{v}} \in S$ there exists $\widetilde{\mathbf{v}} \in S$ and $0 < \lambda < 1$ such that

$$\mathbf{v} = \lambda \overline{\mathbf{v}} + (1 - \lambda) \widetilde{\mathbf{v}} \,.$$

The following lemma states that if a point $\overline{\mathbf{v}}$ belonging to the relative interior of a convex set is a local maximizer of the quadratic form, then it is a global maximizer on the affine hull of the convex set,

**Lemma 3.10.** *Let $\overline{\mathbf{v}}$ be a local maximizer of $q(\mathbf{v}) := \frac{1}{2}\mathbf{v}^T Q \mathbf{v}$ on a convex set $S \subseteq \mathbb{R}^m$ and $\overline{\mathbf{v}} \in$ rint$(S)$, then with aff$(S) = \overline{\mathbf{v}} + V$ we have:*

*i. For all $\mathbf{u} \in V$ it holds*

$$\overline{\mathbf{v}}^T Q \mathbf{u} = 0 \textit{ and } \mathbf{u}^T Q \mathbf{u} \leq 0$$

*and $\overline{\mathbf{v}}$ is a global maximizer of $q$ on aff$(S)$ .*

*ii. If moreover $q(\overline{\mathbf{v}}) = 0$, then $\overline{\mathbf{v}}$ is a global maximizer of $q(\mathbf{v})$ on* Span$\{\overline{\mathbf{v}}\} + V$*, i.e.,*

$$\mathbf{w}^T Q \mathbf{w} \leq 0 \quad \forall \quad \mathbf{w} \in \text{Span}\{\overline{\mathbf{v}}\} + V \,.$$

*Proof. i.* Since $\overline{\mathbf{v}} \in$ rint$(S)$, so for each $\mathbf{u} \in V, \|\mathbf{u}\| = 1$ there exists $\varepsilon > 0$ such that $\overline{\mathbf{v}} \pm \lambda \mathbf{u} \in S$ for all $0 \leq \lambda < \varepsilon$. Moreover $\overline{\mathbf{v}}$ is a local maximizer of $S$, hence,

$$\begin{aligned} q(\overline{\mathbf{v}} \pm \lambda \mathbf{u}) &\leq q(\overline{\mathbf{v}}) \\ (\overline{\mathbf{v}} \pm \lambda \mathbf{u})^T Q (\overline{\mathbf{v}} \pm \lambda \mathbf{u}) &\leq \overline{\mathbf{v}}^T Q \overline{\mathbf{v}} \\ \lambda^2 \mathbf{u}^T Q \mathbf{u} \pm 2\lambda \overline{\mathbf{v}}^T Q \mathbf{u} &\leq 0 \,. \end{aligned}$$

Take $\lambda > 0$ then,

$$\lambda \mathbf{u}^T Q \mathbf{u} \pm 2 \overline{\mathbf{v}}^T Q \mathbf{u} \leq 0 \,.$$

So, for $\lambda \to 0$ we obtain $\overline{\mathbf{v}}^T Q \mathbf{u} = 0$ which implies $\mathbf{u}^T Q \mathbf{u} \leq 0$. Since $\mathbf{u}^T Q \mathbf{u} \leq 0$ holds for all $\mathbf{u} \in V$, hence, $\overline{\mathbf{v}}$ is global maximizer on aff$(S)$.

*ii.* We consider part *i.* and obtain for any $\alpha \in \mathbb{R}$, $\mathbf{u} \in V$,

$$(\alpha\overline{\mathbf{v}} + \mathbf{u})^T Q(\alpha\overline{\mathbf{v}} + \mathbf{u}) = \alpha^2 \overline{\mathbf{v}}^T Q\overline{\mathbf{v}} + 2\alpha\overline{\mathbf{v}}^T Q\mathbf{u} + \mathbf{u}^T Q\mathbf{u} \leq 0 \,. \qquad \square$$

Lemma 3.10 is simple and straightforward, however this lemma has a number of consequences. One consequence is that if $\overline{\mathbf{v}}$ is a local maximizer of $q(\mathbf{v})$ on an open subset of $\mathbb{R}^m$, then $\overline{\mathbf{v}}$ is a global maximizer on $\mathbb{R}^m$. Moreover, in this case $Q$ is negative semidefinite. Another consequence of this lemma is that if some point is a local maximizer on a certain face of the simplex $\Delta_m$ and it also belong to the relative interior of the face, then the point is a global maximizer on the affine hull of the face. Before formally stating this result we will define the face of a convex set.

**Definition 3.11** (Face)**.** Let $S$ be a convex set and $F \subseteq S$. Then $F$ is called a face if for any $\mathbf{v}_1, \mathbf{v}_2 \in S$ it holds that $\lambda\mathbf{v}_1 + (1-\lambda)\mathbf{v}_2 \in F$ for some $0 < \lambda < 1$ implies $\mathbf{v}_1, \mathbf{v}_2 \in F$.

In other words the above definition says that a (convex) subset $F$ of a convex set $S$ is a face of $S$ if any line segment in $S$ with relative interior in $F$ has both end points in $F$ [131, page 162]. From the above definition, it is clear that with $J \subseteq \mathcal{U}$, where as usual $\mathcal{U} := \{1, \cdots, m\}$, the faces of $\Delta_m$ are given by,

$$\mathrm{fc}_J := \{\mathbf{v} \in \mathbb{R}^m_+ : \mathbf{e}^T\mathbf{v} = 1, v_j = 0, \ \forall\, j \notin J\}$$

Moreover, the faces of the standard simplex are itself standard simplices of lower dimensions.

**Corollary 3.12.** *Let $\overline{\mathbf{v}} \in \mathrm{int}(\mathrm{fc}_J)$ be a local maximizer of $(StQP)$ on a face $\mathrm{fc}_J$ then $\overline{\mathbf{v}}$ is global maximizer on aff$(\mathrm{fc}_J)$.*

*Proof.* Follows directly from Lemma 3.10. $\square$

**Theorem 3.13.** *Let $\overline{\mathbf{v}} \in \Delta_m$ be a non-strict local maximizer of $\mathbf{v}^T Q\mathbf{v}$ and $\overline{\mathbf{v}} \in \mathrm{rint}(\Delta_m)$, then $Q$ is singular.*

*Proof.* Since $\overline{\mathbf{v}}$ is a local maximizer, we have from Lemma 3.10,

$$\overline{\mathbf{v}}^T Q\mathbf{u} = 0, \quad \mathbf{u}^T Q\mathbf{u} \leq 0 \quad \forall\, \mathbf{u} \text{ such that } \mathbf{e}^T\mathbf{u} = 0 \tag{3.10}$$

Since $\overline{\mathbf{v}}$ is a non-strict local maximizer from Theorem 3.4 it follows that there exists a $\mathbf{o} \neq \mathbf{d} \in C(\overline{\mathbf{v}})$ such that $\mathbf{d}^T Q\mathbf{d} = 0$. Moreover, since $\overline{\mathbf{v}}$ is a local maximizer it satisfies the KKT conditions and from Remark 3.2 we get,

$$\overline{\mathbf{v}}^T Q\mathbf{d} = 0.$$

Now consider $\mathbf{u}$ such that $\mathbf{e}^T\mathbf{u} = 0$. Then by (3.10) for $\delta > 0$ (small)

$$\begin{aligned} (\mathbf{d} \pm \delta\mathbf{u})^T Q(\mathbf{d} \pm \delta\mathbf{u}) &\leq 0 \\ \Rightarrow \pm 2\delta\mathbf{u}^T Q\mathbf{d} + \delta^2\mathbf{u}^T Q\mathbf{u} &\leq 0 \end{aligned}$$

Dividing by $\delta > 0$ and taking the limit $\delta \to 0$ implies $\mathbf{u}^T Q\mathbf{d} = 0$. Hence we have a $\mathbf{d} \neq \mathbf{o}$ such that $\mathbf{e}^T\mathbf{d} = 0$ with $\mathbf{u}^T Q\mathbf{d} = 0$ for all $\mathbf{u}$ such that $\mathbf{e}^T\mathbf{u} = 0$ and $\overline{\mathbf{v}}^T Q\mathbf{d} = 0$. So, in view of $\overline{\mathbf{v}}^T\mathbf{e} = 1$, $Q\mathbf{d}$ is orthogonal to the whole $\mathbb{R}^m$, which implies, $Q\mathbf{d} = \mathbf{o}$, giving that $Q$ is singular. □

The following is an immediate corollary,

**Corollary 3.14.** *Let $\overline{\mathbf{v}} \in \Delta_m$ be a non-strict local maximizer with $R(\overline{\mathbf{v}}) = \widetilde{S}(\overline{\mathbf{v}})$ then $Q_{R(\overline{\mathbf{v}})}$ is singular.*

*Proof.* If $\overline{\mathbf{v}}$ is a non-strict local maximizer then it is not difficult to see that under the condition $R(\overline{\mathbf{v}}) = \widetilde{S}(\overline{\mathbf{v}})$, $\overline{\mathbf{v}}_{R(\overline{\mathbf{v}})}$ is a non-strict local maximizer with respect to $Q_{R(\overline{\mathbf{v}})}$ and $\overline{\mathbf{v}}_{R(\overline{\mathbf{v}})} \in \mathrm{fc}_{R(\overline{\mathbf{v}})}$. Then by Theorem 3.13, $Q_{R(\overline{\mathbf{v}})}$ is singular. □

## 3.3 Stability Analysis

In the previous section we have presented a characterization for a point $\overline{\mathbf{v}}$ to be a strict local maximizer of $(StQP)$. In this section we shall consider the stability properties of the maximizer. More precisely we will study the effect of small perturbations in the matrix involved in $(StQP)$ on local maximizers. For this purpose we will consider the following parametric optimization problem,

$$(StQP_Q) \qquad \max \quad q_Q(\mathbf{v}) := \frac{1}{2}\mathbf{v}^T Q\mathbf{v} \quad s.t. \quad \mathbf{v} \in \Delta_m$$

where $Q \in \mathcal{S}_m$ is seen as a parameter. First we define for the candidate maximizer $\overline{\mathbf{v}}$ the matrix $I_{\overline{\mathbf{v}}}$ such that $I_{\overline{\mathbf{v}}} = [e_i : i \notin R(\overline{\mathbf{v}})]$. Then the KKT conditions given in (3.1) can be written in the matrix form as follows,

$$\begin{pmatrix} Q & -\mathbf{e} & I_{\overline{\mathbf{v}}} \\ -\mathbf{e}^T & 0 & \mathbf{o}^T \\ I_{\overline{\mathbf{v}}}^T & \mathbf{o} & O \end{pmatrix} \begin{pmatrix} \overline{\mathbf{v}} \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{o} \\ -1 \\ \mathbf{o} \end{pmatrix} \tag{3.11}$$

where $\mu \in \mathbb{R}^{|\mathcal{U}\backslash R(\overline{\mathbf{v}})|}$, $\mathbf{o}$ is a zero vector of appropriate dimension while $O$ is the $|\mathcal{U}\backslash R(\overline{\mathbf{v}})| \times |\mathcal{U}\backslash R(\overline{\mathbf{v}})|$ matrix of zeros. First we will give an auxiliary result.

**Lemma 3.15.** *Let $A \in \mathcal{S}_m$ and $B \in \mathbb{R}^{m\times n}$ be such that $\mathbf{d}^T A\mathbf{d} < 0$ for all $\mathbf{d} \in \ker\{B^T\}\backslash\{\mathbf{o}\}$. Then the matrix $Q := \left(\begin{smallmatrix} A & B \\ B^T & O \end{smallmatrix}\right)$ is nonsingular if and only if the columns of $B$ are linearly independent.*

*Proof.* $\Rightarrow$ Let $Q$ be nonsingular but let the columns of the matrix $B$ be linearly dependent, that is there exists a $\mathbf{w} \neq \mathbf{o}$ such that $B\mathbf{w} = \mathbf{o}$. Then for $\mathbf{u} = \mathbf{o}$ and $\mathbf{v} := \left(\begin{smallmatrix} \mathbf{u} \\ \mathbf{w} \end{smallmatrix}\right)$, we write,

$$Q\mathbf{v} = \begin{pmatrix} A\mathbf{u} + B\mathbf{w} \\ B^T\mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{o} \\ \mathbf{o} \end{pmatrix}$$

which leads to a contradiction to the assumption that $Q$ is nonsingular.
$\Leftarrow$ Let the columns of the matrix $B$ be linearly independent and let the matrix $Q$ be singular,i.e., there exists an $\mathbf{o} \neq \mathbf{v} := \left(\begin{smallmatrix} \mathbf{u} \\ \mathbf{w} \end{smallmatrix}\right) \in \mathbb{R}^{m+n}$ such that $Q\mathbf{v} = \mathbf{o}$. So we have,

$$A\mathbf{u} + B\mathbf{w} = \mathbf{o} \tag{3.12a}$$

$$B^T\mathbf{u} = \mathbf{o} \tag{3.12b}$$

From (3.12b) we can conclude that $\mathbf{u} \in \ker\{B^T\}$. If $\mathbf{u} = \mathbf{o}$ then $\mathbf{w} \neq \mathbf{o}$ (otherwise $\mathbf{v} = \mathbf{o}$). So from (3.12a) we get $B\mathbf{w} = \mathbf{o}$ which is a contradiction to the basic hypothesis that the columns of $B$ are linearly independent. Now if $\mathbf{o} \neq \mathbf{u} \in \ker\{B^T\}$ then from (3.12a) we get $\mathbf{u}^T A\mathbf{u} = -\mathbf{u}^T B\mathbf{w} = -(B^T\mathbf{u})\mathbf{w}$ which using (3.12b) reduces to $\mathbf{u}^T A\mathbf{u} = 0$ hence we arrive at a contradiction to the hypothesis that $\mathbf{d}^T A\mathbf{d} < 0$ for all $\mathbf{o} \neq \mathbf{d} \in \ker\{B^T\}$. □

In the following theorem we will show that for a strict local maximizer satisfying the KKT condition with strict complementarity, locally, the maximizer changes smoothly with the matrix $Q$. The following theorem is a special case of the result given in [64, Theorem 6].

**Theorem 3.16.** *Let $\overline{\mathbf{v}} \in \Delta_m$ satisfies the KKT conditions* (3.1) *with respect to the matrix $\overline{Q}$ with Lagrange multipliers $\overline{\lambda}$ and $\overline{\mu}$ and let $\overline{\mathbf{v}}, \overline{\mu}$ satisfy the strict complementarity conditions, i.e., $\widetilde{S}(\overline{\mathbf{v}}) = R(\overline{\mathbf{v}})$. In addition assume that $\mathbf{d}^T\overline{Q}\mathbf{d} < 0$ holds for all $\mathbf{o} \neq \mathbf{d} \in C(\overline{\mathbf{v}})$ (i.e., $\overline{\mathbf{v}}$ is a strict local maximizer). Then there exits a $C^\infty$ function $f : N_\varepsilon(\overline{Q}) \to N_\delta(\overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu})$, $f(Q) = (\mathbf{v}(Q), \lambda(Q), \mu(Q))$ such that $\mathbf{v}(Q)$ is a strict local maximizer of $(StQP_Q)$ and $f(Q), \mathbf{v}(Q), \mu(Q)$ satisfies strict complementarity.*

*Proof.* Define the system of equations,

$$F(Q, \mathbf{v}, \lambda, \mu) := \begin{pmatrix} Q & -\mathbf{e} & I_{\overline{\mathbf{v}}} \\ -\mathbf{e}^T & 0 & \mathbf{o}^T \\ I_{\overline{\mathbf{v}}}^T & \mathbf{o} & O \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{o} \\ -1 \\ \mathbf{o} \end{pmatrix} \tag{3.13}$$

where $\mu \in \mathbb{R}_+^{|\mathcal{U} \setminus R(\overline{\mathbf{v}})|}$, then the Jacobian of $F$ with respect to $\mathbf{v}, \lambda, \mu$ reads,

$$\nabla_{\mathbf{v},\lambda,\mu} F(\overline{Q}, \overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu}) := \begin{pmatrix} \overline{Q} & -\mathbf{e} & I_{\overline{\mathbf{v}}} \\ -\mathbf{e}^T & 0 & \mathbf{o}^T \\ I_{\overline{\mathbf{v}}}^T & \mathbf{o} & O \end{pmatrix}$$

First we will show that the matrix $\nabla_{\mathbf{v},\lambda,\mu} F(\overline{Q}, \overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu})$ is nonsingular. We take $B := [-\mathbf{e} \;\; I_{\overline{\mathbf{v}}}]$ and note that ker$\{B^T\} = T(\overline{\mathbf{v}})$ (see (3.8)). Recall that from strict complementarity we have $C(\overline{\mathbf{v}}) = T(\overline{\mathbf{v}})$. Moreover the columns of the matrix $B$ are linearly independent, so from the conditions $\mathbf{d}^T Q \mathbf{d} < 0$ for all $\mathbf{o} \neq \mathbf{d} \in C(\overline{\mathbf{v}}) = T(\overline{\mathbf{v}})$ (since $R(\overline{\mathbf{v}}) = \widetilde{S}(\overline{\mathbf{v}})$) we can conclude that the matrix $\nabla_{\mathbf{v},\lambda,\mu} F(\overline{Q}, \overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu})$ is nonsingular. Hence by the inverse function theorem (see e.g. [133, Theorem 9.24]) there exists $\varepsilon > 0$ and $\delta > 0$ and a $C^\infty$ function $f$,

$$f : N_\varepsilon(\overline{Q}) \to N_\delta(\overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu})$$

such that $f(\overline{Q}) = (\overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu})$ and $(\mathbf{v}(Q), \lambda(Q), \mu(Q)) \in N_\delta(\overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu})$ is the unique solution of (3.13).

In order to show that $\mathbf{v}(Q)$ is a strict local maximizer, we will show that $\mathbf{v}(Q)$ satisfies the second order sufficient condition. For the sake of contradiction we assume that there exists an infinite sequence $Q_k \to \overline{Q}$ and critical vectors $\mathbf{d}_k \in C(\mathbf{v}(Q_k))$, $\|\mathbf{d}_k\| = 1$, such that

$$\mathbf{d}_k^T Q_k \mathbf{d}_k \geq 0 \tag{3.14}$$

We can assume that $\mathbf{d}_k \to \mathbf{d}$ (as we did in Theorem 3.4) with $\|\mathbf{d}\| = 1$. First we will show that $\mathbf{d} \in C(\overline{\mathbf{v}})$. For simplicity of notation take $\mathbf{v}_k := \mathbf{v}(Q_k)$ and note that $\mathbf{d}_k \in C(\mathbf{v}_k)$ will imply that $\mathbf{e}^T \mathbf{d}_k = 0$, $(Q_k \mathbf{v}_k)^T \mathbf{d}_k \geq 0$ and $(\mathbf{d}_k)_i \geq 0$ for all $i \notin R(\mathbf{v}_k)$. Taking $k \to \infty$ and noting that from (3.13) we get, $R(\mathbf{v}_k) = R(\overline{\mathbf{v}})$ we arrive at $\mathbf{d} \in C(\overline{\mathbf{v}})$. Now taking $k \to \infty$ in (3.14) gives $\mathbf{d}^T Q \mathbf{d} \geq 0$ which is a clear contradiction to the assumption $\mathbf{d}^T Q \mathbf{d} < 0$ for all $\mathbf{o} \neq \mathbf{d} \in C(\overline{\mathbf{v}})$.

We choose $\varepsilon > 0$, such that $\mu(Q) > 0$ for all $Q \in N_\varepsilon(\overline{Q})$ in order to preserve strict complementarity. □

In the following example we will show that the strict complementarity condition is indeed essential to assure the stability of a strict local maximizer in $(StQP_Q)$.

**Example 3.17** ([17]). *Consider the matrix,*

$$\overline{Q} := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

*It is not difficult to verify that* $\overline{\mathbf{v}} := e_3$ *is a strict local maximizer*[1] *and* $R(e_3) := \{3\}$, $\widetilde{S}(e_3) := \{1, 2, 3\}$. *Now consider the perturbed matrix,*

$$Q_\varepsilon := \begin{pmatrix} 0 & -\varepsilon & 1-\varepsilon \\ -\varepsilon & 0 & 1-\varepsilon \\ 1-\varepsilon & 1-\varepsilon & 1-2\varepsilon \end{pmatrix}$$

*In this case for every* $\mathbf{v} := (v_1, v_2, v_3)^T \in \Delta_3$ *we have,*

$$\begin{aligned} \mathbf{v}^T Q_\varepsilon \mathbf{v} &= (1-\varepsilon)v_1v_3 + (1-\varepsilon)v_2v_3 + (1-\varepsilon)v_3 - \varepsilon(v_3^2 + 2v_1v_2) \\ &= 1 - (v_1+v_2)^2 - 2\varepsilon(1-v_1)(1-v_2) \end{aligned}$$

*It is not difficult to verify that* $v_\alpha := (1-\alpha)(\varepsilon, 0, 1-\varepsilon)^T + \alpha(0, \varepsilon, 1-\varepsilon)^T$ *for* $0 \le \alpha \le 1$, *is a local maximizer.*

Our next result will establish the Lipschitz stability. First we define the Lipschitz continuity,

**Definition 3.18** (Lipschitz Continuity). A function $f : \mathbb{R}^n \to \mathbb{R}^l$ is said to be Lipschitz continuous at a point $\mathbf{v}_0 \in \mathbb{R}^n$ if there exists a constant $L, \varepsilon > 0$ such that

$$\|f(\mathbf{v}) - f(\mathbf{v}_0)\| \le L\|\mathbf{v} - \mathbf{v}_0\| \quad \forall\, \mathbf{v} \in N_\varepsilon(\mathbf{v}_0).$$

The following Lemma is useful for the proof of the next two theorems,

**Lemma 3.19.** *Let* $\overline{Q} \in \mathcal{S}_m$ *and let* $\overline{\mathbf{v}} \in \Delta_m$ *be a strict local maximizer with respect to* $\overline{Q}$ *then there exist* $\varepsilon > 0$, $\delta > 0$ *and* $\mathbf{v}(Q) \in N_\delta(\overline{\mathbf{v}}) \cap \Delta_m$ *such that for all* $Q \in N_\varepsilon\left(\overline{Q}\right)$, *the point* $\mathbf{v}(Q)$ *is a local maximizer with respect to* $Q$.

*Proof.* Since $\overline{\mathbf{v}}$ is a strict local maximizer and by continuity there exists $\varepsilon > 0$, $\alpha > 0$, $\delta > 0$ such that

$$q_Q(\overline{\mathbf{v}}) \ge q_{\overline{Q}}(\overline{\mathbf{v}}) - \frac{\alpha}{2} \quad \forall\, Q \in N_\varepsilon(\overline{Q}) \tag{3.15}$$

$$q_{\overline{Q}}(\mathbf{v}) \le q_{\overline{Q}}(\overline{\mathbf{v}}) - 2\alpha \quad \forall\, \mathbf{v} \in \Delta_m \text{ such that } ||\mathbf{v} - \overline{\mathbf{v}}|| = \delta \tag{3.16}$$

[1] since $e_3^T \overline{Q} e_3 = 1 > \mathbf{v}^T Q \mathbf{v} = 1 - (v_1+v_2)^2$ for all $e_3 \neq \mathbf{v} := (v_1, v_2, 1 - v_1 - v_2) \in \Delta_3$

$$q_Q(\mathbf{v}) \leq q_{\overline{Q}}(\overline{\mathbf{v}}) - \alpha \quad \forall\, \mathbf{v} \in \Delta_m, \text{ such that } ||\mathbf{v} - \overline{\mathbf{v}}|| = \delta, \text{ and } \; Q \in N_\varepsilon(\overline{Q}). \tag{3.17}$$

Combining (3.15) and (3.17) we have,

$$q_Q(\overline{\mathbf{v}}) \geq q_{\overline{Q}}(\overline{\mathbf{v}}) - \frac{\alpha}{2} \geq q_Q(\mathbf{v}) + \frac{\alpha}{2} \qquad \begin{array}{l}\forall\, \mathbf{v} \in \Delta_m, \text{ such that } ||\mathbf{v} - \overline{\mathbf{v}}|| = \delta, \\ \text{and } \; Q \in N_\varepsilon(\overline{Q})\end{array}$$

So, there exists a global maximizer $\mathbf{v} = \mathbf{v}(Q)$ on $\Delta_m \cap N_\delta(\overline{\mathbf{v}})$ with $\|\mathbf{v} - \overline{\mathbf{v}}\| < \delta$ which is a local maximizer. □

In the following theorem and throughout the thesis for a matrix $Q \in \mathbb{R}^{m \times m}$, $\|Q\|$ denotes the Frobenius norm of the matrix $Q$, i.e., $\|Q\| := \sqrt{\text{tr}(QQ^T)}$.

**Theorem 3.20.** *Let $\overline{Q} \in \mathcal{S}_m$ and let $\overline{\mathbf{v}} \in \Delta_m$ be a strict local maximizer with respect to $\overline{Q}$. Then there exist $\varepsilon > 0$, $\delta > 0$ and $L > 0$ such that for all $Q \in N_\varepsilon\left(\overline{Q}\right)$ and for all local maximizers $\mathbf{v}\left(Q\right)$ of $(StQP_Q)$ with $\mathbf{v}\left(Q\right) \in N_\delta\left(\overline{\mathbf{v}}\right)$ the following holds,*

$$\|\mathbf{v}\left(Q\right) - \overline{\mathbf{v}}\| \leq L \left\|Q - \overline{Q}\right\|.$$

*Proof.* Since $\overline{\mathbf{v}}$ is a strict local maximizer by Theorem 3.4 it must satisfy the second order condition. So there exist $\gamma, \delta > 0$ such that,

$$\gamma \left\|\mathbf{v} - \overline{\mathbf{v}}\right\|^2 \leq q_{\overline{Q}}\left(\overline{\mathbf{v}}\right) - q_{\overline{Q}}\left(\mathbf{v}\right) \quad \forall\, \mathbf{v} \in N_\delta\left(\overline{\mathbf{v}}\right) \cap \Delta_m \tag{3.18}$$

From Lemma 3.19 it is clear that there exist a local maximizer $\mathbf{v}\left(Q\right)$ with respect to the matrix $Q \in N_\varepsilon(\overline{Q})$. For the ease of notation we take $\mathbf{v} := \mathbf{v}\left(Q\right)$. Now since $\mathbf{v}$ is a local maximizer with respect to $Q \in N_\varepsilon(\overline{Q})$, we have,

$$q_Q\left(\overline{\mathbf{v}}\right) - q_Q\left(\mathbf{v}\right) \leq 0$$

Define $h\left(\mathbf{v}\right) := q_Q\left(\mathbf{v}\right) - q_{\overline{Q}}\left(\mathbf{v}\right)$. Using the mean value theorem with respect to $\mathbf{v}$, with $0 < \tau < 1$, we find

$$\begin{aligned} h\left(\mathbf{v}\right) - h\left(\overline{\mathbf{v}}\right) &= \nabla_{\mathbf{v}} h\left(\overline{\mathbf{v}} + \tau\left(\mathbf{v} - \overline{\mathbf{v}}\right)\right)\left(\mathbf{v} - \overline{\mathbf{v}}\right) \\ &\leq \left\|\nabla_{\mathbf{v}} h\left(\overline{\mathbf{v}} + \tau\left(\mathbf{v} - \overline{\mathbf{v}}\right)\right)\right\| \left\|\left(\mathbf{v} - \overline{\mathbf{v}}\right)\right\| \\ &= \|(\overline{\mathbf{v}} + \tau\left(\mathbf{v} - \overline{\mathbf{v}}\right))^T (Q - \overline{Q})\| \|\mathbf{v} - \overline{\mathbf{v}}\| \\ &\leq \|\overline{\mathbf{v}} + \tau\left(\mathbf{v} - \overline{\mathbf{v}}\right)\| \|Q - \overline{Q}\| \|\mathbf{v} - \overline{\mathbf{v}}\| \\ &\leq \max_{\mathbf{w} \in N_\delta(\overline{\mathbf{v}})} \|\mathbf{w}\| \left\|Q - \overline{Q}\right\| \|\mathbf{v} - \overline{\mathbf{v}}\| \end{aligned}$$

Finally consider,

$$\begin{aligned}
q_{\overline{Q}}(\overline{\mathbf{v}}) - q_{\overline{Q}}(\mathbf{v}) &= \left[q_Q(\mathbf{v}) - q_{\overline{Q}}(\mathbf{v})\right] - \left[q_Q(\overline{\mathbf{v}}) - q_{\overline{Q}}(\overline{\mathbf{v}})\right] + \left[q_Q(\overline{\mathbf{v}}) - q_Q(\mathbf{v})\right] \\
&= h(\mathbf{v}) - h(\overline{\mathbf{v}}) + \underbrace{\left[q_Q(\overline{\mathbf{v}}) - q_Q(\mathbf{v})\right]}_{\leq 0} \\
&\leq h(\mathbf{v}) - h(\overline{\mathbf{v}}) \\
&\leq \max_{\mathbf{w}\in N_\delta(\overline{\mathbf{v}})} \|\mathbf{w}\| \left\|Q - \overline{Q}\right\| \|\mathbf{v} - \overline{\mathbf{v}}\|
\end{aligned}$$

Take $c := \max_{\mathbf{w}\in N_\delta(\overline{\mathbf{v}})} \|\mathbf{w}\|$ then the above relation together with (3.18) implies,

$$\gamma \|\mathbf{v} - \overline{\mathbf{v}}\|^2 \leq c \left\|Q - \overline{Q}\right\| \|\mathbf{v} - \overline{\mathbf{v}}\|$$

Hence the result follows with $L := \frac{c}{\gamma}$. □

In Example 3.17 the maximizer does not satisfy strict complementarity and we have shown that there exists a matrix in the neighbourhood which does not have a strict local maximizer. But there exists situations where a strict local maximizer behaves stable even if the strict complementarity does not hold. Consider the following example,

**Example 3.21.** *Consider the matrix,*

$$Q := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

*It is not difficult to verify that $e_3$ is a strict local maximizer*[2] *and $R(e_3) := \{3\}$, $\widetilde{S}(e_3) := \{1,2,3\}$. In order to show that each matrix in the neighbourhood has a unique strict local maximizer near $e_3$ it is sufficient (as we will see in Theorem 3.22) to show that $\mathbf{d}^T Q\mathbf{d} < 0$ for all $\mathbf{o} \neq \mathbf{d} \in T^+(e_3)$ (see (3.9)). First we calculate $T^+(e_3)$,*

$$\begin{aligned}
T^+(e_3) &:= \{\mathbf{d} \in \mathbb{R}^3 : \mathbf{e}^T\mathbf{d} = 0, d_i = 0\ \forall i \notin \widetilde{S}(e_3)\} \\
&:= \{\mathbf{d} \in \mathbb{R}^3 : d_1 + d_2 + d_3 = 0\}
\end{aligned}$$

*Now for $\mathbf{o} \neq \mathbf{d} \in T^+(e_3)$ we get,*

$$\mathbf{d}^T Q\mathbf{d} = d_1(d_2 + d_3) + d_2(d_1 + d_3) = -((d_2 + d_3)^2 + d_2^2) = -(d_1^2 + d_2^2) < 0.$$

[2] since $e_3^T Q e_3 = 1 > 1 - (v_1^2 + v_2^2) = \mathbf{v}^T Q\mathbf{v}$ for all $e_3 \neq \mathbf{v} := (v_1, v_2, 1 - v_1 - v_2) \in \Delta_3$

A KKT point $\overline{\mathbf{v}} \in \Delta_m$ with respect to $(StQP_{\overline{Q}})$ is said to satisfy the *strong second order condition* if $\mathbf{d}^T\overline{Q}\mathbf{d} < 0$ for all $\mathbf{d} \in T^+(\overline{\mathbf{v}})$ (see (3.9)). Theorem 3.22 states that if the strict local maximizer $\overline{\mathbf{v}}$ satisfies the strong second order sufficient condition then locally the strict local maximizers, with respect to the matrix $Q$, behave Lipschitz continuous. Note that the result presented in Theorem 3.22 is a special case of the result in [98, Theorem 2]. For the sake of completeness we also provide a proof.

**Theorem 3.22.** *Let $\overline{Q} \in \mathcal{S}_m$ and let $\overline{\mathbf{v}} \in \Delta_m$ satisfy the KKT conditions* (3.1) *with respect to $\mathbf{o} \neq \overline{Q}$, with Lagrange multipliers $\overline{\lambda}$ and $\overline{\mu}$. In addition let $\mathbf{d}^T\overline{Q}\mathbf{d} < 0$ hold for all $\mathbf{d} \in T^+(\overline{\mathbf{v}})$. Then there exist $\varepsilon, \delta > 0$ and a Lipschitz continuous function $f : N_\varepsilon(\overline{Q}) \to N_\delta(\overline{\mathbf{v}}, \overline{\lambda}, \overline{\mu})$, $f(Q) = (\mathbf{v}(Q), \mu(Q), \lambda(Q))$ such that $\mathbf{v}(Q)$ is a strict local maximizer with respect to $Q$. Moreover $\mathbf{v}(Q)$ is the unique local maximizer with respect to $Q$ in a neighbourhood of $\overline{\mathbf{v}}$.*

*Proof.* From Lemma 3.19 it is clear that there exist a local maximizer $\mathbf{v}\,(Q)$ with respect to the matrix $Q \in N_\varepsilon(\overline{Q})$. The maximizer $\mathbf{v}(Q)$ must lie on one of the faces ( say $\mathrm{fc}_R$) of $\Delta_m$ with $R(\overline{\mathbf{v}}) \subseteq R \subseteq \widetilde{S}(\overline{\mathbf{x}})$. So with $R^c := \mathcal{U}\backslash R$, $S^c(\overline{\mathbf{v}}) := \mathcal{U}\backslash\widetilde{S}(\overline{\mathbf{v}}), R^c(\overline{\mathbf{v}}) := \mathcal{U}\backslash R(\overline{\mathbf{v}})$, it is clear that

$$S^c(\overline{\mathbf{v}}) \subseteq R^c \subseteq R^c(\overline{\mathbf{v}}) \tag{3.19}$$

The maximizer property of $\mathbf{v}(Q)$ and continuity will imply that the corresponding multipliers $\lambda(Q) \in \mathbb{R}$, $\mu(Q) \in \mathbb{R}_+^{|R^c|}$ must be the solution of one of the (finitely many) systems of KKT equations,

$$\begin{pmatrix} Q & -\mathbf{e} & I_R \\ -\mathbf{e}^T & 0 & 0 \\ I_R & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{o} \\ -1 \\ \mathbf{o} \end{pmatrix} \tag{3.20}$$

where $I_R := [e_i : i \notin R]$. By the strong second order condition, i.e., $\mathbf{d}^TQ\mathbf{d} < 0$ for all $\mathbf{o} \neq \mathbf{d} \in T^+(\overline{\mathbf{v}})$, for any $R$ in (3.19) the system matrices of (3.20) are nonsingular in $N_\varepsilon(\overline{Q})$ for $\varepsilon > 0$ small enough. So the KKT points $\mathbf{v}(Q), \lambda(Q), \mu(Q)$ must coincide with values of one of the (finitely many) rational $C^\infty$ functions,

$$\begin{pmatrix} \mathbf{v}(Q) \\ \lambda(Q) \\ \mu(Q) \end{pmatrix} = \begin{pmatrix} Q & -\mathbf{e} & I_R \\ -\mathbf{e}^T & 0 & 0 \\ I_R & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{o} \\ -1 \\ \mathbf{o} \end{pmatrix} \tag{3.21}$$

We now show that for $\varepsilon > 0$ small enough any $Q \in N_\varepsilon(\overline{Q})$ can have at most one local maximizer in $N_{\delta_1}(\overline{\mathbf{v}})$, for some $\delta_1 > 0$. Assume that there exists a sequence $Q_\nu \to Q, \nu \to \infty$ and two local maximizer $\mathbf{v}_\nu^1 \neq \mathbf{v}_\nu^2$ of $\frac{1}{2}\mathbf{v}Q_\nu\mathbf{v}$ over $\Delta_m$ such that

$\mathbf{v}_\nu^1, \mathbf{v}_\nu^2 \in N_{\delta_1}(\overline{\mathbf{v}})$. Each sequence $\mathbf{v}_\nu^\rho$, $\rho = 1, 2$ must have a limit point $\mathbf{v}^\rho \in N_\delta(\overline{\mathbf{v}})$. By a continuity argument $q_{\overline{Q}}(\mathbf{v}^\rho) = q_{\overline{Q}}(\overline{\mathbf{v}})$ and thus $\mathbf{v}^\rho = \overline{\mathbf{v}}$, $\rho = 1, 2$. Without loss of generality consider a sequence of solutions such that $\mathbf{v}_\nu^\rho \to \overline{\mathbf{v}}$ and that for all $\rho$, $\mathbf{v}_\nu^\rho$ are solution of (3.20) for (the same) index set $R_1 \neq R_2$ satisfying $\widetilde{S}^c(\overline{\mathbf{v}}) \subseteq R_\rho^c \subseteq R^c(\overline{\mathbf{v}})$ for $\rho = 1, 2$:

$$Q_\nu \mathbf{v}_\nu^\rho - \lambda_\nu^\rho e + I_{R_\rho}\mu_\nu^\rho = 0, \mathbf{e}^T\mathbf{v}_\nu^\rho = 1, [\mathbf{v}_\nu^\rho]_j = 0 \quad j \in R_\rho,\ \rho = 1, 2. \tag{3.22}$$

Since either $q_{Q_\nu}(\mathbf{v}_\nu^1) \leq q_{Q_\nu}(\mathbf{v}_\nu^2)$ or $q_{Q_\nu}(\mathbf{v}_\nu^1) \geq q_{Q_\nu}(\mathbf{v}_\nu^2)$ holds again by selecting a subsequence, without loss of generality, we assume,

$$0 \leq q_{Q_\nu}(\mathbf{v}_\nu^2) - q_{Q_\nu}(\mathbf{v}_\nu^1) \quad \forall \quad \nu \in \mathbb{N} \tag{3.23}$$

By putting $\mathbf{d}_\nu := \frac{\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1}{||\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1||}$ we can assume $\mathbf{d}_\nu \to \mathbf{d}$, $||\mathbf{d}|| = 1$. Observe that $\mathbf{e}^T\mathbf{d}_\nu = 0$ holds and since $(\mathbf{v}_\nu^1)_j = 0$ for all $j \in R_1$ (see (3.22)), then

$$(\mathbf{v}_\nu^2)_j - (\mathbf{v}_\nu^1))_j \geq 0 \quad \forall j \in R_1 \tag{3.24}$$

and thus $(\mathbf{d}_\nu)_j \geq 0$ for all $j \in R_1$. By taking the limit $\nu \to \infty$ we find $\mathbf{e}^T\mathbf{d} = 0$ with $d_j = 0$ for all $j \notin \widetilde{S}(\overline{\mathbf{v}})$ and $d_j \geq 0$ for all $j \in R_1$. In particular $\mathbf{d} \in T^+(\overline{\mathbf{v}})$. In view of (3.23) and using the KKT conditions for $\mathbf{v}_\nu^1$ we obtain,

$$\begin{aligned}
0 &\leq q_{Q_\nu}(\mathbf{v}_\nu^2) - q_{Q_\nu}(\mathbf{v}_\nu^1) \\
&= (\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1)^T Q_\nu \mathbf{v}_\nu^1 + \frac{1}{2}(\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1)^T Q_\nu (\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1) \\
&= -(\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1)^T I_{R_1}\mu_\nu^1 + \frac{1}{2}(\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1)^T Q_\nu (\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1) \\
&\leq \frac{1}{2}(\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1)^T Q_\nu (\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1)
\end{aligned}$$

Since $(\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1)^T I_{R_1}\mu_\nu^1 \geq 0$ by (3.24). By dividing these relations by $||\mathbf{v}_\nu^2 - \mathbf{v}_\nu^1||$ and letting $\nu \to \infty$ it follows $\mathbf{d}^T Q\mathbf{d} \geq 0$ with $\mathbf{o} \neq \mathbf{d} \in T^+(\overline{\mathbf{v}})$ which is a contradiction to the strong second order condition. □

Note that for the matrix $Q$ in Example 3.17 the strong second order sufficient condition does not hold.

## 3.4 Evolutionarily Stable Strategy

The concept of an evolutionarily stable strategy (ESS) was defined by Maynard-Smith and Price [142]. The concept was introduced as the application of a game theoretic model to the conflict among animals. In order to give a mathematical formulation, the conflict among animals is described by an $m \times m$ matrix $Q = (q_{ij})$ where $q_{ij}$ is the expected gain a user of pure strategy $i$ gets whose opponent uses pure strategy $j$. If a user plays each pure strategy with a certain probability then its strategy is called mixed strategy. So the set of all available strategies can be denoted by the unit simplex $\Delta_m$. The mean payoff to a user of strategy $\mathbf{v} \in \Delta_m$ whose opponent plays the strategy $\mathbf{u} \in \Delta_m$ is then $\mathbf{v}^T Q\mathbf{u}$. Now consider (see [91]) an infinite monomorphic population which has achieved a stable state and assume that some new population of size $\varepsilon$ invade the current population. These mutant/migrant are users of a mixed strategy $\mathbf{u}$. Then the ESS conditions for $\mathbf{v}$ say that the average gain of the user of strategy $\mathbf{v}$ is strictly greater than the average gain of the mutant/migrant which are using strategy $\mathbf{u}$ i.e.,

$$(1-\varepsilon)\mathbf{v}^T Q\mathbf{v} + \varepsilon \mathbf{v}^T Q\mathbf{u} > (1-\varepsilon)\mathbf{u}^T Q\mathbf{v} + \varepsilon \mathbf{u}^T Q\mathbf{u} \quad \forall\, \varepsilon > 0 \text{ small} \tag{3.25}$$

Now in the limiting case $\varepsilon \to 0$ we will have for all $\mathbf{u} \neq \mathbf{v}$, $\mathbf{v}^T Q\mathbf{v} \geq \mathbf{u}^T Q\mathbf{v}$ and in the case of equality we will have, $\mathbf{v}^T Q\mathbf{u} > \mathbf{u}^T Q\mathbf{u}$ [91]. Hence an ESS can be defined in the following way,

**Definition 3.23** (Evolutionarily Stable Strategy (ESS))**.** Let $Q \in \mathbb{R}^{m\times m}$, then $\mathbf{v} \in \Delta_m$ is an ESS with respect to the matrix $Q$ if following two conditions hold,

i. $\mathbf{v}^T Q\mathbf{v} \geq \mathbf{u}^T Q\mathbf{v} \quad \forall \quad \mathbf{u} \in \Delta_m$

ii. if $\mathbf{u} \in \Delta_m, \mathbf{v} \neq \mathbf{u}, \mathbf{v}^T Q\mathbf{v} = \mathbf{u}^T Q\mathbf{v}$ then $\mathbf{v}^T Q\mathbf{u} > \mathbf{u}^T Q\mathbf{u}$

For a review of the ESS theory the interested reader is referred to [91]. In [112, 141] the mathematical foundation and its relation to the theory of evolution is discussed. The computational complexity of ESS is discussed in [62, 118]. Some exact and approximate algorithms for finding ESS is the topic of [21].

In the case of symmetric matrices the concept of ESS is directly related to the maximization problem $(StQP)$.

**Proposition 3.24.** *Let $\bar{\mathbf{v}} \in \Delta_m$ and $Q \in \mathcal{S}_m$, then $\bar{\mathbf{v}}$ is a strict local maximizer of $(StQP)$ if and only if $\bar{\mathbf{v}}$ is an ESS with respect to $Q$.*

*Proof.* $\Rightarrow$ Let $\bar{\mathbf{v}}$ be a strict local maximizer. We will show that $\bar{\mathbf{v}}$ is an ESS. In order to prove this first note that $\Delta_m$ is convex and for any $\mathbf{u} \in \Delta_m$, $\mathbf{u} \neq \bar{\mathbf{v}}$ and $\lambda > 0$ small the following holds for $\mathbf{w} = \bar{\mathbf{v}} + \lambda(\mathbf{u} - \bar{\mathbf{v}})$,

$$0 < \mathbf{v}^T Q\mathbf{v} - \mathbf{w}^T Q\mathbf{w} = 2\lambda(\bar{\mathbf{v}}^T Q\bar{\mathbf{v}} - \bar{\mathbf{v}}^T Q\mathbf{u}) - \lambda^2(\mathbf{u} - \bar{\mathbf{v}})^T Q(\mathbf{u} - \bar{\mathbf{v}})$$

$$= \lambda(2-\lambda)(\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} - \overline{\mathbf{v}}^T Q\mathbf{u}) + \lambda^2(\mathbf{u}^T Q\overline{\mathbf{v}} - \mathbf{u}^T Q\mathbf{u}) \quad (3.26)$$

Then by dividing this expression by $\lambda > 0$ and taking $\lambda \to 0$ we get

$$\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} - \overline{\mathbf{v}}^T Q\mathbf{u} \geq 0.$$

If $\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} - \overline{\mathbf{v}}^T Q\mathbf{u} = 0$ then from (3.26) we get $\mathbf{u}^T Q\mathbf{u} - \mathbf{u}^T Q\overline{\mathbf{v}} < 0$. Hence $\overline{\mathbf{v}}$ is an ESS.
$\Leftarrow$ Let $\overline{\mathbf{v}}$ be an ESS we will prove that $\overline{\mathbf{v}}$ is a local maximizer. In order to show this take any $\mathbf{w} \neq \overline{\mathbf{v}}$ near $\overline{\mathbf{v}}$, $\mathbf{w} \in \Delta_m$. We can find a $\mathbf{u} \in \Delta_m$, $\mathbf{u} \neq \overline{\mathbf{v}}$ and (small) $\lambda > 0$ such that $\mathbf{w} := \overline{\mathbf{v}} + \lambda(\mathbf{u} - \overline{\mathbf{v}}) \in \Delta_m$. Now consider,

$$\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} - \mathbf{w}^T Q\mathbf{w} = \lambda(2-\lambda)\underbrace{(\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} - \overline{\mathbf{v}}^T Q\mathbf{u})}_{\geq 0} + \lambda^2 \underbrace{(\mathbf{u}^T Q\overline{\mathbf{v}} - \mathbf{u}^T Q\mathbf{u})}_{>0 \text{ if } \overline{\mathbf{v}}^T Q\overline{\mathbf{v}} - \overline{\mathbf{v}}^T Q\mathbf{u} = 0} > 0$$

The above inequality holds due to the conditions of ESS. □

In the proposition above we have seen that an ESS for symmetric matrices is equivalent to a strict local maximizer of $(StQP)$. In the case of nonsymmetric matrices neither direction holds meaning that if $\overline{\mathbf{v}}$ is an ESS, we cannot say if it is a strict local maximizer and also if $\overline{\mathbf{v}}$ is a strict local maximizer we cannot say that it is an ESS. Consider the following example,

**Example 3.25.** *Let*

$$Q := \begin{pmatrix} 1 & 1 \\ 3 & \frac{1}{2} \end{pmatrix}$$

*First we claim that* $\overline{\mathbf{v}} := \left(\frac{1}{5}, \frac{4}{5}\right)^T \in \Delta_2$ *is an ESS. In order to see this take* $\mathbf{u} := (u_1, 1-u_1)^T \in \Delta_2$. *Then it is not difficult to verify the following,*

$$\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} = \mathbf{u}^T Q\overline{\mathbf{v}} = 1, \;\; \overline{\mathbf{v}}^T Q\mathbf{u} = \frac{3}{5} + 2u_1, \;\; \mathbf{u}^T Q\mathbf{u} = 3u_1 + \frac{1}{2} - \frac{5}{2}u_1^2 \;.$$

*From this we get,* $\overline{\mathbf{v}}^T Q\mathbf{u} - \mathbf{u}^T Q\mathbf{u} = \frac{5}{2}\left(u_1 - \frac{1}{5}\right)^2 > 0$ *for all* $u_1 \neq v_1$. *Hence* $\overline{\mathbf{v}}$ *is an ESS. But* $\overline{\mathbf{v}}$ *is not a strict local maximizer since for* $\varepsilon > 0$ *very small take* $\mathbf{w} := \overline{\mathbf{v}} + \varepsilon\left(e_1 - \overline{\mathbf{v}}\right) = \frac{1}{5}(1+4\varepsilon, 4-4\varepsilon)^T$ *then we have*

$$\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} - \mathbf{w}^T Q\mathbf{w} = 1 - \left(1 + \frac{8}{5}\varepsilon\left(1-\varepsilon\right)\right) = -\frac{8}{5}\varepsilon\left(1-\varepsilon\right) < 0 \;.$$

*Now consider* $\widetilde{\mathbf{v}} := \left(\frac{3}{5}, \frac{2}{5}\right)^T$. *Then it can be readily verified that* $\widetilde{\mathbf{v}}$ *is not an ESS. We* *will show that $\widetilde{\mathbf{v}}$ is a strict local maximizer. For $0 < \lambda < 1$ and $\mathbf{u} \in \Delta_2$ we have,*

$$\mathbf{w} := \widetilde{\mathbf{v}} + \lambda(\mathbf{u} - \widetilde{\mathbf{v}}) = \left(\frac{3}{5} + \lambda\left(u_1 - \frac{3}{5}\right), \frac{2}{5} + \lambda\left(\frac{3}{5} - u_1\right)\right).$$

*Then we have,*

$$\widetilde{\mathbf{v}}^T Q \widetilde{\mathbf{v}} - \mathbf{w}^T Q \mathbf{w} = \frac{7}{5} - \left(\frac{7}{5} - \frac{5}{2}\lambda^2\left(u_1 - \frac{3}{5}\right)^2\right) = \frac{5}{2}\lambda^2\left(u_1 - \frac{3}{5}\right)^2.$$

*Clearly the above expression is positive for all $u_1 \neq \frac{3}{5}$. Hence $\widetilde{\mathbf{v}}$ is a strict local maximizer.*

An interesting property of ESS, as observed by Bishop and Cannings [14], is that if we add a constant to the columns of a matrix then the original and the resulting matrix have the same set of ESS.

**Lemma 3.26.** *Let $A \in \mathbb{R}^{m\times m}$ and $\mathbf{a}_i$ are the columns of $A$. Define the matrix $B$ with columns $\mathbf{b}_i = \mathbf{a}_i + \alpha_i \mathbf{e}$, where $\alpha_i \in \mathbb{R}$. Then $\overline{\mathbf{v}}$ is an ESS of $A$ if and only if $\overline{\mathbf{v}}$ is an ESS of $B$.*

*Proof.* Let $\mathbf{u}, \overline{\mathbf{v}} \in \Delta_m$ and $\alpha = (\alpha_1, \alpha_2, ..., \alpha_m)^T \in \mathbb{R}^m$. First note that $\mathbf{u}^T B = \mathbf{u}^T A + \alpha^T$. Then

$$\mathbf{u}^T B\overline{\mathbf{v}} = \mathbf{u}^T A\overline{\mathbf{v}} + \alpha^T\overline{\mathbf{v}}, \quad \overline{\mathbf{v}}^T B\overline{\mathbf{v}} = \overline{\mathbf{v}}^T A\overline{\mathbf{v}} + \alpha^T\overline{\mathbf{v}}$$

and

$$\overline{\mathbf{v}}^T B\overline{\mathbf{v}} - \mathbf{u}^T B\overline{\mathbf{v}} = \overline{\mathbf{v}}^T A\overline{\mathbf{v}} + \alpha^T\overline{\mathbf{v}} - \mathbf{u}^T A\overline{\mathbf{v}} - \alpha^T\overline{\mathbf{v}} = \overline{\mathbf{v}}^T A\overline{\mathbf{v}} - \mathbf{u}^T A\overline{\mathbf{v}}$$

Note that $\overline{\mathbf{v}}^T B\overline{\mathbf{v}} = \mathbf{u}^T B\overline{\mathbf{v}}$ if and only if $\overline{\mathbf{v}}^T A\overline{\mathbf{v}} = \mathbf{u}^T A\overline{\mathbf{v}}$. The second condition of ESS can be shown to hold in a similar way. □

For the case of symmetric matrices the above result is not useful in the sense that adding a different constant to each column may result in a matrix which is no more symmetric. So the following corollary is more useful in the case of symmetric matrices.

**Corollary 3.27.** *Let $\alpha \in \mathbb{R}$, then $\overline{\mathbf{v}} \in \Delta_m$ is an ESS of $A \in \mathcal{S}_m$ if and only if $\overline{\mathbf{v}}$ is an ESS of $A + \alpha\mathbf{e}\mathbf{e}^T$.*

*Proof.* Follows immediately from Lemma 3.26. □

### 3.4.1 Existence of ESS

In this subsection we will discuss necessary conditions for the existence of an ESS. We shall show that there exist matrices with no ESS, moreover every nonsingular $2 \times 2$ matrix has at least one ESS. We start the existence analysis of ESS with the following necessary conditions initially formulated by Haigh [79]. First we define the set,

$$S(\mathbf{v}) := \left\{ i : (Q\mathbf{v})_i = \max_j (Q\overline{\mathbf{v}})_j \right\}.$$

**Lemma 3.28.** *([79, Theorem 3]). Let $Q \in \mathbb{R}^{m\times m}$ and let $\overline{\mathbf{v}} \in \Delta_m$ be an ESS with respect to $Q$, then*

$$(Q\overline{\mathbf{v}})_i = \max_j (Q\overline{\mathbf{v}})_j \quad \forall\, i \in R(\overline{\mathbf{v}}) \;\; \text{``or equivalently''} \;\; R(\overline{\mathbf{v}}) \subseteq S(\overline{\mathbf{v}}), \tag{3.27}$$

*where $R(\overline{\mathbf{v}})$ denotes the support of the vector $\overline{\mathbf{v}}$ (cf.* (3.2)*).*

*Proof.* We suppose to the contrary, that there exists $j \in R(\overline{\mathbf{v}})$ such that, $s := \max_i (Q\overline{\mathbf{v}})_i > (Q\overline{\mathbf{v}})_j$. Then for $\mathbf{u} \in \Delta_m$ with $R(\mathbf{u}) \subseteq S(\overline{\mathbf{v}})$ it follows $\mathbf{u}^T Q\overline{\mathbf{v}} = \sum u_i (Q\overline{\mathbf{v}})_i = s$ and

$$\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} = \overline{\mathbf{v}}_j (Q\overline{\mathbf{v}})_j + \sum_{i\neq j} \overline{\mathbf{v}}_i (Q\overline{\mathbf{v}})_i < \sum_i \overline{\mathbf{v}}_i s = s = \mathbf{u}^T Q\overline{\mathbf{v}}$$

leading to a contradiction that $\overline{\mathbf{v}}$ is an ESS. □

It can be readily verified that for $Q \in \mathcal{S}_m$ the necessary conditions given in the above lemma are equivalent to the KKT conditions (3.1). This can be shown by taking $\lambda = \max_j (Q\mathbf{v})_j$ where $\lambda$ is Lagrange multiplier in (3.1). From this observation and for the same value of $\lambda$ we can also conclude that the set $\widetilde{S}(\overline{\mathbf{v}})$ (see (3.3)) is equal to $S(\overline{\mathbf{v}})$, i.e., $S(\overline{\mathbf{v}}) = \widetilde{S}(\overline{\mathbf{v}})$.

**Remark 3.29.** *It is interesting to note that if $\overline{\mathbf{v}}$ is an ESS with respect to the matrix $Q$, then equality in the first condition of ESS precisely occurs for those $\mathbf{u} \in \Delta_m$ for which $R(\mathbf{u}) \subset S(\overline{\mathbf{v}})$. This follows from Lemma 3.28. To see this let $s = \max_j (Q\overline{\mathbf{v}})_j$. Then* (3.27) *implies $\overline{\mathbf{v}}^T Q\overline{\mathbf{v}} = s$. Moreover $(Q\overline{\mathbf{v}})_i < s$ for all $i \notin S(\overline{\mathbf{v}})$, and if $R(\mathbf{u}) \not\subset S(\overline{\mathbf{v}})$ then*

$$\mathbf{u}^T Q\overline{\mathbf{v}} = s \sum_{i\in S(\overline{\mathbf{v}})} u_i + \sum_{i\notin S(\overline{\mathbf{v}})} u_i (Q\overline{\mathbf{v}})_i$$

$$< s \sum_{i \in S(\overline{\mathbf{v}})} u_i + s \sum_{i \notin S(\overline{\mathbf{v}})} u_i = s = \overline{\mathbf{v}}^T Q \overline{\mathbf{v}}$$

*so in the case $R(\mathbf{u}) \not\subset S(\overline{\mathbf{v}})$, we always get strict inequality in the first condition of ESS.*

In the following lemma we give sufficient conditions for the existence of an ESS for a matrix $Q$.

**Lemma 3.30** ([79]). *Let $Q \in \mathbb{R}^{m \times m}$. If for $i$ it holds that $q_{ii} > q_{ji}$ for all $j \neq i$ then $e_i$ is an ESS of $Q$.*

*Proof.* First note that $Qe_i$ will give the $i^{th}$ column of the matrix $Q$ and for $\mathbf{u} \in \Delta_m$ consider,

$$\begin{aligned} \mathbf{u}^T Q e_i &= q_{1i} u_1 + \ldots + q_{ii} u_i + \ldots + q_{mi} u_m \\ &< q_{ii}(u_1 + u_2 + \ldots + u_m) = q_{ii} = e_i^T Q e_i \ . \end{aligned}$$

Hence $e_i$ is an ESS. □

As mentioned before, for every nonsingular $2 \times 2$ matrix there exists an ESS.

**Lemma 3.31** ([79]). *Every nonsingular $2 \times 2$ matrix has at least one ESS.*

*Proof.* Let

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \ .$$

Now if $q_{11} > q_{21}$ or $q_{22} > q_{12}$, then there is an ESS due to Lemma 3.30. Now suppose that $q_{11} \leq q_{21}$ and $q_{22} \leq q_{12}$, in this case it can be readily verified that $\overline{\mathbf{v}} := (v_1, 1 - v_1)$ is an ESS with,

$$v_1 := \frac{q_{22} - q_{12}}{q_{12} + q_{21} - q_{11} - q_{22}} \ .$$

The denominator in the above expression can only be zero when $q_{11} = q_{21}, q_{22} = q_{12}$, which corresponds to a singular matrix. □

It is claimed by Vickers and Cannings [151, page 389] that every nonsingular symmetric matrix has an ESS. A proof of the claim is not given. As a matter of fact the example given below, with a nonsingular matrix without an ESS, provides a counter example to the claim.

**Example 3.32.** *Consider the nonsingular matrix*

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & -2 \end{pmatrix} .$$

*In order to show that the matrix $A$ does not have an ESS first note that $\overline{\mathbf{v}} := (v_1, 1-v_1, 0) \in \Delta_m$ cannot be an ESS since $\overline{\mathbf{v}}^T A\overline{\mathbf{v}} = 2$ for all $0 \le v_1 \le 1$. Moreover $\overline{\mathbf{v}} := (v_1, v_2, 1 - v_1 - v_2) \in \Delta_m$ with $0 \le v_1 + v_2 < 1$ cannot be an ESS. In order to see this consider*

$$A\overline{\mathbf{v}} = \begin{pmatrix} 1+v_1+v_2 \\ 1+v_1+v_2 \\ 3v_1+3v_2-2 \end{pmatrix}$$

*and note that if $\overline{\mathbf{v}}$ is an ESS then from Lemma 3.28 we should have, $1 + v_1 + v_2 = 3v_1 + 3v_2 - 2$ implying $v_1 + v_2 = \frac{3}{2}$, which is a clear contradiction.*

However the following is true,

**Corollary 3.33.** *Let $Q \in \mathcal{S}_m$ be such that each principle submatrix is nonsingular. Then $Q$ has an ESS.*

*Proof.* It is clear from Proposition 3.24 that an ESS corresponds to a strict local maximizer. Assume that the global maximizer is not a strict local maximizer. Choose such a global maximizer $\overline{\mathbf{v}}$ on rint($\mathrm{fc}_{R(\overline{\mathbf{v}})}$) with maximal support $R(\overline{\mathbf{v}})$. Then $\overline{\mathbf{v}}$ is not a strict local maximizer on $\mathrm{fc}_{R(\overline{\mathbf{v}})}$ and by Theorem 3.13 we have $\det(Q_{R(\overline{\mathbf{v}})}) = 0$, which is a contradiction. □

### 3.4.2 Patterns of ESS

In this subsection we shall give a brief survey of results related to the patterns of ESS. A pattern, roughly speaking, is a set of supports of ESS. In this subsection results on the patterns are provided for which it is known that they cannot exist. A complete enumeration of patterns for a matrix of order up to four is also presented.

**Definition 3.34** (Pattern of ESS). A *pattern* $P$ is a set of distinct subsets of the set $\mathcal{U} := \{1, \ldots, m\}$. We call a pattern *attainable* if there exists a matrix $Q \in \mathbb{R}^{m\times m}$ with ESS whose support corresponds to each of the subsets of $\mathcal{U}$ present in the pattern. An attainable pattern $P$ is called *maximal* if there is no $P^* \supset P$ which is attainable.

More specifically a pattern is a set $P = \{P_1, P_2, \ldots, P_k\}$ such that $P_i \subset \mathcal{U}$ and $P_i \neq P_j$ for all $i \neq j = 1, \ldots, k$. We call $P$ an attainable pattern if there exists a matrix $Q$ with $k$ ESS namely $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ whose set of supports is $R(\mathbf{v}_1), R(\mathbf{v}_2)$, ..., $R(\mathbf{v}_k)$ such that for each $i$, $R(\mathbf{v}_i) = P_i$. In this subsection we shall use $P$ to denote the pattern. For example $P = \{(1,2,3),(3,4),(1,3,5,6)\}$ means that the pattern consist of three subsets of $\mathcal{U}$ with set of supports $\{1,2,3\}$, $\{3,4\}$ and $\{1,3,5,6\}$.

The following is a well known conjecture concerning the attainable patterns.

**Conjecture 3.35** ([151]). *Let $P$ be an attainable pattern and $P^* \subset P$ then $P^*$ is also attainable.*

A weaker result is proven by Broom [34]. The conjecture is useful when it is required to find a complete list of attainable patterns. It is worth mentioning that there exists some patterns which are not attainable by symmetric matrices. The pattern $P := \{(1,2),(1,3),(2,3,4),(3,5),(4,5)\}$ is attained by a nonsymmetric matrix but it is not attainable by any $5 \times 5$ symmetric matrix [43, page. 197]. For the matrices of order 2,3,4, a complete list of attainable patterns is known and is discussed by Vickers and Cannings in a series of papers [42, 43, 151]. The description is based on some exclusion results stated and proven by Vickers and Canning. Here we will enlist these exclusion results starting with a simple result of Bishop and Cannings [15],

**Lemma 3.36** ([15]). *Let $Q \in \mathbb{R}^{m\times m}$ and let $\mathbf{v}, \mathbf{u} \in \Delta_m$ be two ESS with respect to $Q$, then*

$$R(\mathbf{v}) \not\subseteq S(\mathbf{u}) \quad \textit{and} \quad R(\mathbf{u}) \not\subseteq S(\mathbf{v})$$

*Proof.* Let us suppose to the contrary that $R(\mathbf{v}) \subseteq S(\mathbf{u})$. Since $\mathbf{u} \in \Delta_m$ is an ESS we have $\mathbf{u}^T Q\mathbf{u} \geq \mathbf{v}^T Q\mathbf{u}$, and due to the arguments given in Remark 3.29 only the equality is possible (i.e. $\mathbf{u}^T Q\mathbf{u} = \mathbf{v}^T Q\mathbf{u}$), in which case we get $\mathbf{u}^T Q\mathbf{v} > \mathbf{v}^T Q\mathbf{v}$, which contradicts $\mathbf{v}$ being an ESS. $\square$

The immediate consequence of the above lemma is that it excludes the possibility for the existence of two ESS such that the support of one is contained in the other. Another consequence of the above result is that, if there exists an ESS $\bar{\mathbf{v}} \in \Delta_m$ such that $|R(\bar{\mathbf{v}})| = m$, then it is unique.

Now we turn our attention to the exclusion results related to certain patterns. We start this discussion with matrices of size 3. Lemma 3.37 says that for a $3 \times 3$ matrix there cannot exists three ESS of support size two, simultaneously.

**Lemma 3.37.** *Let $Q \in \mathbb{R}^{3\times 3}$, then $P := \{(1,2),(2,3),(1,3)\}$ is not attainable.*

*Proof.* See [151]. □

In the following theorem we will enlist all patterns which are known to be not possible.

**Theorem 3.38.** *Let $Q \in \mathbb{R}^{m\times m}$, then the following patterns are not attainable,*

*i.* *with* $S \subseteq \mathcal{U}\backslash\{1,2,3\}$, $P := \{(1,2,S),(2,3,S),(1,3,S)\}$

*ii.* *with* $S \subseteq \mathcal{U}\backslash\{1,2,...,k\}$,
$P := \{(1,k,S),(2,k,S),...,(k-1,k,S),(1,2,...,k-1,S)\}$

*iii.* *with* $S \subseteq \mathcal{U}\backslash\{1,2,...,k\}$, $P := \{(1,S),(2,S),...,(k,S),(1,2,...,k)\}$

*iv.* *with* $k < m$, $P := \{(1,k+1,k+2,\cdots,m),(2,3,\cdots,k,k+1,k+2,\cdots,m),$
$(1,2),(1,3),\cdots,(1,k)\}$

*Proof.* Part *i.,ii.* and *iii.* are proven in [151] while part *iv.* is proven in [41, Theorem 6]. □

By applying the above exclusion results, Vicker and Cannings [151] have provided a complete list of maximal patterns for matrices of order up to 4. For matrices of order 5 a partial list is provided. In the following theorem the maximal attainable patterns for matrices of order up to four are enlisted. For examples showing the attainability, the interested reader is referred to [151].

**Theorem 3.39.** *For $m = 2,3,4$ the following is the complete list of maximal patterns which are attainable,*

$$\begin{aligned}
m=2: &\quad \{(1,2)\},\{(1),(2)\}\\
m=3: &\quad \{(1,2,3)\},\{(1,2),(1,3)\},\{(1,2),(3)\},\{(1),(2),(3)\}\\
m=4: &\quad \{(1,2,3,4)\},\{(1,2,3),(1,2,4)\},\{(1,2,3),(2,4),(3,4)\},\{(1,2,3),(4)\},\\
&\quad \{(1,2),(1,3),(1,4)\},\{(1,2),(2,3),(3,4),(1,4)\},\{(1,2),(1,3),(4)\},\\
&\quad \{(1,2),(3),(4)\},\{(1),(2),(3),(4)\}
\end{aligned}$$

The description of the maximal attainable patterns given in the above theorem is minimal in a sense that when we write $\{(1,2),(1,3)\}$ is attainable then any permutation is also attainable, i.e., $\{(1,2),(2,3)\}$,$\{(2,3),(1,3)\}$ are also attainable.

### 3.4.3 ESS in $\{0, \pm 1\}$ Matrices

In this subsection the existence of ESS in a special class of matrices is discussed. The matrix class is denoted by $\mathcal{M}$,

$$\mathcal{M} := \{Q \in \mathcal{S}_m : q_{ij} = \pm 1, q_{ii} = 0\}.$$

To each $Q \in \mathcal{M}$ one can associate a graph $G = (\mathcal{V}, \mathcal{E})$ with the set of vertices $\mathcal{V} = \{1, 2, ..., m\}$ and the set of edges $\mathcal{E}$ such that $\{i, j\} \in \mathcal{E}$ if and only if $q_{ij} = 1$. For these matrices and the corresponding graph, an ESS can be characterized by the maximal clique property.

**Definition 3.40** (Maximal Clique)**.** Let $\mathcal{V}$ be the set of vertices in the graph $G$. A subset $\overline{\mathcal{V}} \subseteq \mathcal{V}$ is called a clique if every pair of vertices in $\overline{\mathcal{V}}$ is connected by an edge. A clique is called maximal if it is not properly contained in some other clique.

The following results have appeared in [42]. For the sake of completeness, here, we include the proof of the theorem,

**Theorem 3.41** ([42])**.** *Let $Q \in \mathcal{M}$. Then there is an ESS (say $\overline{\mathbf{v}}$) with support $S := R(\overline{\mathbf{v}})$ if and only if S forms a maximal clique in the corresponding graph G.*

*Proof.* Let us suppose that $S \subset \mathcal{U}$ is a maximal clique in $G$. We define $\overline{\mathbf{v}} \in \Delta_m$ such that,

$$v_i = \begin{cases} \frac{1}{|S|} & \text{if } i \in S \\ 0 & otherwise \end{cases}$$

Now it is sufficient to prove that $\overline{\mathbf{v}}$ is a strict local maximizer. Let $\mathbf{u} \in \Delta_m$ with $\mathbf{u} \neq \overline{\mathbf{v}}$. Considering the convex combination, $\overline{\mathbf{w}} := \overline{\mathbf{v}} + \lambda(\mathbf{u} - \overline{\mathbf{v}}) \in \Delta_m$, then $\overline{\mathbf{v}}$ is a strict local maximizer if and only if, for all such $\mathbf{u} \in \Delta_m$ and $\lambda > 0$ small the following holds,

$$\overline{\mathbf{w}}^T Q \overline{\mathbf{w}} - \overline{\mathbf{v}}^T Q \overline{\mathbf{v}} = 2\lambda \overline{\mathbf{v}}^T Q(\mathbf{u} - \overline{\mathbf{v}}) + \lambda^2 (\mathbf{u} - \overline{\mathbf{v}})^T Q(\mathbf{u} - \overline{\mathbf{v}}) < 0 \tag{3.28}$$

First note that for the case $R(\mathbf{u}) \not\subset S$ we have $\overline{\mathbf{v}}^T Q \overline{\mathbf{v}} > \mathbf{u}^T Q \overline{\mathbf{v}}$. In order to see this observe that the principle submatrix $Q_S$ corresponding to $S$ does not contain negative entries so we have $(Q\overline{\mathbf{v}})_i = \frac{1}{|S|}(|S| - 1)$ for all $i \in S$. Moreover since $S$ forms a maximal clique, for all $i \notin S$ it follows that $(Q\overline{\mathbf{v}})_i < \frac{1}{|S|}(|S| - 1)$, which results in

$$\mathbf{u}^T Q \overline{\mathbf{v}} = \sum_{i \in S} (Q\overline{\mathbf{v}})_i u_i + \sum_{i \notin S} (Q\overline{\mathbf{v}})_i u_i$$

$$
\begin{aligned}
&= \frac{1}{|S|}(|S|-1)\sum_{i\in S} u_i + \sum_{i\notin S}(Q\overline{\mathbf{v}})_i u_i \\
&< \frac{1}{|S|}(|S|-1) = \overline{\mathbf{v}}^T Q\overline{\mathbf{v}}\,.
\end{aligned}
$$

Now consider the case $R(\mathbf{u}) \subset S$. Then $Q_S = E - I$ where $E$ is the $|S| \times |S|$ matrix consisting of all ones while $I$ is the identity matrix of order $|S|$. Observe also that for $\overline{\mathbf{v}}$ the following holds,

$$
\begin{aligned}
\overline{\mathbf{v}}^T Q(\mathbf{u}-\overline{\mathbf{v}}) &= \frac{1}{|S|}\mathbf{e}_S^T Q_S(\mathbf{u}_S - \frac{1}{|S|}\mathbf{e}_S) = \frac{1}{|S|}\mathbf{e}_S^T(E-I)(\mathbf{u}_S - \frac{1}{|S|}\mathbf{e}_S) \\
&= \frac{1}{|S|}\mathbf{e}_S^T(E\mathbf{u}_S - \frac{1}{|S|}E\mathbf{e}_S - I\mathbf{u}_S + \frac{1}{|S|}I\mathbf{e}_S) \\
&= \frac{1}{|S|}\mathbf{e}_S^T(\mathbf{e}_S\mathbf{e}_S^T\mathbf{u}_S - \frac{|S|}{|S|}\mathbf{e}_S - \mathbf{u}_S + \frac{1}{|S|}\mathbf{e}_S) \\
&= \frac{1}{|S|}\mathbf{e}_S^T(\mathbf{e}_S - \mathbf{e}_S - \mathbf{u}_S + \frac{1}{|S|}\mathbf{e}_S) \text{ since } \mathbf{e}^T\mathbf{u} = 1 \\
&= \frac{1}{|S|}(-\mathbf{e}_S^T\mathbf{u}_S + \frac{1}{|S|}\mathbf{e}_S^T\mathbf{e}_S) = 0 \qquad (3.29)
\end{aligned}
$$

where $\mathbf{e}_S \in \mathbb{R}_+^{|S|}$ is the vector of all ones.

Since $R(\mathbf{u}) \subset S$ for $\mathbf{o} \neq \mathbf{w} = \mathbf{u} - \overline{\mathbf{v}}$ it holds $w_i = 0$ for all $i \notin S$. So

$$
\begin{aligned}
\mathbf{w}^T Q\mathbf{w} &= \sum_{\substack{i,j\in S\\ i\neq j}} q_{ij}w_iw_j + \sum_{\substack{i,j\notin S\\ i\neq j}} q_{ij}w_iw_j \\
&= \sum_{\substack{i,j\in S\\ i\neq j}} q_{ij}w_iw_j = \sum_{\substack{i,j\in S\\ i\neq j}} w_iw_j \\
&= \left(\left(\sum_{i\in S} w_i\right)^2 - \sum_{i\in S} w_i^2\right) = -\sum_{i\in S} w_i^2 < 0. \qquad (3.30)
\end{aligned}
$$

The last equality follows since $\mathbf{e}^T\mathbf{w} = 0$. So the inequality in (3.28) follows from (3.29) and (3.30).

For the converse let $\overline{\mathbf{v}}$ be an ESS with support $R(\overline{\mathbf{v}})$. If $R(\overline{\mathbf{v}})$ does not form a maximal clique in the corresponding graph then there may occur two possibilities

i. $R(\overline{\mathbf{v}})$ forms a clique but not a maximal clique. In this case there exists some $S \subset \mathcal{U}$ such that $R(\overline{\mathbf{v}}) \subset S$ where $S$ forms a maximal clique in the graph $G$.

ii. $R(\bar{\mathbf{v}})$ does not form a clique at all, meaning that there exists some $i, j \in R(\bar{\mathbf{v}})$ such that $q_{ij} = -1$. In this case there exists some $S \subset R(\bar{\mathbf{v}})$, such that $S$ forms a maximal clique in the corresponding graph.

Now from the first part of the theorem, $S$ corresponds to the support of some ESS **u**. Since $\bar{\mathbf{v}}$ is an ESS and the support $R(\bar{\mathbf{v}})$ is either contained in $S$ or it is containing $S$, we obtain a contradiction to Lemma 3.36. □

Interestingly all patterns given in Theorem 3.39 are attainable by matrices from the class $\mathcal{M}$ with the exception of the pattern $\{(1,2,3),(2,4),(3,4)\}$ (for details see [43, page 196]).

**Remark 3.42.** *The result of Theorem 3.41 can be generalized to matrices with elements from $\{\alpha, \beta, \gamma\}$ such that $\alpha < \beta < \gamma$, with $\beta$ on the main diagonal. For these matrices the graph $G := (\mathcal{V}, \mathcal{E})$ can be associated with $\mathcal{V} = \mathcal{U}$ and $\{i, j\} \in \mathcal{E}$ if and only if $q_{ij} = \gamma$.*

### 3.4.4 Number of ESS

In this subsection the question of the maximum number of ESS which can coexists in a matrix $Q \in \mathbb{R}^{m\times m}$ is discussed. The subsection starts with a well known lemma from combinatorics which is helpful to obtain a bound on the maximum number of ESS. The bound for the special case of the matrix class $\mathcal{M}$ is also provided with an example proving that the bound is sharp. We will also state results for ESS with specific support size.

**Lemma 3.43** (Sperner's Lemma)**.** *Let $S$ be a Sperner set of subsets of $\mathcal{U} := \{1, \ldots, m\}$ (i.e. for $A, B \in S$, if $A \neq B$, then $A \not\subset B$ and $B \not\subset A$). Then $|S| \leq \binom{m}{\lfloor \frac{1}{2}m \rfloor}$, where for $a \in \mathbb{R}_+$, $\lfloor a \rfloor$ gives the largest integer less then or equal to $a$.*

*Proof.* See e.g. [36]. □

In view of Sperner's Lemma and Lemma 3.36 a bound on the maximum number of ESS becomes apparent. For sake of completeness, the bound is provided in the following proposition,

**Proposition 3.44.** *Let $Q \in \mathbb{R}^{m\times m}$. Then $Q$ can have at most $\binom{m}{\lfloor \frac{m}{2} \rfloor}$ ESS.*

*Proof.* Let $\mathbf{w}_1, \cdots, \mathbf{w}_N$ be ESS. Then from Lemma 3.36 we get $R(\mathbf{w}_i) \not\subset R(\mathbf{w}_j)$ for all $i \neq j, i, j = 1, \cdots, N$. By Sperner's Lemma the maximum number of $R(w_j)$'s in $\mathcal{U}$ such that no one is contained in some other is given by $\binom{m}{\lfloor \frac{m}{2} \rfloor}$. □

**Remark 3.45.** *From Theorem 3.39 it can be concluded that the bound given above is not sharp for the matrices of order* $4$*.*

Proposition 3.44 provides a bound on the maximum number of ESS that may coexists for $Q \in \mathbb{R}^{m\times m}$. For the class of matrices $\mathcal{M}$ the upper bound is different and achievable as is shown in the following lemma and example.

**Lemma 3.46** ([42]). *Let $Q \in \mathcal{M}$. If $m \geq 4$ and $m = r + 3s$, where $r = 2, 3, 4$, then the greatest number of ESS that can coexist in $Q$ is $r3^s$.*

*Proof.* Due to Lemma 3.41 the support of each ESS corresponds to a maximal clique in the associated graph $G$. The maximum number of maximal cliques in a graph associated with $Q$ is bounded by $r3^s$ by the result of Moon and Moser [113]. □

The following example shows that the bound given in Lemma 3.46 is tight.

**Example 3.47.** *Consider the matrix $Q \in \mathcal{S}_m$, $m = 3n$, $n \geq 2$*

$$Q = \begin{pmatrix} I-E & E & \cdots & E \\ E & I-E & \cdots & E \\ \vdots & \vdots & \ddots & \vdots \\ E & E & \cdots & I-E \end{pmatrix}$$

*where $E$ is the $3\times 3$ matrix of all ones and $I$ is the $3\times 3$ identity matrix. Define $n$ sets $S_1, S_2, \cdots, S_n$ such that $S_i := \{1+3(i-1), 2+3(i-1), 3+3(i-1)\}$ with $i = 1, \cdots, n$ and $P := \{S_1 \times S_2 \times \cdots \times S_n\}$. Now for any $S \in P$ we have*

$$(Q_S)_{i,j} := \begin{cases} 0 & i = j \\ 1 & otherwise \end{cases}$$

*i.e., $S$ forms a (maximal) clique in the corresponding graph, so each $S \in P$ is associated with an ESS, and there will be in total $|P| = 3^n$ of them. Now define (as we did in Theorem 3.41),*

$$v_i := \begin{cases} \frac{1}{|S|} & i \in S \\ 0 & otherwise \end{cases}$$

*then $\mathbf{v}^T Q\mathbf{v} = \frac{|S|-1}{|S|} = 1 - \frac{1}{n}$ will be the value of each ESS.*

**Number of ESS with Fixed Support**

In the literature, instead of giving sharp bounds on the maximum number of ESS much emphasis is given on the number of ESS with specific support size. Note that sharper bounds for each support size then lead to sharper bounds for the maximum number of ESS. In the context of $(StQP)$ this gives the number of strict local maximizers on a face of $\Delta_m$ of a certain dimension.

Let us denote by $u_m(r)$ the maximum number of ESS for $Q \in \mathbb{R}^{m\times m}$ with support of length $r$ with equality used when it is known that the bound described is tight. In the following theorem all known results for $u_m(r)$ are summarized,

**Theorem 3.48.** *Let $Q \in \mathbb{R}^{m\times m}$ then,*

*i.* $u_m(2) = \left\lfloor \frac{1}{4}m^2 \right\rfloor$

*ii.* $u_m(3) \leq \left\lfloor \frac{m^3-3m^2+6m-13}{13} \right\rfloor$

*iii.* $u_m(m-1) = 2$

*iv.* $u_m(m-2) = m$

*v.* $u_m(m-3) \leq \left\lfloor \frac{1}{3}m(m-1) \right\rfloor$

*Proof.* See [34]. □

### 3.4.5 ESS in Random Matrices

It is shown in the previous subsections that a matrix can have many ESS, as well as, that there exist matrices with no ESS. This situation leads one to think in probabilistic terms, meaning one might think about the probability for a given matrix to have an ESS? The question has been analysed for randomly generated symmetric and nonsymmetric matrices. Here we will briefly summarise these results.

The question of existence of ESS in randomly generated matrices is dealt indirectly. Instead of analysing the existence of ESS the attention is given to the question of existence of ESS with certain support size. The first result in this direction is obtained for the ESS of support size one. It has been shown that the probability for the existence of an ESS of support size one goes to $1 - \frac{1}{e}$ as the size of the matrix goes to infinity [80, 102]. Here it is worth mentioning that the result is independent of the distribution used for generating the elements of the matrix. The only other case analysed is for the support of size two. In this case the distribution for the generation of elements of the matrix play an important role since the results are dependent on the distribution. Hart et al. [85] have shown that if the elements of the matrix $Q \in \mathbb{R}^{m\times m}$ are generated using a distribution F then,

- for distributions F with ``exponential and faster decreasing tails'' (e.g. uniform, normal, exponential), we have

$$\lim_{m\to\infty} Pr(\exists \text{ ESS with support size} = 2) = 1$$

- for distributions F with ``slower than exponential decreasing tails'' (e.g. lognormal, Pareto, Cauchy) we have

$$\lim_{m\to\infty} Pr(\exists \text{ ESS with support size} = 2) = 1 - \frac{1}{\sqrt{e}}$$

Kingman [102] analysed the question of support size in large randomly generated symmetric matrices. He has shown that for symmetric matrices whose elements are drawn randomly using the uniform distribution, the probability of existence of an ESS with support size greater than or equal to $2.49m^{\frac{1}{2}}$ goes to zero as $m$ goes to infinity [81, 102]. Here the bound is in dependence of the probability distribution used. Haigh [80] extended the work of Kingman by proving a similar result for nonsymmetric matrices. For the case of nonsymmetric matrices the probability of the existence of an ESS with support size greater than $1.63m^{\frac{2}{3}}$ goes to zero as $m$ goes to infinity.

Less attention has been given to the question of finding bounds on the number of ESS in randomly generated matrices. The only result known is for the number of ESS with support size two. It is shown that the number of ESS with suport size two goes to $\frac{1}{3}log(\frac{m}{2})$ as $m$ goes to infinity[81].

## 3.5 Vector Iterations

In this section we consider vector iterations for solving $(StQP)$. We will also discuss a similar well-known algorithm to solve a similar program. We start with a special program which is used for finding the maximum eigenvalue of a matrix $Q \in \mathcal{S}_m$.

$$(E\text{-}Max) \qquad \max \quad \mathbf{v}^T Q\mathbf{v} \quad s.t. \qquad \|\mathbf{v}\| = 1, \quad \mathbf{v} \in \mathbb{R}^m,$$

where, as usual, $\|.\|$ denotes the Euclidean norm. Here without loss of generality we assume that $Q$ is positive definite i.e. $Q \in \mathcal{S}_m^{++}$, since $Q$ and $Q + \alpha I, \alpha \in \mathbb{R}$, $(E\text{-}Max)$ will have the same maximizers. A solution of the above program gives the eigen vector corresponding to the maximum eigenvalue of the matrix $Q$ (for details see e.g. [67, Section 4.8]). Although $(E\text{-}Max)$ is a special instance of a quadratic program it is well known to be polynomial time solvable since it can be reformulated as a semidefinite programming problem. In fact $(E\text{-}Max)$ can be used as a feasibility test for semidefinite programming problems.

The power method is a well known method for solving $(E\text{-}Max)$ (see e.g. [75]). The power method starts with an initial vector $\mathbf{v}(0)$ with $\|\mathbf{v}(0)\| = 1$, and iterates:

$$\mathbf{v}(t+1) := \frac{Q\mathbf{v}(t)}{\|Q\mathbf{v}(t)\|} \tag{3.31}$$

$$\lambda^{t+1} := \mathbf{v}(t+1)^T Q\mathbf{v}(t+1) \quad t = 0, 1, \ldots \; . \tag{3.32}$$

Here $t$ is the variable for the iteration. The power iteration is guaranteed to converge under the conditions that the matrix $Q$ has a unique dominating (positive) eigenvalue and the initial vector does not have a nonzero component in the direction of the eigenvector associated with the dominating eigenvalue. Here by dominating eigenvalue we mean the eigenvalue with the largest absolute value.

In the following theorem we will show that the iteration (3.32) is monotonically increasing for positive definite matrices. The proof is based on an unpublished manuscript [135].

**Theorem 3.49.** *Let $Q \in \mathcal{S}_m^{++}$, then $\lambda^t \leq \lambda^{t+1}$.*

*Proof.* First observe that $\mathbf{v}(t)$ can be written as follows,

$$\mathbf{v}(t) := \frac{Q^t\mathbf{v}(0)}{\|Q^t\mathbf{v}(0)\|}.$$

Then $\lambda^t$ can be written as,

$$\lambda^t := \frac{\mathbf{v}(0)^T Q^{2t+1}\mathbf{v}(0)}{\|Q^t\mathbf{v}(0)\|^2} \tag{3.33}$$

For the sake of clarity in the proof we will use $\mathbf{v}$ instead of $\mathbf{v}(0)$. By defining $s_k := \mathbf{v}^T Q^k \mathbf{v}$, (3.33) can be written as $\lambda^t = \frac{s_{2t+1}}{s_{2t}}$. Note also that $s_k > 0$ holds since $Q \in \mathcal{S}_m^{++}$. So in order to prove the theorem we will show the following:

$$\frac{s_{2t+1}}{s_{2t}} \leq \frac{s_{2t+3}}{s_{2t+2}} \tag{3.34}$$

From $L_1 := \|s_{2t+1}Q^t\mathbf{v} - s_{2t}Q^{t+1}\mathbf{v}\|^2$ we get,

$$\begin{aligned}
0 \leq L_1 &= \|s_{2t+1}Q^t\mathbf{v} - s_{2t}Q^{t+1}\mathbf{v}\|^2 \\
&= (s_{2t+1}Q^t\mathbf{v} - s_{2t}Q^{t+1}\mathbf{v})^T(s_{2t+1}Q^t\mathbf{v} - s_{2t}Q^{t+1}\mathbf{v}) \\
&= s_{2t+1}^2\mathbf{v}^T Q^{2t}\mathbf{v} - 2s_{2t+1}s_{2t}\mathbf{v}^T Q^{2t+1}\mathbf{v} + s_{2t}^2\mathbf{v}^T Q^{2t+2}\mathbf{v} \\
&= s_{2t+1}^2 s_{2t} - 2s_{2t+1}^2 s_{2t} + s_{2t}^2 s_{2t+2} \\
&= s_{2t+2}s_{2t}^2 - s_{2t+1}^2 s_{2t}
\end{aligned} \tag{3.35}$$

Divide (3.35) by $s_{2t}^2 s_{2t+1}$ to obtain,

$$\frac{s_{2t+1}}{s_{2t}} \leq \frac{s_{2t+2}}{s_{2t+1}} \tag{3.36}$$

Now consider $L_2 := \langle s_{2t+3}Q^{t+1}\mathbf{v} - s_{2t+2}Q^{t+2}\mathbf{v}, s_{2t+3}Q^t\mathbf{v} - s_{2t+2}Q^{t+1}\mathbf{v}\rangle$ and note that $L_2 \geq 0$ since $Q \in \mathcal{S}_m^{++}$, then we have the following,

$$\begin{aligned}
0 \leq L_2 &= s_{2t+3}^2 s_{2t+1} - 2s_{2t+2}^2 s_{2t+3} + s_{2t+2}^2 s_{2t+3} \\
&= s_{2t+1}s_{2t+3}^2 - s_{2t+2}^2 s_{2t+3}
\end{aligned} \tag{3.37}$$

Divide (3.37) by $s_{2t+1}s_{2t+2}s_{2t+3}$ to obtain,

$$\frac{s_{2t+2}}{s_{2t+1}} \leq \frac{s_{2t+3}}{s_{2t+2}} \tag{3.38}$$

Combining (3.36), (3.38) we find (3.34). □

We consider a similar iteration which can be associated (as we will see) with $(StQP)$. Start with $\mathbf{v}(0) \in \Delta_m$ and iterate:

$$v_i(t+1) = v_i(t)\frac{[Q\mathbf{v}(t)]_i}{\mathbf{v}(t)^T Q\mathbf{v}(t)} \quad i \in \mathcal{U} := \{1, 2, \ldots, m\}, \;\; t = 0, 1, \ldots . \tag{3.39}$$

Throughout this section the matrix $Q \in \mathcal{S}_m$ is assumed to be positive, since from Corollary 3.27 it is clear that $Q$ and $Q + \alpha E$ have the same strict local maximizers.

A point $\mathbf{v} = \mathbf{v}(t)$ is said to be a *fixed point* of (3.39) if $\mathbf{v}(t+1) = \mathbf{v}(t)$ holds in (3.39). It is not difficult to verify that the solution of the following system of equations gives the set of all fixed points of (3.39),

$$v_i[[Q\mathbf{v}]_i - \mathbf{v}^T Q\mathbf{v}] = 0, \quad \forall\, i \in \mathcal{U}\,. \tag{3.40}$$

**Remark 3.50.** *It is worth mentioning that the iteration* (3.39) *has a nice interpretation in* theoretical biology *and* population genetics *which goes as follows (see e.g. [119, 147]). Consider an infinite population of the same species contesting for a particular limited resource. If we take randomly chosen members of the population as players then this kind of conflicts can be modelled as a game, where each player acts according to a pre programmed behaviour termed as pure strategy. As usual let $\mathcal{U}$ denote all pure strategies and let $v_i(t)$ be the relative frequency of the members of the population playing strategy $i$, at time $t$. Then the vector $\mathbf{v}(t) = (v_1(t), \cdots, v_m(t))^T$ will denote the state of the system at time $t$. We further assume that the sum of relative frequencies is one, i.e., $\mathbf{v}(t) \in \Delta_m$. If we denote the advantage or payoff for a user of strategy $i$ whose opponent is playing strategy $j$ by $q_{ij}$ then the complete set of payoffs are denoted by a matrix $Q = (q_{ij})$. In this context the average payoff for the user of strategy $i$ will be $e_i Q\mathbf{v} = (Q\mathbf{v})_i$ [119, 147]. In theoretical biology iterations* (3.39) *are known as* replicator dynamics *while in population genetics they are called* selection equations*.*

Here, from a mathematical point of view, we are interested to know if the iteration (3.39) has some monotonicity properties, and whether starting with an initial point the iteration converges to a strict local maximizer of $(StQP)$ or not. The answer to the first question is positive. In the literature there exists many proofs for the monotonicity of the iteration (3.39). Here we will reproduce the elegant proof of Kingman [103]. Before the proof we provide some auxiliary results.

**Lemma 3.51** (Jensen Inequality)**.** *Let $f$ be a convex function on a convex set $S \subseteq \mathbb{R}$. Then for all $\lambda_i \geq 0, u_i \in S, i = 1, \ldots, N$, with $\sum_{i=1}^{N} \lambda_i = 1$ for $N \in \mathbb{N}$ we have,*

$$f\left(\sum_{i=1}^{N} \lambda_i u_i\right) \leq \sum_{i=1}^{N} \lambda_i f(u_i)$$

*Equality holds if and only if either $f$ is linear or $u_1 = u_2 = \cdots = u_N$.*

*Proof.* See e.g. [134]. □

If we take $f(u) = u^l, u \in \mathbb{R}_+, l \geq 0$ in the Jensen inequality, then we obtain,

$$\left[\sum_{i=1}^{N} \lambda_i u_i\right]^l \leq \sum_{i=1}^{N} \lambda_i (u_i)^l \tag{3.41}$$

As mentioned before, for $l > 1$ equality in (3.41) is possible if and only if $u_1 = u_2 = \cdots = u_N$. Note also that for $a, b \geq 0$ we have,

$$\frac{a+b}{2} \geq \sqrt{ab} \tag{3.42}$$

**Theorem 3.52.** *Let $Q \in \mathcal{S}_m$ be positive and let $\mathbf{v}(t) \in \Delta_m$ not be a fixed point. Then*

$$\mathbf{v}(t+1)^T Q \mathbf{v}(t+1) > \mathbf{v}(t)^T Q \mathbf{v}(t).$$

*Proof.* First observe that,

$$\begin{aligned}
\mathbf{v}(t+1) Q \mathbf{v}(t+1) &= \sum_{i,j} v_i(t+1) v_j(t+1) q_{ij} \\
&= \sum_{i,j} v_i(t) \frac{[Q\mathbf{v}(t)]_i}{\mathbf{v}(t)^T Q \mathbf{v}(t)} v_j(t) \frac{[Q\mathbf{v}(t)]_j}{\mathbf{v}(t)^T Q \mathbf{v}(t)} q_{ij} \\
&= \frac{1}{[\mathbf{v}(t)^T Q \mathbf{v}(t)]^2} \sum_{i,j} v_i(t) v_j(t) [Q\mathbf{v}(t)]_i [Q\mathbf{v}(t)]_j q_{ij}
\end{aligned}$$

In view of the above observation it is sufficient to prove:

$$\sum_{i,j} v_i(t) v_j(t) [Q\mathbf{v}(t)]_i [Q\mathbf{v}(t)]_j q_{ij} > [\mathbf{v}(t)^T Q \mathbf{v}(t)]^3$$

For the sake of clarity we will use $\mathbf{v}$ instead of $\mathbf{v}(t)$ in the rest of the proof. Take $L := \sum_{i,j} v_i v_j [Q\mathbf{v}]_i [Q\mathbf{v}]_j q_{ij}$ and note that $[Q\mathbf{v}]_i = \sum_k q_{ik} v_k$. Then we have,

$$L = \sum_{i,j,k} v_i v_j v_k [Q\mathbf{v}]_j q_{ij} q_{ik}$$

By interchanging $j$ with $k$ we obtain the following two equivalent forms of $L$,

$$L = \sum_{i,j,k} v_i v_j v_k [Q\mathbf{v}]_j q_{ij} q_{ik} = \sum_{i,j,k} v_i v_j v_k [Q\mathbf{v}]_k q_{ij} q_{ik}$$

Now adding the two expressions for $L$ we get,

$$
\begin{aligned}
L &= \sum_{i,j,k} v_i v_j v_k \frac{1}{2}([Q\mathbf{v}]_j + [Q\mathbf{v}]_k) q_{ij} q_{ik} \\
&\geq \sum_{i,j,k} v_i v_j v_k ([Q\mathbf{v}]_j [Q\mathbf{v}]_k)^{\frac{1}{2}} q_{ij} q_{ik} \quad \text{using (3.42)} \\
&= \sum_i v_i \left[ \sum_{j,k} v_j v_k ([Q\mathbf{v}]_j [Q\mathbf{v}]_k)^{\frac{1}{2}} q_{ij} q_{ik} \right] \\
&= \sum_i v_i \left[ \sum_j v_j ([Q\mathbf{v}]_j)^{\frac{1}{2}} q_{ij} \right]^2 \\
&\geq \left[ \sum_i v_i \sum_j v_j ([Q\mathbf{v}]_j)^{\frac{1}{2}} q_{ij} \right]^2 \quad \text{Using (3.41) with } l = 2 \qquad (3.43) \\
&= \left[ \sum_j v_j ([Q\mathbf{v}]_j)^{\frac{1}{2}} \left[ \sum_i v_i q_{ij} \right] \right]^2 \\
&= \left[ \sum_j v_j ([Q\mathbf{v}]_j)^{\frac{3}{2}} \right]^2 \\
&> \left[ \left[ \sum_j v_j [Q\mathbf{v}]_j \right]^{\frac{3}{2}} \right]^2 \quad \text{Using (3.41) with } l = 3/2 \qquad (3.44) \\
&= \left[ \sum_j v_j ([Q\mathbf{v}]_j) \right]^3 = \left[ \mathbf{v}^T Q \mathbf{v} \right]^3
\end{aligned}
$$

Here we would like to emphasize that the inequality (3.44) is strict. In order to see this, assume that equality holds. Then (in view of Lemma 3.51) for all $v_j > 0$ we have $[Q\mathbf{v}]_j = \alpha$ giving that:

$$
\left[ \sum_j v_j ([Q\mathbf{v}]_j)^{\frac{3}{2}} \right]^2 = \left[ \sum_j v_j (\alpha)^{\frac{3}{2}} \right]^2 = \left[ \alpha^{\frac{3}{2}} \sum_j v_j \right]^2 = \alpha^3,
$$

where we have used $\mathbf{v} \in \Delta_m$. So, also

$$\alpha^3 = \left[\sum_j v_j([Q\mathbf{v}]_j)^{\frac{3}{2}}\right]^2 = \left[\sum_j v_j([Q\mathbf{v}]_j)\right]^3 = (\mathbf{v}^T Q\mathbf{v})^3 .$$

Hence, we have $\alpha = \mathbf{v}^T Q\mathbf{v}$ and thus $[Q\mathbf{v}]_j = \mathbf{v}^T Q\mathbf{v} = \alpha$ for all $v_j > 0$. This contradicts the assumption that $\mathbf{v}$ is not a fixed point.

For alternative proofs see [22, 116, 137]. □

From the results of Losert and Akin [109] it follows that the iterations (3.39) will converge to a fixed point. Now the question arises if the iteration converges to a strict local maximizer or an ESS. The answer to this question is negative, since there exist matrices with no strict local maximizers. So it is clear that starting with some initial point the iteration may not converge to a strict local maximizer. Now the following question arises. Let the initial point $\mathbf{v}(0)$ of the iteration be very close to an ESS (say $\overline{\mathbf{v}}$). Will then the iteration converge to $\overline{\mathbf{v}}$? The answer to this question is positive.

**Theorem 3.53.** *Let $Q \in \mathcal{S}_m$ be positive and let $\overline{\mathbf{v}} \in \Delta_m$ be an ESS. Then there exists $\varepsilon > 0$ such that for all $\mathbf{v}(0)$ satisfying $\|\mathbf{v}(0) - \overline{\mathbf{v}}\| \leq \varepsilon$ we have $\mathbf{v}(t) \to \overline{\mathbf{v}}$ as $t \to \infty$.*

*Proof.* See [22, Theorem 3]. □

## 3.6 Genericity

In this section we will discuss so called genericity results for $(StQP)$. First we will specify what is exactly meant by genericity.

**Definition 3.54.** We say that a property is generic in the problem set $\mathcal{S}_m$, if the property holds for a (generic) subset $\mathcal{Q}_r$ of $\mathcal{S}_m$ such that $\mathcal{Q}_r$ is open and $\mathcal{S}_m \backslash \mathcal{Q}_r$ has (Lebesgue) measure zero.(So genericity implies density and stability of the set $\mathcal{Q}_r$ of ``nice'' problem instances).

First consider the following lemma required in the proof of the next theorem,

**Lemma 3.55.** *Let $p : \mathbb{R}^k \to \mathbb{R}$ be a polynomial mapping, $p \neq 0$. Then the set of zeros of $p$, $p^{-1}(0) = \{v \in \mathbb{R}^k : p(v) = 0\}$, has (Lebesgue) measure zero (in $\mathbb{R}^k$).*

*Proof.* See e.g. [16, Lemma 2.8]. □

In the next theorem we show that generically any local maximizer $\overline{\mathbf{v}}$ of $(StQP)$ is a strict local maximizer, i.e., an ESS that furthermore satisfies $R(\overline{\mathbf{v}}) = S(\overline{\mathbf{v}})$.

**Theorem 3.56.** *There is a generic subset $\mathcal{Q}_r \subset \mathcal{S}_m$ such that for any $Q \in \mathcal{Q}_r$ the following holds: For any $\overline{\mathbf{v}} \in \Delta_m$ such that $\overline{\mathbf{v}}$ is a local maximizer we have,*

*i.* $R(\overline{\mathbf{v}}) = S(\overline{\mathbf{v}})$

*ii.* $\overline{\mathbf{v}}$ *is a strict local maximizer*

*Proof. i.* Since $\overline{\mathbf{v}}$ is a local maximizer, from the KKT conditions it follows that $R(\overline{\mathbf{v}}) \subseteq S(\overline{\mathbf{v}})$. Suppose that this inclusion is strict i.e. $R(\overline{\mathbf{v}}) \neq S(\overline{\mathbf{v}})$. Then there exists some $j \in S(\overline{\mathbf{v}}) \backslash R(\overline{\mathbf{v}})$. This means that with $R := R(\overline{\mathbf{v}})$ the point $\overline{\mathbf{v}}_R \in \mathbb{R}^{|R|}_{++}$ solves the system of linear equations,

$$\begin{pmatrix} Q_R \\ q_{j,R} \end{pmatrix} \overline{\mathbf{v}}_R = \lambda \begin{pmatrix} e_R \\ 1 \end{pmatrix} \quad \text{with } \lambda := \max_i (Q\overline{\mathbf{v}})_i \tag{3.45}$$

where $q_{j,R} := (q_{j,l}, l \in R)$. This implies that the determinant of the $(|R|+1) \times (|R|+1)$ matrix $\left( \begin{smallmatrix} Q_R & e_R \\ q_{j,R} & 1 \end{smallmatrix} \right)$ is zero.

Consider now the polynomial function $p(Q, q_{j,R}) := \det \left( \begin{smallmatrix} Q_R & e_R \\ q_{j,R} & 1 \end{smallmatrix} \right)$. Since $p(I_R, 0) = 1$ this function is nonzero and according to Lemma 3.55 for almost all $Q_R$, $q_{j,R} \in \mathbb{R}^{|R| \times (|R|+1)}$ the relation $p(Q_R, q_{j,R}) \neq 0$ holds, i.e., there is no solution of the equation (3.45). Moreover since the function $p(Q_R, q_{j,R})$ is continuous the set of parameters $(Q_R, q_{j,R})$ with $p(Q_R, q_{j,R}) \neq 0$ is open. Since there is only a finite selection of subsets $R \subset \mathcal{U}$ and elements $j \notin R$ possible, also the set of parameters $Q$ such that for all $R, j$ the condition $p(Q_R, q_{j,R}) \neq 0$ holds is generic.

*ii.* Now suppose that a local maximizer $\overline{\mathbf{v}}$ (by the above analysis we can assume $R(\overline{\mathbf{v}}) = S(\overline{\mathbf{v}})$) is not a strict local maximizer. Then in view of Corollary 3.14 we have, $\det(Q_{R(\overline{\mathbf{v}})}) = 0$. But by defining the non-zero polynomial $p(Q) := \det(Q_{R(\overline{\mathbf{v}})})$ and using Lemma 3.55 the conditions $\det(Q_{R(\overline{\mathbf{v}})}) = 0$ can be excluded for almost all $Q$. By noticing that also the condition $\det(Q_{R(\overline{\mathbf{v}})}) \neq 0$ is stable with respect to small perturbations of $Q$ the condition $\det(Q_{R(\overline{\mathbf{v}})}) = 0$ is generically excluded. $\square$

The above theorem immediately implies that generically every symmetric matrix has an ESS, i.e. :

**Corollary 3.57.** *The set $\{A \in \mathcal{S}_m : A$ has an ESS$\}$ contains a generic subset.*

*Proof.* For every matrix $A$, $(StQP)$ has a global maximizer. By Theorem 3.56 generically it is an ESS. $\square$

# 4

# Nonconvex Quadratic Programming[1]

ZOOMING into the intersection of convex and nonlinear programming problems we study a list of problems which are originally nonconvex but by the use of relaxation techniques these programs are reformulated (approximately) as convex programming problems. In this chapter we investigate how sharp the set-semidefinite relaxations of nonconvex quadratic programs are.

## 4.1 Introduction

As mentioned before it is common to solve/approximate programs with binary or general quadratic constraints by considering their semidefinite or copositive programming relaxations. It is interesting to know how sharp these relaxations of general quadratic programs are. For the SDP relaxation this has been answered by [105]. In this section we give the corresponding result for the copositive programming relaxations and more generally for set-semidefinite programming, i.e., a cone program over the cone $\mathcal{C}_m^*(K)$ (see (2.2)).

The results obtained are somewhat negative. They roughly speaking say that

[1]This chapter is based on [2]

without adding extra restrictions into the relaxation we cannot expect the copositive programming or set-semidefinite programming relaxation of (nonconvex) quadratic programs to be sharp. To obtain sharper relaxations one has to consider additional restrictions, e.g., by adding new (convex quadratic) constraints which are redundant in the original quadratic program. Recent research has revealed that for several special classes of 0-1 programs such a sharpening leads to exact copositive programming representations (see e.g., [39, 47, 123, 124, 125]). The results in [39] have been extended to set-semidefinite programs ($K$-SD) by [40]. In this chapter $K \subseteq \mathbb{R}^m$ is a given cone (see (2.2),(2.1)). Future research should show which other classes of (non-convex) quadratic programs allow similar sharp copositive programming (or set - semidefinite programming) relaxations. In [10], a set of extra conditions on the original quadratic constraints is presented which guaranty that the $K$-SD relaxation is exact.

Note that exact set-semidefinite programming relaxations of NP-hard problems evidently are NP-hard. However the set-semidefinite programming relaxations are convex problems and one may hope that this extra structure leads to new insight and better algorithms for solving hard (non-convex) problems.

## 4.2 Set-Semidefinite Relaxation

We consider $(QCQP)$ given in Section 1.5. Although some arguments given below have already been mentioned in Section 1.5 we shall repeat them for the sake of completeness. Consider the (nonconvex) quadratic program with linear objective function and quadratic constraints:

$$(QP_0) \qquad \min\ \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \begin{array}{l} q_j(\mathbf{u}) \leq 0, \quad j \in J \\ \text{with also:} \quad \mathbf{u} \in K \ \text{ in K-SD case} \end{array}$$

with quadratic functions $q_j(\mathbf{u}) = \gamma_j + 2\mathbf{c}_j^T\mathbf{u} + \mathbf{u}^T C_j\mathbf{u},\ \ C_j \in \mathcal{S}_m,\ j \in J$, and $J$, a finite index set. We can write $q_j(\mathbf{u}) = \gamma_j + 2\mathbf{c}_j^T\mathbf{u} + \mathbf{u}^T C_j\mathbf{u}$ in the form

$$q_j(\mathbf{u}) = \left\langle Q_j, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{pmatrix} \right\rangle \quad \text{where } Q_j = \begin{pmatrix} \gamma_j & \mathbf{c}_j^T \\ \mathbf{c}_j & C_j \end{pmatrix} .$$

Recall that the relation $U = \mathbf{u}\mathbf{u}^T$ is equivalent to

$$\begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix}^T .$$

In this setting the original program $(QP_0)$ takes the equivalent lifted form:

$$(QP) \quad \min \quad \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \begin{array}{l} \left\langle Q_j, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \right\rangle \le 0,\ j \in J, \\ \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix}^T \\ \text{with also: } \mathbf{u} \in K \quad \text{in K-SD case} \end{array}$$

By replacing the (nonconvex) relation $\left(\begin{smallmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ \mathbf{u} \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ \mathbf{u} \end{smallmatrix}\right)^T$ by the SDP relaxation, $\left(\begin{smallmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{smallmatrix}\right) \in \mathcal{S}^+_{m+1}$, or the K-SD relaxation, $\left(\begin{smallmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{smallmatrix}\right) \in \mathcal{C}^*_{m+1}(\mathbb{R}_+ \times K)$ (see (2.2)), we are led to the relaxations of $(QP)$:

$$(SDP) \qquad \min\ \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \begin{array}{l} \left\langle Q_j, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \right\rangle \le 0,\ j \in J \\ \text{and} \quad \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \in \mathcal{S}^+_{m+1} \end{array}$$

$$(K\text{-}SD) \qquad \min\ \mathbf{c}_0^T\mathbf{u} \quad s.t. \quad \begin{array}{l} \left\langle Q_j, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \right\rangle \le 0,\ j \in J \\ \text{and} \quad \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \in \mathcal{C}^*_{m+1}(\mathbb{R}_+ \times K) \end{array}$$

In case of a K-SD relaxation of QP we always tacitly assume that the original program $(QP_0)$ and thus $(QP)$ contains the constraint $\mathbf{u} \in K$ (explicitly or implicitly). For optimality conditions and more details on K-SD programs and their dual we refer to [60]. We introduce some notation. Let $S$ denote the set of quadratic functions defining the feasible set of $(QP_0)$ and $(QP)$:

$$S = \{Q_j : j \in J\} \equiv \{q_j(\mathbf{u}) : j \in J\} .$$

Note that a quadratic function $q(\mathbf{u}) = \gamma + 2\mathbf{c}^T\mathbf{u} + \mathbf{u}^TC\mathbf{u}$ can be identified with the coefficient matrix $Q = \left(\begin{smallmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{smallmatrix}\right)$. In this chapter, $\mathcal{F}^{\text{QP}_0}, \mathcal{F}^{\text{QP}} = \mathcal{F}^{\text{QP}}(S)$, $\mathcal{F}^{\text{SDP}}(S)$ and $\mathcal{F}^{\text{K-SD}}(S)$ denote the feasible sets of $(QP_0)$, $(QP)$, the $(SDP)$ and the $(K\text{-}SD)$ relaxation, respectively. By $\mathcal{F}^{\text{QP}}_{\mathbf{u}}(S)$, $\mathcal{F}^{\text{SDP}}_{\mathbf{u}}(S)$ and $\mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(S)$ we denote the projections onto the **u**-space $\mathbb{R}^m$. Notice that all these feasible sets defined by a set $S$ of quadratic inequalities coincide with the feasible sets given by the conic combinations cone $(S)$, i.e., $\mathcal{F}^{\text{QP}}_{\mathbf{u}}(S) = \mathcal{F}^{\text{QP}}_{\mathbf{u}}(\text{cone}\,(S))$ etc. From these definitions we find

$$\mathcal{F}^{QP_0} = \mathcal{F}^{QP}_{\mathbf{u}}(S) = \mathcal{F}^{QP}_{\mathbf{u}}(\text{cone}\,(S)) \subset \text{ conv}\,\mathcal{F}^{QP}_{\mathbf{u}}(S)\ .$$

Since the objective of $(QP)$ is linear, the minimum value on $\mathcal{F}^{QP}_{\mathbf{u}}(S)$ and on conv $\mathcal{F}^{QP}_{\mathbf{u}}(S)$ coincide. By relaxation properties we have:

$$\text{conv}\,\mathcal{F}^{QP}_{\mathbf{u}}(S) \subset \mathcal{F}^{\text{SDP}}_{\mathbf{u}}(S)\ , \quad \text{conv}\,\mathcal{F}^{QP}_{\mathbf{u}}(S) \subset \mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(S)$$

and also $\mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(S) \subset \mathcal{F}^{\text{SDP}}_{\mathbf{u}}(S)$ in case $(QP)$ contains the constraint $\mathbf{u} \in K$.

We wish to know how sharp these inclusions are. Defining the set of convex quadratic functions,

$$\mathcal{Q}_+ := \left\{ Q = \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} : C \in \mathcal{S}^+_m \right\}$$

for the $(SDP)$ relaxation this question has been answered by Kojima and Tunçel in [105].

**Theorem 4.1.** *[105] conv* $[\mathcal{F}^{QP}_{\mathbf{u}}(S)] \subset \mathcal{F}^{QP}_{\mathbf{u}}(\text{cone}\,(S) \cap \mathcal{Q}_+) = \mathcal{F}^{SDP}_{\mathbf{u}}(S)$ .

We emphasize that in general the set $\mathcal{F}^{QP}_{\mathbf{u}}(\text{cone}\,(S) \cap \mathcal{Q}_+)$ is strictly smaller than the set $\mathcal{F}^{QP}_{\mathbf{u}}(S \cap \mathcal{Q}_+)$.

**Remark 4.2.** *In [105], based on the theorem above a conceptual algorithm is discussed which generates a sequence of sets $\mathcal{F}^{SDP}_{\mathbf{u}}(S_k)$ which converges to the set conv $[\mathcal{F}^{QP}_{\mathbf{u}}(S)]$. In each step by solving an SDP a ``cutting'' convex, quadratic constraint $\left\langle Q^K, \left(\begin{smallmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{smallmatrix}\right)\right\rangle = \gamma^k + 2(\mathbf{c}^k)^T\mathbf{u} + \mathbf{u}^T C^k \mathbf{u} \le 0$ with $Q^k \in \mathcal{Q}_+$ is constructed in such a way that for $S_{k+1} := S_k \cup \{Q^k\}$ we still have conv $[\mathcal{F}^{QP}_{\mathbf{u}}(S)] \subset \mathcal{F}^{SDP}_{\mathbf{u}}(S_{k+1})$ but the set $\mathcal{F}^{SDP}_{\mathbf{u}}(S_{k+1})$ is strictly smaller than $\mathcal{F}^{SDP}_{\mathbf{u}}(S_k)$. In the context of our generalization such a procedure is no more useful. For example in the case of $K = \mathbb{R}^m_+$, instead of a SDP, in each step we would have to solve a (NP-hard) ``completely positive program''.*

We now are able to extend (partially) the result of Theorem 4.1 to the K-SD relaxation of QP. The set $\mathcal{Q}_+$ in the SDP relaxation has now to be replaced by the set of ``$K$-semidefinite quadratic functions'':

$$\mathcal{Q}_{K-\mathbf{SD}} := \left\{ Q = \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} : C \in \mathcal{C}_m(K) \right\}$$

Let us first present an instructive example. Let $\mathcal{F}_{\mathbf{u}}^{\text{QP}}(\{Q\})$ be the feasible set defined by only one inequality

$$q(\mathbf{u}) = \left\langle Q, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{pmatrix} \right\rangle \leq 0,$$

$Q = \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix}$ ( and $\mathbf{u} \in K$) then:

if $C \notin \mathcal{C}_m(K)$ (i.e., $q$ is not ``K-semidefinite'') $\quad \Rightarrow \quad \mathcal{F}_{\mathbf{u}}^{\text{K-SD}}(\{Q\}) = K$ .

To see this, note that for $C \notin \mathcal{C}_m(K)$ there exists a vector $\mathbf{d} \in K$ such that $\mathbf{d}^T C \mathbf{d} < 0$. So, for any fixed $\mathbf{u} \in K$ with $U := \lambda \mathbf{d}\mathbf{d}^T + \mathbf{u}\mathbf{u}^T$ it holds

$$\left\langle Q, \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \right\rangle = \gamma + 2\mathbf{c}^T\mathbf{u} + \lambda \mathbf{d}^T C \mathbf{d} + \mathbf{u}^T C \mathbf{u} < 0 \quad \text{for } 0 < \lambda, \;\; \lambda \text{ large}$$

Since $U - \mathbf{u}\mathbf{u}^T = \lambda \mathbf{d}\mathbf{d}^T \in \mathcal{C}_m^*(K)$, Lemma 2.4 implies $\mathbf{u} \in \mathcal{F}_{\mathbf{u}}^{\text{K-SD}}(\{Q\})$. So, the K-SD relaxation does not provide any restriction apart from $\mathbf{u} \in K$. Generally, the following holds.

**Theorem 4.3.** *conv* $[\mathcal{F}_{\mathbf{u}}^{QP}(S)] \subset$ *conv* $[\mathcal{F}_{\mathbf{u}}^{QP}(\text{cone}\,(S) \cap \mathcal{Q}_{\textit{K-SD}})] \subset \mathcal{F}_{\mathbf{u}}^{\textit{K-SD}}(S)$ .

*Proof.* The first inclusion holds trivially. To prove the second, we begin by showing

$$\mathcal{Q}_{\text{K-SD}}^* = \left\{ \begin{pmatrix} 0 & \mathbf{o}^T \\ \mathbf{o} & B \end{pmatrix} : B \in \mathcal{C}_m^*(K) \right\} . \tag{4.1}$$

In fact, $\begin{pmatrix} \beta & \mathbf{b}^T \\ \mathbf{b} & B \end{pmatrix} \in \mathcal{Q}_{\text{K-SD}}^*$ holds if and only if for all $\begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} \in \mathcal{Q}_{\text{K-SD}}$, i.e., for all $\gamma \in \mathbb{R}, \mathbf{c} \in \mathbb{R}^m, C \in \mathcal{C}_m(K)$ we have

$$\left\langle \begin{pmatrix} \beta & \mathbf{b}^T \\ \mathbf{b} & B \end{pmatrix}, \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} \right\rangle = \beta\gamma + 2\mathbf{c}^T b + \langle C, B \rangle \geq 0 .$$

This obviously implies $\beta = 0, \mathbf{b} = \mathbf{o}$ and $B \in \mathcal{C}_m^*(K)$. On the other hand for any

$\begin{pmatrix} 0 & \mathbf{o}^T \\ \mathbf{o} & B \end{pmatrix}$, $B \in \mathcal{C}_m^*(K)$ it holds,

$$\left\langle \begin{pmatrix} 0 & \mathbf{o}^T \\ \mathbf{o} & B \end{pmatrix}, \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} \right\rangle = \langle B, C \rangle \geq 0$$

since $C \in \mathcal{C}_m(K)$. To compare the feasible sets we can write

$$\mathcal{F}_{\mathbf{u}}^{\text{K-SD}}(S) = \left\{ \mathbf{u} : \exists U \text{ such that } \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \in -S^* \cap \mathcal{C}_{m+1}^*(\mathbb{R}_+ \times K) \right\}$$

and by using the relations $(\text{cone}(S))^* = S^*$, $(K_1 \cap K_2)^* = K_1^* + K_2^*$ (for closed convex cones) and (4.1) we obtain

$$\begin{aligned} \mathcal{F}_{\mathbf{u}}^{\text{QP}}(\text{cone}\,(S) \cap \mathcal{Q}_{\text{K-SD}}) &= \left\{ \mathbf{u} : \left\langle \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{pmatrix}, Q \right\rangle \leq 0 \;\; \forall Q \in \text{cone}\,(S) \cap \mathcal{Q}_{\text{K-SD}} \right\} \\ &= \left\{ \mathbf{u} : \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{pmatrix} \in \; -(\text{cone}\,(S) \cap \mathcal{Q}_{\text{K-SD}})^* \right\} \\ &= \left\{ \mathbf{u} : \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{pmatrix} \in \; -(S^* + \mathcal{Q}_{\text{K-SD}}^*) \right\} \\ &= \left\{ \mathbf{u} : \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{pmatrix} \in \; -S^* - \begin{pmatrix} 0 & \mathbf{o}^T \\ \mathbf{o} & \mathcal{C}_m^*(K) \end{pmatrix} \right\} \end{aligned}$$

Consequently, $\mathbf{u} \in \mathcal{F}_{\mathbf{u}}^{\text{QP}}(\text{cone}\,(S) \cap \mathcal{Q}_{\text{K-SD}})$ holds if and only if with some $H \in \mathcal{C}_m^*(K)$ we have $\begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{o}^T \\ \mathbf{o} & H \end{pmatrix} \in -S^*$. But since $\mathbf{u}\mathbf{u}^T + H - \mathbf{u}\mathbf{u}^T \in \mathcal{C}_m^*(K), \mathbf{u} \in K$, by Lemma 2.4 it follows

$$\begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & H + \mathbf{u}\mathbf{u}^T \end{pmatrix} \in -S^* \cap \mathcal{C}_{m+1}^*(\mathbb{R}_+ \times K).$$

So (with $U = H + \mathbf{u}\mathbf{u}^T$), the vector $\mathbf{u}$ is contained in the set $\mathcal{F}_{\mathbf{u}}^{\text{K-SD}}(S)$. Since this set is convex the second inclusion follows. $\square$

To see the difference with the SDP case (in Theorem 4.1) let us chose $\mathbf{u} \in \mathcal{F}_{\mathbf{u}}^{\text{K-SD}}(S)$, i.e., with some $U \in \mathcal{S}_m$ the relation

$$\left\langle \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix}, Q \right\rangle \leq 0 \quad \text{for all } Q = \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} \in S$$

must hold. Then we also obtain

$$\left\langle \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix}, Q \right\rangle = \left\langle \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix}, Q \right\rangle + \left\langle \begin{pmatrix} 0 & \mathbf{o}^T \\ \mathbf{o} & \mathbf{u}\mathbf{u}^T - U \end{pmatrix}, Q \right\rangle$$
$$\leq \langle C, (\mathbf{u}\mathbf{u}^T - U) \rangle$$

Unfortunately the converse of Lemma 2.4 is not generally true. So here, even with $Q \in \text{cone}(S) \cap \mathcal{Q}_{\text{K-SD}}$, i.e., with $C \in \mathcal{C}_m(K)$, the relation $\langle C, \mathbf{u}\mathbf{u}^T - U \rangle \leq 0$ need not hold and $\mathbf{u}$ need not satisfy the corresponding original constraint $\langle C, \mathbf{u}\mathbf{u}^T \rangle + 2\mathbf{c}^T\mathbf{u} + \gamma \leq 0$. We give some examples to illustrate the statement of Theorem 4.3 and to show that in general (for $K \neq \mathbb{R}^m$) the situation is more complicated than in the SDP case (for $K = \mathbb{R}^m$).

**Example 4.4.** *We chose $K = \mathbb{R}^m_+$, i.e., the completely positive relaxation. Let us take the special case $S \subset \mathcal{Q}_{\text{K-SD}}$. In contrast to the SDP relaxation the set $\mathcal{F}^{QP}_{\mathbf{u}}(S)$ need not be convex. So, an inclusion $\mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(S) \subset \mathcal{F}^{QP}_{\mathbf{u}}(S)$ is not true in general. Even $\mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(S) \subset$ conv $[\mathcal{F}^{QP}_{\mathbf{u}}(S)]$ need not hold as we shall show. Theorem 4.3 only assures conv $[\mathcal{F}^{QP}_{\mathbf{u}}(S)] \subset \mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(S)$. Even in the case $S = \{Q\}$ with $Q = \begin{pmatrix} \gamma & \mathbf{c}^T \\ \mathbf{c} & C \end{pmatrix} \in \mathcal{Q}_{\text{K-SD}}$ the inclusion can be strict. Take for example*

$$C = \begin{pmatrix} \frac{1}{2} & 1 \\ 1 & \frac{1}{2} \end{pmatrix} \quad , \quad \mathbf{c} = (-2.5\ ,\ -2 + \rho)\ , \quad \gamma = 8.$$

*The feasibility conditions read:*

$$\text{for } \mathcal{F}^{QP}_{\mathbf{u}}(\{Q\}) : \frac{1}{2}(u_1^2 + u_2^2) + 2u_1u_2 - 5u_1 - (2-\rho)u_2 + 8 \leq 0, \quad \text{and } \mathbf{u} \in \mathbb{R}^2_+$$
$$\text{for } \mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(\{Q\}) : \frac{1}{2}(U_{11} + U_{22}) + 2U_{12} - 5u_1 - (2-\rho)u_2 + 8 \leq 0,$$
$$\text{and } \begin{pmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & U \end{pmatrix} \in \mathcal{C}^*_3(\mathbb{R}^m_+)$$

*We have computed the feasible sets. For $\rho = 0$ the set $\mathcal{F}^{QP}_{\mathbf{u}}(\{Q\})$ consists of the point $(0, 4)$ together with the convex (black) set (see Figure 4.1). The set $\mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(\{Q\})$ equals the (grey) triangle conv$[\mathcal{F}^{QP}_{\mathbf{u}}(\{Q\})]$. For $\rho > 0$ (small) the point $(0, 4)$ is no more feasible for $\mathcal{F}^{QP}_{\mathbf{u}}(\{Q\})$ (black) and the (convex) set $\mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(\{Q\})$ (grey) (depending continuously on $\rho$) is as sketched in Figure 4.2 (for $\rho = 0.2$). Obviously in this example $\rho = 0.2$ we have*

$$\mathcal{F}^{QP}_{\mathbf{u}}(\{Q\}) = \text{conv}\,[\mathcal{F}^{QP}_{\mathbf{u}}(\{Q\})] \subsetneq \mathcal{F}^{\text{K-SD}}_{\mathbf{u}}(\{Q\}).$$

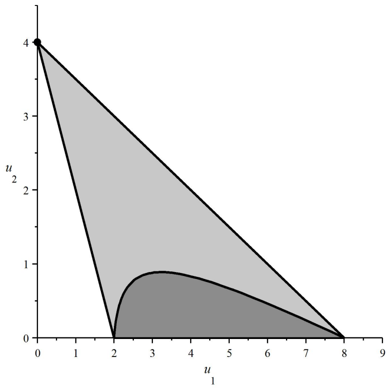



Figure 4.1: $\mathcal{F}_{\mathbf{u}}^{\mathrm{QP}}(\{Q\})$ for $\rho = 0$

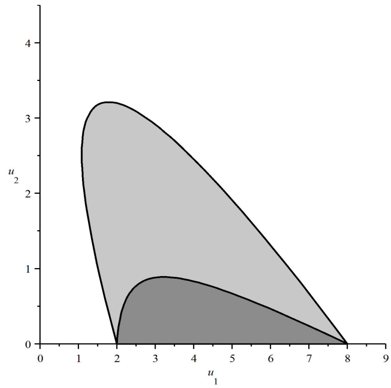



Figure 4.2: $\mathcal{F}_{\mathbf{u}}^{\mathrm{QP}}(\{Q\})$ for $\rho = 0.2$

*For the other special case $S \cap \mathcal{Q}_{\textit{K-SD}} = \emptyset$ we have:*

$$\textit{conv}\,[\mathcal{F}_{\mathbf{u}}^{\textit{QP}}(\mathrm{cone}(S) \cap \mathcal{Q}_{\textit{K-SD}})] \subset \mathcal{F}_{\mathbf{u}}^{\textit{K-SD}}(S) \subset \textit{conv}\,[\mathcal{F}_{\mathbf{u}}^{\textit{QP}}(S \cap \mathcal{Q}_{\textit{K-SD}})] = \mathbb{R}_{+}^{m}\ .$$

*The equality on the right-hand side follows by the assumption $S \cap \mathcal{Q}_{\textit{K-SD}} = \emptyset$, so that the feasibility condition for $\mathcal{F}_{\mathbf{u}}^{\textit{QP}}(S \cap \mathcal{Q}_{\textit{K-SD}})$ reduces to $\mathbf{u} \in K = \mathbb{R}_{+}^{m}$.*

# 5

# Copositive Programming via Semi-infinite Optimization[1]

A standard way to tackle new problems in mathematics is to formulate them in a well known form and utilize the machinery available to solve the problem. In this chapter copositive programming (COP) is viewed as the special case of linear semi-infinite programming. We start in the first section by formulating a copositive program as a linear semi-infinite program(LSIP). In section two, first order optimality conditions and duality results of LSIP are applied to COP leading to known results but also to new insight. In section three, we reinterpret approximation schemes for solving COP as discretization methods in LSIP. This leads to new explicit error bounds between the approximate and the original problem. Section five gives error bounds for the maximizers in dependence on the order of the maximizer of the original program. We also show by examples that maximizers of arbitrarily large order can occur in copositive programming.

[1]This chapter is based on [1]

## 5.1 LSIP Representation of COP

In this section we shall reformulate COP as LSIP. First recall the pair of primal/ dual copositive programs (COP) from Chapter 1,

$$(COP_P) \qquad \max_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \qquad s.t. \quad B - \sum_{i=1}^n x_i A_i \in \mathcal{C}_m$$

$$(COP_D) \qquad \min_{Y\in\mathcal{S}_m} \langle Y, B\rangle \qquad s.t. \quad \langle Y, A_i\rangle = c_i \ (i = 1,\dots,n), \ \ Y \in \mathcal{C}_m^*,$$

We assume throughout that the matrices $A_i$ ($i = 1,\dots,n$) are linearly independent. Recall also our standard form of linear semi-infinite primal/dual programs,

$$(SIP_P) \qquad \max_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \qquad s.t. \quad b(\mathbf{z}) - a(\mathbf{z})^T\mathbf{x} \geq 0 \quad \forall\, \mathbf{z} \in Z,$$

$$(SIP_D) \qquad \min_{y_z} \sum_{\mathbf{z}\in Z} y_z b(\mathbf{z}) \qquad s.t. \quad \sum_{\mathbf{z}\in Z} y_z a(\mathbf{z}) = \mathbf{c}, \ y_z \geq 0,$$

Note that the condition $A \in \mathcal{C}_m$ can be equivalently expressed by either of the conditions:

$$\mathbf{z}^T A\mathbf{z} \geq 0 \quad \forall\, \mathbf{z} \in \mathcal{B}_m := \left\{\mathbf{z} \in \mathbb{R}^m_+ : \|\mathbf{z}\| = 1\right\} \quad \text{(unit orthant)},$$
$$\mathbf{z}^T A\mathbf{z} \geq 0 \quad \forall\, \mathbf{z} \in \Delta_m := \left\{\mathbf{z} \in \mathbb{R}^m_+ : \sum_{i=1}^m \mathbf{z}_i = 1\right\} \quad \text{(unit simplex)}.$$

In view of this relation, the primal COP can be written as a $(SIP_P)$ with

$$a(\mathbf{z}) = (\mathbf{z}^T A_1\mathbf{z}, \dots, \mathbf{z}^T A_n\mathbf{z})^T, \quad b(\mathbf{z}) = \mathbf{z}^T B\mathbf{z}, \quad \text{and} \quad Z \in \{\mathcal{B}_m, \Delta_m\}. \tag{5.1}$$

In this chapter, we always take $Z = \Delta_m$, and defining

$$F(\mathbf{x}) := B - \sum_{i=1}^n x_i A_i$$

we write the copositive primal problem $(COP_P)$ in the form:

$$(COP_P) \qquad \max_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \qquad s.t. \quad \mathbf{z}^T F(\mathbf{x})\mathbf{z} \geq 0 \ \ \forall \mathbf{z} \in Z := \Delta_m. \tag{5.2}$$

In view of (5.1), the feasibility condition for $(SIP_D)$ becomes

$$c_i = \sum_{\mathbf{z}\in Z} y_z \langle \mathbf{z}\mathbf{z}^T, A_i\rangle \quad (i = 1, \ldots, n) \ , \ y_z \geq 0$$

and with $Y := \sum_{\mathbf{z}\in Z} y_z \ \mathbf{z}\mathbf{z}^T \in \mathcal{C}_m^*$ this coincides with the feasibility condition $c_i = \langle Y, A_i\rangle \quad (i = 1, \ldots, n)$ of $(COP_D)$. Moreover,

$$\sum_{\mathbf{z}\in Z} y_z b(\mathbf{z}) = \sum_{\mathbf{z}\in Z} y_z \langle \mathbf{z}\mathbf{z}^T, B\rangle = \langle Y, B\rangle.$$

So, the dual $(SIP_D)$ of $(COP_P)$ in LSIP form (5.2) is equivalent to the COP dual $(COP_D)$ and we simply denote both versions by $(COP_D)$.

We shall close this section with an observation on the number of isolated active indices a copositive program can have. Consider the following program,

$$(COP_Q) \qquad \max_{x\in\mathbb{R}} \ x \quad s.t. \qquad -Q - xE \in \mathcal{C}_m$$

where $Q \in \mathcal{S}_m$ with $m = 3n$ for $n \geq 2$ as considered in Example 3.47. The LSIP formulation of the above program is,

$$\max_{x\in\mathbb{R}} \ x \quad s.t. \qquad \mathbf{b}(\mathbf{z}) - \mathbf{a}(\mathbf{z})x \geq 0 \ \forall \ \mathbf{z} \in \Delta_m$$

where $\mathbf{b}(\mathbf{z}) = -\mathbf{z}^T Q\mathbf{z}$, $\mathbf{a}(\mathbf{z}) = \mathbf{z}^T E\mathbf{z} = 1$. Note that in $x \leq -\mathbf{z}^T Q\mathbf{z}$, for all $\mathbf{z} \in \Delta_m$ equality holds if and only if $x = -\max_{\mathbf{z}\in\Delta_m} \mathbf{z}^T Q\mathbf{z}$. Recall from Example 3.47 that $\max_{\mathbf{z}\in\Delta_m} \mathbf{z}^T Q\mathbf{z} = \frac{n-1}{n}$. Moreover, there are $3^{\frac{m}{3}}$ strict local maximizers of $\max_{\mathbf{z}\in\Delta_m} \mathbf{z}^T Q\mathbf{z}$ with value $\frac{n-1}{n}$.

In view of Definition 1.14 it is clear that the set of active indices of the solution $\overline{x} = -\frac{n-1}{n}$ of the above program reads:

$$Z(\overline{x}) = \{\overline{\mathbf{z}} \in \Delta_m : \overline{x} = -\overline{\mathbf{z}}^T Q\overline{\mathbf{z}}\} \tag{5.3}$$

From (5.3) it is clear that the isolated active indices of the copositive program $(COP_Q)$ are precisely the strict local maximizers of $\mathbf{z}^T Q\mathbf{z}$ over $\Delta_m$, which are $3^{\frac{m}{3}}$ in total as mentioned above implying that $|Z(\overline{x})| = 3^{\frac{m}{3}}$. Hence, the copositive program can have an exponential number of active indices. This fact also indicates that solving $(COP_P)$ is ``hard''.

## 5.2 Optimality Conditions and Duality

From the LSIP form of COP, clearly, any result for LSIP can directly be translated to COP. We will do this for some optimality conditions and duality results.

As mentioned before, optimality conditions for LSIP are usually presented in terms of KKT conditions for a feasible candidate maximizer $\overline{\mathbf{x}}$.

Using (5.1), for the copositive problem in LSIP-form (5.2), the KKT conditions for LSIP (1.5) translate to

$$\mathbf{c} = \sum_{j=1}^{k} y_j a(\mathbf{z}_j) = \sum_{j=1}^{k} y_j \begin{pmatrix} \mathbf{z}_j^T A_1 \mathbf{z}_j \\ \vdots \\ \mathbf{z}_j^T A_n \mathbf{z}_j \end{pmatrix}, \quad \mathbf{z}_j \in Z(\overline{\mathbf{x}}), y_j \geq 0 \ (j = 1, \ldots, k). \tag{5.4}$$

It is important to note that any solution of the KKT system with feasible $\overline{\mathbf{x}}$, automatically yields a minimizer $\overline{Y}$ of the dual program $(COP_D)$:

$$\overline{Y} := \sum_{j=1}^{k} y_j \mathbf{z}_j \mathbf{z}_j^T \in \mathcal{C}_m^*. \tag{5.5}$$

Observe that, by Carathéodory's Lemma for cones (see, e.g., [63]), we can assume that

$$\text{the KKT condition (5.4) is satisfied with } k \leq n \text{ active points } \mathbf{z}_j \in Z(\overline{\mathbf{x}}). \tag{5.6}$$

This implies that the dual minimizer $\overline{Y}$ allows a representation (5.5) with $k \leq n$, i.e., $\overline{Y} \in \mathcal{C}_m^*$ has CP-rank $\leq n$.

Before applying the standard results of LSIP to copositive programming, we have to translate the primal/dual constraint qualification (Slater condition) from LSIP (see Definition 1.12) to the copositive terminology.

**Lemma 5.1.** *Consider the copositive problem in its LSIP-formulation* (5.2)*. The primal LSIP constraint qualification*

$$(CQ_P): \qquad \mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} \geq \sigma_0 > 0, \ \textit{for all } \mathbf{z} \in Z \textit{ and for some } \sigma_0 > 0 \tag{5.7}$$

*is satisfied for $\mathbf{x}_0 \in \mathbb{R}^n$ if and only if $F(\mathbf{x}_0) \in \text{int}(\mathcal{C}_m)$. The dual LSIP constraint qualification*

$$(CQ_D): \qquad \mathbf{c} \in \text{int}(M), \quad \textit{with } M := \text{cone}\{a(\mathbf{z}) : \mathbf{z} \in Z\}$$

*holds if and only if there exists $Y_0$ feasible for $(COP_D)$ such that $Y_0 \in \text{int}(\mathcal{C}_m^*)$.*

*Proof.* The fact that (5.7) implies,

$$F(\mathbf{x}_0) \in \text{int}(\mathcal{C}_m),$$

follows immediately from (2.6).

For the converse let $F(\mathbf{x}_0) \in \text{int}(\mathcal{C}_m)$. Then there exists $\varepsilon > 0$ such that $F \in \mathcal{C}_m$ for all $F$ with $\|F - F(\mathbf{x}_0)\| \leq \varepsilon$. Define $F := F(\mathbf{x}_0) - \frac{\varepsilon}{\sqrt{m}} I$. Then $\|F - F(\mathbf{x}_0)\| \leq \varepsilon$ and thus $F \in \mathcal{C}_m$. Consequently,

$$0 \leq \mathbf{z}^T F \mathbf{z} = \mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} - \frac{\varepsilon}{\sqrt{m}} \mathbf{z}^T \mathbf{z} \quad \text{for all } \mathbf{z} \in Z.$$

Using $\mathbf{z}^T \mathbf{z} \geq \frac{1}{m}$ for $\mathbf{z} \in Z$, we obtain $\mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} \geq \frac{\varepsilon}{\sqrt{m}m} =: \sigma_0 > 0$ for all $\mathbf{z} \in Z$. To prove the equivalence of the dual constraint qualifications we define the mapping $c(Y) := (\langle A_1, Y\rangle, \ldots, \langle A_n, Y\rangle)^T$. We first show that

$$Y \in \mathcal{C}_m^* \quad \Rightarrow \quad c(Y) \in M. \tag{5.8}$$

To see this, note that $Y \in \mathcal{C}_m^*$ has a rank-one representation $Y = \sum_{j=1}^k \mathbf{v}_j \mathbf{v}_j^T$ with $\mathbf{o} \neq \mathbf{v}_j \in \mathbb{R}_+^m$ for all $j = 1, \ldots, k$. Define $\mathbf{z}_j := \mathbf{v}_j / (\mathbf{v}_j^T \mathbf{e})$ to obtain $\mathbf{z}_j \in Z$, and $y_j := (\mathbf{v}_j^T \mathbf{e})^2 > 0$. Then $Y = \sum_{j=1}^k y_j \mathbf{z}_j \mathbf{z}_j^T$. Therefore we get

$$c(Y) = \sum_{j=1}^k y_j \big(\langle A_1, \mathbf{z}_j \mathbf{z}_j^T\rangle, \ldots, \langle A_n, \mathbf{z}_j \mathbf{z}_j^T\rangle\big)^T = \sum_{j=1}^k y_j a(\mathbf{z}_j) \in M,$$

and (5.8) is proved. Now let $Y_0 \in \text{int}(\mathcal{C}_m^*)$ be feasible for $(COP_D)$, i.e., $c(Y_0) = \mathbf{c}$. To prove $\mathbf{c} \in \text{int}(M)$ we assert that there exists some $\varepsilon > 0$ such that, for any $\gamma \in \mathbb{R},\ |\gamma| < \varepsilon$, the relation

$$\mathbf{c} + \gamma e_k \in M \text{ holds for all (standard basis) vectors } e_k \ (k = 1, \ldots, n). \tag{5.9}$$

To show this we note that, since the $A_i$'s are linearly independent, for any $k$ the linear system $c(Y_k) = (\langle A_1, Y_k\rangle, \ldots, \langle A_n, Y_k\rangle)^T = e_k$ has a solution $\widetilde{Y}_k \in \mathcal{S}_m$. Since $Y_0 \in \text{int}(\mathcal{C}_m^*)$, there exists some $\varepsilon > 0$ such that for all $\gamma,\ |\gamma| < \varepsilon$:

$$Y_k := Y_0 + \gamma \widetilde{Y}_k \in \mathcal{C}_m^* \quad \text{for all } k = 1, \ldots, n.$$

Using (5.8) and $c(Y_0) = \mathbf{c}$ we get $M \ni c(Y_k) = c(Y_0) + \gamma c(\widetilde{Y}_k) = \mathbf{c} + \gamma e_k$, which proves (5.9).

We finally show that $(CQ_D)$ yields some $Y_0$ feasible for $(COP_D)$ with $Y_0 \in$

$\text{int}(\mathcal{C}_m^*)$. To do so, choose any $Y_* \in \text{int}(\mathcal{C}_m^*)$ and define:

$$\mathbf{b} := c(Y_*) = (\langle A_1, Y_* \rangle, \ldots, \langle A_n, Y_* \rangle)^T.$$

Since $\mathbf{c} \in \text{int}(M)$, we have for some $\varepsilon > 0$ that $\mathbf{c} - \varepsilon\mathbf{b} \in M$, which means that for some $y_j \geq 0, \mathbf{z}_j \in Z$ $(j = 1, \ldots, k)$ we have

$$\mathbf{c} - \varepsilon\mathbf{b} = \sum_{j=1}^{k} y_j a(\mathbf{z}_j) = \sum_{j=1}^{k} y_j c(\mathbf{z}_j \mathbf{z}_j^T).$$

Defining $Y := \sum_{j=1}^{k} y_j \mathbf{z}_j \mathbf{z}_j^T \in \mathcal{C}_m^*$, we find that $c(Y) = \mathbf{c} - \varepsilon\mathbf{b}$ by construction. Next, define $Y_0 := Y + \varepsilon Y_*$. Then $Y_0 \in \text{int}(\mathcal{C}_m^*)$ because $Y \in \mathcal{C}_m^*, Y_* \in \text{int}(\mathcal{C}_m^*)$ and $\mathcal{C}_m^*$ is a convex cone. Moreover, $c(Y_0) = c(Y) + \varepsilon c(Y_*) = \mathbf{c} - \varepsilon\mathbf{b} + \varepsilon\mathbf{b} = \mathbf{c}$, which means that $Y_0$ is feasible for $(COP_D)$. This completes the proof. □

We emphasize that relation (5.6) implies that, under $(CQ_P)$ to any maximizer $\overline{\mathbf{x}}$ of $(COP_P)$, there always exists a corresponding (complementary) optimal solution $\overline{Y}$ of $(COP_D)$ that has CP-rank $\leq n$. Similarly the duality result for LSIP, Theorem 1.13, can be applied to copositive programming.

## 5.3 Discretization Methods for COP

Due to the LSIP representation of COP, any solution method of LSIP can directly be applied to COP. In this chapter, we only consider discretization methods. An inner and outer approximation algorithm for COP has been proposed and analysed by Bundfuss and Dür [38]. We re-analyse this approach in the light of discretization methods in LSIP as outlined in [144]. This will lead to additional insight and explicit error bounds.

We start with the COP in LSIP-form (5.2) with $Z = \Delta_m$. The approach in [38] is based on the following partition of $\Delta_m$.

**Definition 5.2.** We partition the unit simplex $Z = \Delta_m$ into finitely many sub-simplices $\Delta^1, \ldots, \Delta^k$ of $\Delta_m$ such that

$$\Delta_m = \bigcup_{l=1}^{k} \Delta^l \quad \text{and} \quad \text{int}(\Delta^l) \cap \text{int}(\Delta^p) = \emptyset \;\text{ for }\; l \neq p.$$

This partition defines a meshsize $d$, a discretization $Z_d$ and a set $E_d$ of ``edges'' (pairs of vertices):

$$
\begin{aligned}
Z_d =& \{\mathbf{v}_j : \mathbf{v}_j \text{ is a vertex of } \Delta^l \text{ for some } l\} \\
E_d =& \{(\mathbf{v}_i, \mathbf{v}_j) : \mathbf{v}_i, \mathbf{v}_j \text{ are vertices in the same } \Delta^l \text{ for some } l \text{ (possibly } i = j)\} \\
d =& \max\{\|\mathbf{v}_i - \mathbf{v}_j\| : (\mathbf{v}_i, \mathbf{v}_j) \in E_d\}.
\end{aligned}
$$

In [38], the following outer and inner approximation schemes for (5.2) are given:

$$
\begin{aligned}
&(P_d) && \max_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \quad s.t. \quad \mathbf{z}^T F(\mathbf{x})\mathbf{z} \geq 0 \ \ \forall \mathbf{z} \in Z_d, \\
&(\widetilde{P}_d) && \max_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \quad s.t. \quad \mathbf{u}^T F(\mathbf{x})\mathbf{v} \geq 0 \ \ \forall (\mathbf{u}, \mathbf{v}) \in E_d.
\end{aligned}
$$

Note that $(P_d)$ represents a special instance of a discretization scheme in LSIP. ($\widetilde{P}_d$) provides feasible points for the original copositive problem $(COP_P)$, see [38] and below. Observe that both $(P_d)$ and ($\widetilde{P}_d$) are linear programming problems.

**Remark 5.3.** *Note that any point* $\mathbf{z} \in Z = \Delta_m$ *is contained in one of the sub-simplices* $\Delta^l$ *and thus* $\mathbf{z} \in \Delta^l$ *can be written as a convex combination* $\mathbf{z} = \sum_\nu \lambda_\nu \mathbf{v}_\nu$*, with* $\sum_\nu \lambda_\nu = 1, \lambda_\nu \geq 0$ *of vertices* $\mathbf{v}_\nu$ *of* $\Delta^l$*. Consequently, for any* $\mathbf{z} \in Z$*, the inequality* $\min_{\mathbf{z}_j \in Z_d} \|\mathbf{z} - \mathbf{z}_j\| \leq d$ *holds so that* $d$ *above really defines a meshsize:*

$$
d \geq \max_{\mathbf{z}\in Z} \min_{\mathbf{z}_j \in Z_d} \|\mathbf{z} - \mathbf{z}_j\|.
$$

In the following, the vector $\overline{\mathbf{x}}$ is always a maximizer of $(COP_P)$ and $\overline{\mathbf{x}}_d, \widetilde{\mathbf{x}}_d$ are feasible points (possibly maximizers) of $(P_d), (\widetilde{P}_d)$. We are now going to discuss some of the convergence results of [144] for our special program $(COP_P)$ in terms of the meshsize $d$ in an explicit form. The proofs are independent and mainly based on the following two relations: For any $F \in \mathcal{S}_m$ and $\mathbf{z}, \mathbf{u} \in \mathbb{R}^m$ we have

$$
\mathbf{z}^T F \mathbf{u} = \tfrac{1}{2}\left[\mathbf{z}^T F \mathbf{z} + \mathbf{u}^T F \mathbf{u} - (\mathbf{z} - \mathbf{u})^T F (\mathbf{z} - \mathbf{u})\right]. \tag{5.10}
$$

Moreover, as mentioned earlier, for every $\mathbf{z} \in \Delta^l \subseteq Z$ we have the representation $\mathbf{z} = \sum_\nu \lambda_\nu \mathbf{v}_\nu$ with $\mathbf{v}_\nu$ the vertices of $\Delta^l$, $\lambda_\nu \geq 0$, and $\sum_\nu \lambda_\nu = 1$. This gives:

$$
\mathbf{v}_\nu^T F \mathbf{v}_\mu \geq \gamma, \ \ \forall (\mathbf{v}_\nu, \mathbf{v}_\mu) \in E_d \quad \Rightarrow \quad \mathbf{z}^T F \mathbf{z} = \sum_{\nu,\mu} \lambda_\nu \lambda_\mu \mathbf{v}_\nu^T F \mathbf{v}_\mu \geq \gamma \ \ \forall \mathbf{z} \in \Delta_m. \tag{5.11}
$$

Clearly, $\mathcal{F}(COP_P) \subset \mathcal{F}(P_d)$ holds, and using (5.11) for $\gamma = 0$ we obtain the relations

$$\mathcal{F}(\widetilde{P}_d) \subset \mathcal{F}(COP_P) \subset \mathcal{F}(P_d) \quad \text{and thus} \quad val(\widetilde{P}_d) \leq val(COP_P) \leq val(P_d). \tag{5.12}$$

We are interested in accurate bounds, e.g., for $val(P_d) - val(COP_P)$ and $val(COP_P) - val(\widetilde{P}_d)$, depending explicitly on the meshsize $d$. From [144], we know that even for nonlinear LSIP under a constraint qualification, the approximation error between $\mathcal{F}(COP_P), val(COP_P)$ and $\mathcal{F}(P_d), val(P_d)$ behaves like $\mathcal{O}(d^2)$ in the meshsize $d$, provided that the discretization $Z_d$ of $Z$ *``covers all boundary parts of $Z$ of all dimensions''*. In the above discretization scheme this is automatically fulfilled.

The next lemma shows that the inner approximation $(\widetilde{P}_d)$ yields points feasible for the original program $(COP_P)$ and the outer approximation $(P_d)$ generates points with an infeasibility error of order $\mathcal{O}(d^2)$.

**Lemma 5.4.** *Let $\overline{\mathbf{x}}_d, \widetilde{\mathbf{x}}_d$ be feasible for $(P_d), (\widetilde{P}_d)$. Then for all $\mathbf{z} \in Z$ and for all $d$ we have:*

*(a)* $\mathbf{z}^T F(\overline{\mathbf{x}}_d)\mathbf{z} \geq -\frac{1}{2}\|F(\overline{\mathbf{x}}_d)\| \cdot d^2$

*(b)* $\mathbf{z}^T F(\widetilde{\mathbf{x}}_d)\mathbf{z} \geq 0.$

*So $\widetilde{\mathbf{x}}_d$ is feasible for $(COP_P)$, and $\overline{\mathbf{x}}_d$ is feasible up to an error of order $\mathcal{O}(d^2)$.*

*Proof.* Let $F = F(\overline{\mathbf{x}}_d)$. Using $\mathbf{z}^T F\mathbf{z} \geq 0$ for all $\mathbf{z} \in Z_d$, we find from (5.10) that for all $(\mathbf{z}, \mathbf{u}) \in E_d$

$$\begin{aligned} \mathbf{z}^T F\mathbf{u} &= \tfrac{1}{2}\left[\mathbf{z}^T F\mathbf{z} + \mathbf{u}^T F\mathbf{u} - (\mathbf{z}-\mathbf{u})^T F(\mathbf{z}-\mathbf{u})\right] \\ &\geq -\tfrac{1}{2}(\mathbf{z}-\mathbf{u})^T F(\mathbf{z}-\mathbf{u}) \geq -\tfrac{1}{2}\|F\|\|\mathbf{z}-\mathbf{u}\|^2 \\ &\geq -\tfrac{1}{2}\|F\| \cdot d^2. \end{aligned}$$

The second inequality follows from the fact that with the 2- norms the relation $\|Fz\| \leq \|F\|\|\mathbf{z}\|$ holds. In view of (5.11), this shows (a). Letting $F := F(\widetilde{\mathbf{x}}_d)$, (b) follows from (5.11) with $\gamma = 0$. □

Assuming a strictly feasible point $\mathbf{x}_0$ we show that small perturbations of any feasible point $\overline{\mathbf{x}}_d$ for $(P_d)$ leads to points in $\mathcal{F}(COP_P)$ or even $\mathcal{F}(\widetilde{P}_d)$.

**Lemma 5.5.** *Let $(CQ_P)$ be satisfied for $\mathbf{x}_0 \in \mathcal{F}(COP_P)$ with $\sigma_0 > 0$ (see* (5.7)*). Then for any $\overline{\mathbf{x}}_d$, feasible for $(P_d)$ and $d$ small enough we have:*

*(a)* $\overline{\mathbf{x}}_d^* := \overline{\mathbf{x}}_d + \rho d^2(\mathbf{x}_0 - \overline{\mathbf{x}}_d) \in \mathcal{F}(COP_P)$ *for* $\rho \geq \frac{\|F(\overline{\mathbf{x}}_d)\|}{2\sigma_0}$ *and* $0 < \rho d^2 < 1$

*(b)* $\widetilde{\mathbf{x}}_d^* := \overline{\mathbf{x}}_d + \tau d^2(\mathbf{x}_0 - \overline{\mathbf{x}}_d) \in \mathcal{F}(\widetilde{P}_d)$ *for* $\tau \geq \frac{\|F(\overline{\mathbf{x}}_d)\|}{2\sigma_0 + d^2(\|F(\overline{\mathbf{x}}_d)\| - \|F(\mathbf{x}_0)\|)}$ *and* $0 < \tau d^2 < 1$
*Recall that* $\mathcal{F}(\widetilde{P}_d) \subset \mathcal{F}(COP_P)$ *holds, cf.,* (5.12).

*(c) If* $\overline{\mathbf{x}}_d$ *is a solution of* $(P_d)$*, i.e.,* $\mathbf{c}^T\overline{\mathbf{x}}_d = val(P_d)$ *it follows*
$0 \leq val(P_d) - val(\widetilde{P}_d) \leq \tau[\mathbf{c}^T(\overline{\mathbf{x}}_d - \mathbf{x}_0)] \cdot d^2$ *for* $\tau$ *satisfying the bound in (b).*

*Proof.* Recall that $(CQ_P)$ means that $\mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} \geq \sigma_0 > 0$ for all $\mathbf{z} \in Z$. Using this, the fact that $F(\overline{\mathbf{x}}_d^*) = (1-\rho d^2)F(\overline{\mathbf{x}}_d) + \rho d^2 F(\mathbf{x}_0)$, and Lemma 5.4, we see that for any $\mathbf{z} \in Z$ and $0 \leq 1 - \rho d^2$,

$$\begin{aligned}
\mathbf{z}^T F(\overline{\mathbf{x}}_d^*)\mathbf{z} &= (1-\rho d^2)\mathbf{z}^T F(\overline{\mathbf{x}}_d)\mathbf{z} + \rho d^2 \mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} \\
&\geq -\tfrac{1}{2}(1-\rho d^2)\|F(\overline{\mathbf{x}}_d)\| \cdot d^2 + \rho d^2\sigma_0 \\
&\geq d^2(\rho\sigma_0 - \tfrac{1}{2}\|F(\overline{\mathbf{x}}_d)\|).
\end{aligned}$$

which shows (a). Part (b) is proven similarly. Here, observing

$$F(\widetilde{\mathbf{x}}_d^*) = (1-\tau d^2)F(\overline{\mathbf{x}}_d) + \tau d^2 F(\mathbf{x}_0),$$

for any pair $(\mathbf{z}, \mathbf{u}) \in E_d$, we find using (5.10), $\mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} \geq \sigma_0$ and $\mathbf{z}^T F(\overline{\mathbf{x}}_d)\mathbf{z} \geq 0$

$$\begin{aligned}
\mathbf{z}^T F(\widetilde{\mathbf{x}}_d^*)\mathbf{u} &= (1-\tau d^2)\frac{1}{2}[\mathbf{z}^T F(\overline{\mathbf{x}}_d)\mathbf{z} + \mathbf{u}^T F(\overline{\mathbf{x}}_d)\mathbf{u} - (\mathbf{z}-\mathbf{u})^T F(\overline{\mathbf{x}}_d)(\mathbf{z}-\mathbf{u})] + \\
&\quad + \tau d^2 \frac{1}{2}[\mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} + \mathbf{u}^T F(\mathbf{x}_0)\mathbf{u} - (\mathbf{z}-\mathbf{u})^T F(\mathbf{x}_0)(\mathbf{z}-\mathbf{u})] \\
&\geq -(1-\tau d^2)\frac{1}{2}\|F(\overline{\mathbf{x}}_d)\|\ d^2 + \tau d^2\left(\sigma_0 - d^2\frac{\|F(\mathbf{x}_0)\|}{2}\right) \\
&= d^2\left(-\frac{\|F(\overline{\mathbf{x}}_d)\|}{2} + \tau\big[\sigma_0 + \frac{d^2}{2}(\|F(\overline{\mathbf{x}}_d)\| - \|F(\mathbf{x}_0)\|)\big]\right) \quad \geq 0
\end{aligned}$$

if $\tau$ is chosen as stated (assuming $\|F(\mathbf{x}_0)\|d^2 \leq \sigma_0$, implying $\tau > 0$). The inequality (c) for the maximum values is deduced easily using that $\widetilde{\mathbf{x}}_d^*$ is feasible for ($\widetilde{P}_d$):

$$0 \leq val(P_d) - val(\widetilde{P}_d) \leq \mathbf{c}^T(\overline{\mathbf{x}}_d - \widetilde{\mathbf{x}}_d^*) = [\mathbf{c}^T(\overline{\mathbf{x}}_d - \mathbf{x}_0)\tau] \cdot d^2. \qquad \square$$

Observe that the bounds in Lemma 5.5(c) depend on the actual solutions $\overline{\mathbf{x}}_d$ of $(P_d)$. In order to use these bounds (a-priori) we must assure that the solutions $\overline{\mathbf{x}}_d$ exist and that they are bounded. As we shall see below, the key assumption here is a dual constraint qualification. We define the distance between a point $\mathbf{x}$ and the set $\mathcal{S}(COP_P)$ of maximizers of $(COP_P)$,

$$\delta(\mathbf{x}, \mathcal{S}(COP_P)) := \min\{\|\mathbf{x} - \overline{\mathbf{x}}\| : \overline{\mathbf{x}} \in \mathcal{S}(COP_P)\}.$$

Under feasibility of $(COP_P)$ the existence of solutions $\overline{\mathbf{x}}_d$ of $(P_d)$ and the convergence towards $\mathcal{S}(COP_P)$ follow by only assuming the dual constraint qualification $(CQ_D)$, or equivalently, the boundedness of the level sets $\mathcal{F}_\alpha(COP_P)$ (or the condition $\emptyset \neq \mathcal{S}(COP_P)$ compact), see Theorem 1.13.

**Theorem 5.6.** *Let $(COP_P)$ be feasible and let $(CQ_D)$ be satisfied. Then for any meshsize $d$ small enough, the sets $\mathcal{S}(P_d)$ of optimal solutions of $(P_d)$ are nonempty and compact. Moreover, for any sequence of solutions $\overline{\mathbf{x}}_d \in \mathcal{S}(P_d)$ we have $\delta(\overline{\mathbf{x}}_d, \mathcal{S}(COP_P)) \to 0$ for $d \to 0$.*

*Proof.* See [108, Theorem 9] for a proof. See also [38, Theorem 4.2(b),(c)]) for a proof under slightly stronger assumptions. □

Since the feasible set $\mathcal{F}(COP_P)$ may consists of a single point, it is clear that, in order to ensure the existence of a feasible point the inner approximation $(\widetilde{P}_d)$, we have to assume that $\mathcal{F}(COP_P)$ has interior points (see also [38, Theorem 4.2]).

**Theorem 5.7.** *Let $(CQ_P)$ and $(CQ_D)$ hold. Then for any meshsize $d$ small enough the sets $\mathcal{S}(\widetilde{P}_d)$ of optimal solutions of $(\widetilde{P}_d)$ are nonempty and compact. Moreover, for any sequence of solutions $\widetilde{\mathbf{x}}_d \in \mathcal{S}(\widetilde{P}_d)$ we have $\delta(\widetilde{\mathbf{x}}_d, \mathcal{S}(COP_P)) \to 0$ for $d \to 0$.*

*Proof.* If $(CQ_P)$ holds for $\mathbf{x}_0$, then we find from (5.10) that for all $(\mathbf{u}, \mathbf{v}) \in E_d$ and $d$ small enough

$$\begin{aligned} \mathbf{u}^T F(\mathbf{x}_0)\mathbf{v} &= \tfrac{1}{2}\left[\mathbf{u}^T F(\mathbf{x}_0)\mathbf{u} + \mathbf{v}^T F(\mathbf{x}_0)\mathbf{v} - (\mathbf{u}-\mathbf{v})^T F(\mathbf{x}_0)(\mathbf{u}-\mathbf{v})\right] \\ &\geq \sigma_0 - \tfrac{1}{2}\|F(\mathbf{x}_0)\| \cdot d^2 \geq 0. \end{aligned}$$

Hence $\mathbf{x}_0 \in \mathcal{F}(\widetilde{P}_d)$ if $d$ is small. By $(CQ_D)$ the level sets $\mathcal{F}_\alpha(COP_P)$ are bounded (compact) (see Theorem 1.13). Since $\mathcal{F}_\alpha(\widetilde{P}_d) \subset \mathcal{F}_\alpha(COP_P)$ (see (5.12)), also the level sets $\mathcal{F}_\alpha(\widetilde{P}_d)$ are bounded. Therefore, solutions $\widetilde{\mathbf{x}}_d$ of the linear programs $(\widetilde{P}_d)$ exist and the sets $\mathcal{S}(\widetilde{P}_d)$ of maximizers are nonempty and compact.

Suppose now that a sequence $\widetilde{\mathbf{x}}_{d_k}$ of such solutions does not satisfy

$$\delta(\widetilde{\mathbf{x}}_{d_k}, \mathcal{S}(COP_P)) \to 0 \text{ for } k \to \infty.$$

Then there exists $\varepsilon > 0$ and a subsequence $\widetilde{\mathbf{x}}_{d_{k_\nu}}$ such that

$$\delta(\widetilde{\mathbf{x}}_{d_{k_\nu}}, \mathcal{S}(COP_P)) \geq \varepsilon \quad \forall \nu. \tag{5.13}$$

Since the minimizers $\widetilde{\mathbf{x}}_{d_{k_\nu}}$ are elements of a compact set $\mathcal{F}_\alpha(COP_P)$ we can select a convergent subsequence and without loss of generality we can assume,

$$\widetilde{\mathbf{x}}_{d_{k_\nu}} \to \hat{\mathbf{x}} \in \mathcal{F}_\alpha(COP_P) \text{ for } \nu \to \infty.$$

In view of Lemma 5.5 (c) we have $val(P_d) - val(\widetilde{P}_d) \to 0$ and thus, by (5.12),

$$val(\widetilde{P}_d) \to val(COP_P), \;\; d \to 0.$$

This yields,

$$\mathbf{c}^T\widetilde{\mathbf{x}}_{d_{k_\nu}} = val(P_{d_{k_\nu}}) \to \mathbf{c}^T\hat{\mathbf{x}} = val(COP_P), \;\; \nu \to \infty$$

and since $\hat{\mathbf{x}} \in \mathcal{F}_\alpha(COP_P)$ is feasible for $(COP_P)$ we obtain $\hat{\mathbf{x}} \in \mathcal{S}(COP_P)$ contradicting (5.13). □

The next example shows that it may happen that every program $(P_d)$ and $(\widetilde{P}_d)$ has a solution while no solution of the original program $(COP_P)$ exists.

**Example 5.8.** *Consider the copositive program ( based on [30, Theorem 3.1]) with* $\mathbf{c} = (1,1,0)^T$ *and*

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

*Then* $(COP_P)$ *becomes:*

$$\max\; x_1 + x_2 \;\; s.t. \;\; F(x_1,x_2,x_3) := \begin{pmatrix} 1-x_1 & 0 & 0 \\ 0 & -x_2 & -1 \\ 0 & -1 & -x_3 \end{pmatrix} \in \mathcal{C}_3\,.$$

*The feasibility conditions for this program read:*

$$x_1 \le 1\,, \;\; x_2 \le 0\,, \;\; x_3 \le 0\,, \;\; x_2x_3 \ge 1\,.$$

*Obviously,* $x_1 + x_2 \le 1$ *holds for any feasible* $\mathbf{x}$ *and for any* $\epsilon > 0$ *the point* $\mathbf{x} = (1, -\epsilon, -1/\epsilon)^T$ *is feasible with objective value* $x_1 + x_2 = 1 - \epsilon$*. On the other hand, no feasible* $\overline{\mathbf{x}}$ *exists with objective* $\overline{x}_1 + \overline{x}_2 = 1$ *(*$\overline{x}_2 = 0$ *is excluded). So, the sup value of* $(COP_P)$ *is* $val(COP_P) = 1$ *but a maximizer does not exist. Now, consider the program* $(P_d)$*:*

$$(P_d) \qquad \max\; x_1 + x_2 \quad s.t. \quad \mathbf{z}^T F(x_1,x_2,x_3)\mathbf{z} \ge 0 \quad \forall \mathbf{z} \in Z_d.$$

*where* $Z_d$ *is any (finite) discretization of* $\Delta_3$ *containing the basis vectors* $\mathbf{z} = e_i \in \mathbb{R}^3,\ i = 1,2,3$. *Then* $(P_d)$ *in particular contains the constraints*

$$e_i^T F(x) e_i \geq 0,\ i = 1,2,3 \quad \text{or} \quad 1 - x_1 \geq 0\ ,\ x_2 \leq 0\ ,\ x_3 \leq 0\ .$$

*This implies* $x_1 + x_2 \leq 1$. *So, the linear program* $(P_d)$ *is bounded and a solution exists. In fact, any program* $(P_d)$ *has a solution* $\overline{\mathbf{x}}_d = (1, 0, \overline{x}_3(d))^T$ *with objective value* $val(P_d) = 1$ *(and* $\overline{x}_3(d) \to -\infty$ *for* $d \to 0$*).*

*Note that also the inner LP-approximations* $(\widetilde{P}_d)$ *have solutions. Indeed, since the feasible sets* $\mathcal{F}(\widetilde{P}_d)$ *are contained in* $\mathcal{F}(COP_P)$, *the values* $val(\widetilde{P}_d)$ *are bounded by 1. Moreover the feasible sets are non-empty. To see this take e.g. the* $(CQ_P)$ *-point* $\mathbf{x}_0 = (0, -2, -2)^T$ *in the interior of* $\mathcal{F}(COP_P)$. *Then, as in the proof of Theorem 5.7, it follows* $\mathbf{x}_0 \in \mathcal{F}(\widetilde{P}_d)$, *provided* $d$ *is small enough.*

We finish this section with some remarks. Note that for any solution $\overline{\mathbf{x}}_d$ of the standard linear program $(P_d)$ the KKT condition holds:

$$\mathbf{c} = \sum_{j=1}^{k} y_j \cdot \left( \mathbf{z}_j^T A_1 \mathbf{z}_j, \ldots, \mathbf{z}_j^T A_n \mathbf{z}_j \right), \quad \text{for some } y_j \geq 0,\ \mathbf{z}_j \in Z_d(\overline{\mathbf{x}}_d)\ , \tag{5.14}$$

where $Z_d(\overline{\mathbf{x}}_d) := \{\mathbf{z} \in Z_d : \mathbf{z}^T F(\overline{\mathbf{x}}_d) \mathbf{z} = 0\}$. Again, any such solution $\overline{\mathbf{x}}_d$ generates a dual feasible matrix

$$\overline{Y}_d := \sum_{j=1}^{k} y_j \mathbf{z}_j \mathbf{z}_j^T \in \mathcal{F}(COP_D),$$

such that

$$\langle \overline{Y}_d, B \rangle = val(P_d) \geq val(COP_D) \geq val(COP_P).$$

**Remark 5.9.** *Any solution* $\widetilde{\mathbf{x}}_d$ *of* $(\widetilde{P}_d)$ *also satisfies the KKT condition*

$$\mathbf{c} = \sum_{j=1}^{s} \widetilde{y}_j \cdot \left( \mathbf{u}_j^T A_1 \mathbf{v}_j, \ldots, \mathbf{u}_j^T A_n \mathbf{v}_j \right), \quad \widetilde{y}_j \geq 0,\ (\mathbf{u}_j, \mathbf{v}_j) \in E_d(\widetilde{\mathbf{x}}_d),\ s \in \mathbb{N},$$

*where* $E_d(\widetilde{\mathbf{x}}_d) := \{(\mathbf{u}, \mathbf{v}) \in E_d : \mathbf{u}^T F(\widetilde{\mathbf{x}}_d) \mathbf{v} = 0\}$. *Such a solution* $\widetilde{\mathbf{x}}_d$ *generates the matrix* $\widetilde{Y}_d := \sum_{j=1}^{s} \widetilde{y}_j \cdot \frac{1}{2}(\mathbf{u}_j \mathbf{v}_j^T + \mathbf{v}_j \mathbf{u}_j^T)$ *which satisfies the constraints* $\langle \widetilde{Y}_d, A_i \rangle = c_i$ *for all* $i$. *However, in general,* $\widetilde{Y} \notin \mathcal{C}_m^*$, *so* $\widetilde{Y}$ *is not necessarily feasible for* $(COP_D)$. *Using* (5.10), *we see (under the assumption of Theorem 5.6) that* $\widetilde{Y}_d$ *is in* $\mathcal{C}_m^*$ *up to an error of order* $\mathcal{O}(d^2)$.

### 5.3.1 Comparison with an Inner Approximation

In this subsection, we consider a special discretization scheme first considered in [47] which is connected to an inner approximation of $\mathcal{C}_m$ by subsets $\mathcal{C}_m^r \subset \mathcal{C}_m$. For $r \in \mathbb{N}$, let us define

$$\mathcal{C}_m^r := \left\{ A \in \mathcal{S}_m : \sum_{i,j=1}^m a_{ij} x_i^2 x_j^2 \left( \sum_{k=1}^m x_k^2 \right)^r \text{ has non-negative coefficients} \right\}.$$

The following is shown in [47] :

$$\mathcal{C}_m^r \subset \mathcal{C}_m^{r+1} \subset \ldots \subset \mathcal{C}_m \quad \text{and} \quad \text{cl}\left( \lim_{r \to \infty} \mathcal{C}_m^r \right) = \mathcal{C}_m.$$

The interesting connection with the discretization approach above is based on the following description of the sets $\mathcal{C}_m^r$(see[28]),

$$\mathcal{C}_m^{r-2} = \{ A \in \mathcal{S}_m : \mathbf{v}^T A \mathbf{v} - \mathbf{v}^T \operatorname{diag}(A) \geq 0 \quad \text{for all } \mathbf{v} \in \mathbb{I}_m^r \}, \tag{5.15}$$

where $\mathbb{I}_m^r$ is the grid $\mathbb{I}_m^r = \{ \mathbf{v} \in \mathbb{N}^m : \sum_{j=1}^m \mathbf{v}_j = r \}$. By (5.15), we can write,

$$\mathcal{C}_m^{r-2} = \{ A \in \mathcal{S}_m : \mathbf{z}^T A \mathbf{z} - \tfrac{1}{r} z^T \operatorname{diag}(A) \geq 0 \ \text{ for all } \mathbf{z} \in Z_d^0 := \tfrac{1}{r}\mathbb{I}_m^r \}. \tag{5.16}$$

**Remark 5.10.** *Note that the cone $\mathcal{C}_m^{r-2}$ can be seen as the special instance of the generalised cone, $\mathcal{C}_m(K, \alpha)$ (see (2.4)) by taking $K = \text{cone}(\frac{1}{r}\mathbb{I}_m^r), \alpha = \frac{1}{r}$. In this setting the dual of $\mathcal{C}_m^{r-2}$ is given by*

$$\mathcal{C}_m^{r-2*} = \left\{ U \in \mathcal{S}_m : U = \sum_i (\mathbf{u}_i \mathbf{u}_i^T - \frac{1}{r} \operatorname{Diag}(\mathbf{u}_i)), \mathbf{u}_i \in K \right\}.$$

It is not difficult to see that the set $Z_d^0 := \frac{1}{r}\mathbb{I}_m^r$ defines a uniform discretization of the simplex $Z = \Delta_m$ with meshsize of $Z_d^0$ given by

$$d = \max_{\mathbf{z}_j \in \frac{1}{r}\mathbb{I}_m^r} \ \min_{\mathbf{z}_i \in \frac{1}{r}\mathbb{I}_m^r,\ \mathbf{z}_i \neq \mathbf{z}_j} \|\mathbf{z}_j - \mathbf{z}_i\| = \tfrac{\sqrt{2}}{r}.$$

So it is natural to compare the outer and inner approximations $(P_d), (\widetilde{P}_d)$ of $(COP_P)$ in Section 5.3 with the following approximations, where $d = \sqrt{2}/r$, $r \in \mathbb{N}$:

$$(\widehat{P}_d) \quad \max_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^T \mathbf{x} \quad s.t. \quad \mathbf{z}^T F(\mathbf{x}) \mathbf{z} - \tfrac{d}{\sqrt{2}} \mathbf{z}^T \operatorname{diag}(F(\mathbf{x})) \geq 0 \ \ \forall \mathbf{z} \in Z_d^0. \tag{5.17}$$

Note that, by (5.16), a point $\mathbf{x}$ is feasible for $(\widehat{P}_d)$ if and only if $F(\mathbf{x}) \in \mathcal{C}_m^{r-2}$. So $(\widehat{P}_d)$ provides an inner approximation, i.e., $\mathcal{F}(\widehat{P}_d) \subset \mathcal{F}(COP_P)$ and $val(\widehat{P}_d) \leq val(COP_P)$. Similar to Lemma 5.5 we obtain

**Lemma 5.11.** *Let $(CQ_P)$ be satisfied for $\mathbf{x}_0 \in \mathcal{F}(COP_P)$. Then with the solutions $\overline{\mathbf{x}}_d$ of $(P_d)$ (with discretization $Z_d = Z_d^0$) the following holds for all $d = \frac{\sqrt{2}}{r},\ r \in \mathbb{N}$, $d$ small enough:*

$$\hat{\mathbf{x}}_d^* = \overline{\mathbf{x}}_d + \tau d(\mathbf{x}_0 - \overline{\mathbf{x}}_d) \in \mathcal{F}(\widehat{P}_d) \subset \mathcal{F}(COP_P)$$

*and*

$$0 \leq val(P_d) - val(\widehat{P}_d) \leq \tau[\mathbf{c}^T(\overline{\mathbf{x}}_d - \mathbf{x}_0)] \cdot d$$

*if* $\tau \geq \frac{\|\operatorname{diag}(F(\overline{\mathbf{x}}_d))\|}{\sqrt{2}\sigma_0 - d\|\operatorname{diag}(F(\mathbf{x}_0))\|}$ *and* $0 < \tau d < 1$.

*Proof.* We use the relation $F(\hat{\mathbf{x}}_d^*) = (1-\tau d)F(\overline{\mathbf{x}}_d) + \tau d F(\mathbf{x}_0)$ and proceed as in the proof of Lemma 5.5. By Lemma 5.4, using the relation $\|\mathbf{z}\| \leq 1$ for $\mathbf{z} \in Z_d^0$ and $\mathbf{z}^T F(\overline{\mathbf{x}}_d)\mathbf{z} \geq 0$ for $\mathbf{z} \in Z_d = Z_d^0$, we obtain for any $\mathbf{z} \in Z_d^0$:

$$\begin{aligned}
\mathbf{z}^T F(\hat{\mathbf{x}}_d^*)\mathbf{z} - \tfrac{d}{\sqrt{2}}\mathbf{z}^T \operatorname{diag}(F(\hat{\mathbf{x}}_d^*)) &= [(1-\tau d)\mathbf{z}^T F(\overline{\mathbf{x}}_d)\mathbf{z} + \tau d \mathbf{z}^T F(\mathbf{x}_0)\mathbf{z} \\
&\quad - \tfrac{d}{\sqrt{2}}(1-\tau d)\mathbf{z}^T \operatorname{diag}(F(\overline{\mathbf{x}}_d)) - \tau \tfrac{d^2}{\sqrt{2}}\mathbf{z}^T \operatorname{diag}(F(\mathbf{x}_0))] \\
&\geq \tau d \sigma_0 - \tfrac{d}{\sqrt{2}}(1-\tau d)\|\operatorname{diag}(F(\overline{\mathbf{x}}_d))\| - \tau \tfrac{d^2}{\sqrt{2}}\|\operatorname{diag}(F(\mathbf{x}_0))\| \\
&\geq d\left[\tau\left(\sigma_0 - \frac{d}{\sqrt{2}}\|\operatorname{diag}(F(\mathbf{x}_0))\|\right) - \frac{\|\operatorname{diag}(F(\overline{\mathbf{x}}_d))\|}{\sqrt{2}}\right] \geq 0
\end{aligned}$$

for any $d > 0$ (small enough) if $\tau$ is as given above. This shows the first relation. The inequality for the maximum values follows again easily using that $\hat{\mathbf{x}}_d^*$ is feasible for $(\widehat{P}_d)$ :

$$0 \leq val(P_d) - val(\widehat{P}_d) \leq \mathbf{c}^T(\overline{\mathbf{x}}_d - \hat{\mathbf{x}}_d^*) = [\mathbf{c}^T(\overline{\mathbf{x}}_d - \mathbf{x}_0)\tau] \cdot d \qquad \square$$

According to the analysis above, under the assumption that the sequence $\overline{\mathbf{x}}_d$, $d \to 0$, is bounded (cf., Theorem 5.6), we have established the following error bounds (the last bound holds for $Z_d^0 = \frac{1}{r}\mathbb{I}_m^r$ with $d = \sqrt{2}/r,\ r \in \mathbb{N}$):

$$\begin{aligned}
0 \leq & val(P_d) - val(COP_P) \leq \mathcal{O}(d^2), \\
0 \leq & val(COP_P) - val(\widetilde{P}_d) \leq \mathcal{O}(d^2), \\
0 \leq & val(COP_P) - val(\widehat{P}_d) \leq \mathcal{O}(d).
\end{aligned}$$

The next example shows that the bound $\mathcal{O}(d)$ for $(\widehat{P}_d)$ is sharp.

**Example 5.12.** *We consider the program,*

$$(P) \qquad \max_{x\in\mathbb{R}} x \quad s.t.\ F(x) := \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + x \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \in \mathcal{C}_2.$$

*The maximizer of $(COP_P)$ is $\overline{x} = 0$ with $val(COP_P) = \overline{x} = 0$. The corresponding unique active index is $\overline{\mathbf{z}} = (\frac{1}{2}, \frac{1}{2})^T$. For odd $r = 2l + 1$ and $d = \sqrt{2}/r$, the discretization $Z_d^0$ of $Z = \Delta_2 = \{\mathbf{z} \in \mathbb{R}_+^2 : z_1 + z_2 = 1\}$ is given by*

$$Z_d^0 = \left\{ \mathbf{z}(\lambda) := \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (1-\lambda) \begin{pmatrix} 0 \\ 1 \end{pmatrix} : \lambda = \tfrac{i}{r},\ i = 0, \dots, r \right\}.$$

*It is not difficult to see that the optimal solutions of $(P_d), (\widehat{P}_d)$ are given by the solutions $\overline{x}_d, \hat{x}_d$ of the equations*

$$\begin{aligned} \mathbf{z}\left(\frac{l}{2l+1}\right)^T F(x)\mathbf{z}\left(\frac{l}{2l+1}\right) &= 0, \\ \mathbf{z}\left(\frac{l}{2l+1}\right)^T F(x)\mathbf{z}\left(\frac{l}{2l+1}\right) - \frac{d}{\sqrt{2}}\mathbf{z}\left(\frac{l}{2l+1}\right)^T \operatorname{diag} F(x) &= 0, \end{aligned}$$

*respectively. After some calculations we obtain $val(P_d) = \overline{x}_d = \frac{1}{2}\frac{1}{l(l+1)} = \mathcal{O}(d^2)$ and*

$$val(\widehat{P}_d) = \hat{x}_d = -\frac{\sqrt{2}d}{2}\left[1 + \frac{2l^2+1}{2l(l+1)}\right] + \frac{1}{2l(l+1)} = -\sqrt{2}d + \mathcal{O}(d^2) = \mathcal{O}(d).$$

Let us finally compare the inner approximations $(\widetilde{P}_d)$ and $(\widehat{P}_d)$. It is not difficult to show that the number of points in the discretization $Z_d^0$ for $d = \frac{\sqrt{2}}{(r+2)}$ (approximation by $\mathcal{C}_m^r$ see [154]) are given by $N := \binom{m+r-1}{r}$. To obtain a corresponding inner approximation $(\widetilde{P}_d)$ one could think of the so-called Delauney triangulation (by simplices) of the point set $Z_d^0$. The number of edges in such a triangulation is ``much smaller'' than $N^2$ (edge from each point to each other, instead of edges only to ``neighbouring points''). So (for fixed $m$) the same order of approximation $O(\frac{1}{r^2})$ (wrt. $r$) would require ``much less'' than $N^2 = \binom{m+r-1}{r}^2$ constraints in $(\widetilde{P}_d)$ and $\binom{m+r^2-1}{r^2}$ constraints in $(\widehat{P}_d)$. This can be seen to be in favor of the scheme $(\widetilde{P}_d)$.

Interested in an inner approximation, one could also avoid both inner approximations $(\widetilde{P}_d), (\widehat{P}_d)$ and only make use of $(P_d)$. Indeed, the a-posteriori error bound of Lemma 5.5 allows us to construct a feasible point

$\overline{x}_d^* = \overline{x}_d + O(d^2)$ from the ``outer approximation'' $\overline{x}_d$ if a strictly feasible point $\mathbf{x}_0$ is available.

We wish to emphasize that in practice, the pure discretization methods have to be modified to a so-called exchange method where (as in [38]) during the computation only those grid points in $Z_d$ are kept in the discretization which still play a role as candidates for the active points $\mathbf{z}_j \in Z(\overline{\mathbf{x}})$ of a solution $\overline{\mathbf{x}}$ of $(COP_P)$ (see also [108]). For such exchange methods the bounds obtained above remain valid.

## 5.4 Order of Convergence for the Maximizers

In this section, we shortly discuss error bounds for $\|\overline{\mathbf{x}} - \overline{\mathbf{x}}_d\|, \|\overline{\mathbf{x}} - \widetilde{\mathbf{x}}_d\|, \|\overline{\mathbf{x}} - \hat{\mathbf{x}}_d\|$ for the maximizers of $(P_d), (\widetilde{P}_d), (\widehat{P}_d)$, respectively. These bounds are based on the concept of the order of a maximizer. A feasible point $\overline{\mathbf{x}} \in \mathcal{F}(COP_P)$ is a maximizer of $(COP_P)$ of order $p > 0$, iff with some $\gamma > 0, \varepsilon > 0$

$$\mathbf{c}^T\overline{\mathbf{x}} \geq \mathbf{c}^T\mathbf{x} + \gamma\|\mathbf{x} - \overline{\mathbf{x}}\|^p \qquad \text{for all } \mathbf{x} \in \mathcal{F}(COP_P),\ \|\mathbf{x} - \overline{\mathbf{x}}\| < \varepsilon \tag{5.18}$$

holds. Note that, if $\overline{\mathbf{x}}$ is a maximizer of order $0 < p$, in particular, $\mathcal{S}(COP_P) = \{\overline{\mathbf{x}}\}$ is nonempty and compact. So, by Theorem 1.13 the condition $(CQ_D)$ is satisfied and we can apply Theorem 5.6.

**Corollary 5.13.** *Let $(CQ_P)$ be satisfied and let $\overline{\mathbf{x}}$ be a maximizer of $(COP_P)$ of order $p \geq 1$. Then for the maximizers $\overline{\mathbf{x}}_d, \widetilde{\mathbf{x}}_d, \hat{\mathbf{x}}_d$ of $(P_d), (\widetilde{P}_d), (\widehat{P}_d)$, respectively, we have:*

$$\|\overline{\mathbf{x}} - \overline{\mathbf{x}}_d\| = \mathcal{O}(d^{2/p}), \qquad \|\overline{\mathbf{x}} - \widetilde{\mathbf{x}}_d\| = \mathcal{O}(d^{2/p}), \qquad \|\overline{\mathbf{x}} - \hat{\mathbf{x}}_d\| = \mathcal{O}(d^{1/p}).$$

*Proof.* Recall from Lemma 5.5(a) that $\overline{\mathbf{x}}_d^* := \overline{\mathbf{x}}_d + \rho d^2(\mathbf{x}_0 - \overline{\mathbf{x}}_d) \in \mathcal{F}(COP_P)$ for $\rho$ large enough. Using (5.18) and $\mathbf{c}^T(\overline{\mathbf{x}} - \overline{\mathbf{x}}_d) \leq 0$ we get ($\overline{\mathbf{x}}_d^*$ is feasible for $(COP_P)$),

$$\begin{aligned}\|\overline{\mathbf{x}} - \overline{\mathbf{x}}_d^*\|^p \leq \tfrac{1}{\gamma}\mathbf{c}^T(\overline{\mathbf{x}} - \overline{\mathbf{x}}_d^*) \ &= \ \tfrac{1}{\gamma}\mathbf{c}^T(\overline{\mathbf{x}} - \overline{\mathbf{x}}_d) - \tfrac{\rho}{\gamma}d^2\mathbf{c}^T(\mathbf{x}_0 - \overline{\mathbf{x}}_d) \\ \leq \tfrac{\rho}{\gamma}d^2\mathbf{c}^T(\overline{\mathbf{x}}_d - \mathbf{x}_0) \ &\leq \ \mathcal{O}(d^2)\end{aligned}$$

or $\|\overline{\mathbf{x}} - \overline{\mathbf{x}}_d^*\| \leq \mathcal{O}(d^{2/p})$. We thus find using $1 \leq p$,

$$\begin{aligned}\|\overline{\mathbf{x}} - \overline{\mathbf{x}}_d\| &\leq \|\overline{\mathbf{x}} - \overline{\mathbf{x}}_d^*\| + \|\overline{\mathbf{x}}_d^* - \overline{\mathbf{x}}_d\| \\ &\leq \mathcal{O}(d^{2/p}) + \mathcal{O}(d^2) = \mathcal{O}(d^{2/p})\ .\end{aligned}$$

The other bounds are proven in the same way. For $\hat{\mathbf{x}}_d$, e.g., we obtain using Lemma 5.11

$$\begin{aligned}\|\overline{\mathbf{x}} - \hat{\mathbf{x}}_d\| &\leq \|\overline{\mathbf{x}} - \hat{\mathbf{x}}_d^*\| + \|\hat{\mathbf{x}}_d^* - \hat{\mathbf{x}}_d\| \\ &= \mathcal{O}(d^{1/p}) + \mathcal{O}(d) = \mathcal{O}(d^{1/p})\ .\end{aligned}$$

According to this corollary, the smaller the order $p$ of the maximizer $\overline{\mathbf{x}}$, the faster is the convergence. The following examples show that for copositive programs $(COP_P)$ (unique) maximizer of order 1,2 and of arbitrarily large order can occur.

**Example 5.14.** *Obviously, in Example 5.8 the maximizer $\overline{x} = 0$ is of order $p = 1$. Considering the copositive program:*

$$(P) \qquad \max\ x_1 \quad \text{s.t.}\ \ F(\mathbf{x}) := \begin{pmatrix} -x_1 & x_2 & 0 \\ x_2 & 1 & 0 \\ 0 & 0 & -x_2 \end{pmatrix} \in \mathcal{C}_3\ ,$$

*we see that $\mathbf{x}$ is feasible if and only if $-x_1 \geq 0, -x_2 \geq 0$ and $-x_1 - x_2^2 \geq 0$ hold, or*

$$x_1 \leq 0\ , \quad x_2 \leq 0\ , \quad x_1 \leq -x_2^2.$$

*The maximum value is $x_1 = 0$ implying $x_2 = 0$. So $\overline{\mathbf{x}} = (0,0)^T$ is the unique maximizer. For the feasible points $\mathbf{x} = (-x_2^2, x_2)^T,\ x_2 < 0$ ($|x_2|$ small) we find with $||\mathbf{x}||_\infty := \max\{|x_1|, |x_2|\}$:*

$$\mathbf{c}^T\overline{\mathbf{x}} - \mathbf{c}^T\mathbf{x} = x_2^2 = ||\mathbf{x}||_\infty^2.$$

*and $\overline{\mathbf{x}}$ is a maximizer of order 2. Now, we take the program,*

$$(P) \qquad \max\ x_1 \quad s.t.\ \ F(\mathbf{x}) := \begin{pmatrix} -x_1 & x_2 & 0 & 0 & 0 \\ x_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -x_2 & x_3 & 0 \\ 0 & 0 & x_3 & 1 & 0 \\ 0 & 0 & 0 & 0 & -x_3 \end{pmatrix} \in \mathcal{C}_5.$$

*In view of the block structure of $F(\mathbf{x})$ a vector $\mathbf{x} \in \mathbb{R}^3$ is feasible if and only if:*

$$x_1 \leq 0\ , \quad x_2 \leq 0\ , \quad x_3 \leq 0\ , \quad x_1 \leq -x_2^2\ , \quad x_2 \leq -x_3^2$$

*Thus $\overline{\mathbf{x}} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$ is the (unique) maximizer and with feasible vectors $\mathbf{x} =$*

$\begin{pmatrix}-x_3^4 & -x_3^2 & x_3\end{pmatrix}^T$ $x_3 < 0$ *($|x_3|$ small) we find ,*

$$\mathbf{c}^T\overline{\mathbf{x}} - \mathbf{c}^T\mathbf{x} = x_3^4 = ||\mathbf{x}||_\infty^4 \ ,$$

*showing that $\overline{\mathbf{x}}$ is a maximizer of order 4. Similarly we can construct copositive programs with maximizer of arbitrarily large order.*

**Remark 5.15.** *In [28, Section 3], approximation results have been established for the values $v^* = \min_{\mathbf{z}\in\Delta_m} \mathbf{z}^T A\mathbf{z}$ with $A \in \mathcal{S}_m$. Note that $v^*$ is in fact the value of $(StQP)$ where maximization is replaced with minimization. We briefly show that these bounds appear in our result above as special instances. Obviously $v^*$ is the value of*

$$(COP_P) \qquad \max_{x\in\mathbb{R}} \ x \qquad s.t. \quad \mathbf{z}^T(A - xI)\mathbf{z} \geq 0 \ \ \forall \mathbf{z} \in Z := \Delta_m,$$

*with dual*

$$(D) \qquad \min_{Y\in\mathcal{S}_m} \ \langle Y, A\rangle \qquad s.t. \quad \langle Y, I\rangle = 1, \ \ Y \in \mathcal{C}_m^*.$$

*Obviously, $(COP_P)$ satisfies $(CQ_P)$ with some $\mathbf{x}_0$ (small enough) and also $(COP_D)$ has strictly feasible matrices $Y_0$ (with any $Y \in \text{int}(\mathcal{C}_m^*)$ take $Y_0 = Y/\langle Y, I\rangle$). Let $\overline{\mathbf{x}}$ be the solution of $(COP_P)$ and consider the approximations $(P_d), (\widetilde{P}_d)$ defined by the grids $Z_d^0$ ($d = \sqrt{2}/r$) with corresponding values $\mathbf{v}_d, \widetilde{\mathbf{v}}_d$ and solutions $\overline{\mathbf{x}}_d, \widetilde{\mathbf{x}}_d$. It is easy to see that these solutions must be unique, satisfy $\mathbf{x}_0 \leq \widetilde{\mathbf{x}}_d \leq \overline{\mathbf{x}} \leq \overline{\mathbf{x}}_d$ and are monotonic, i.e., $\widetilde{\mathbf{x}}_d \uparrow \overline{\mathbf{x}}, \overline{\mathbf{x}}_d \downarrow \overline{\mathbf{x}}$ for $d \to 0$. Then by Lemma 5.5(a) we obtain the bound*

$$0 \leq \mathbf{v}_d - v^* = \overline{\mathbf{x}}_d - \overline{\mathbf{x}} \leq \overline{\mathbf{x}}_d - \overline{\mathbf{x}}_d^* = \frac{\|F(\overline{\mathbf{x}}_d)\|}{2\sigma_0}\big(\overline{\mathbf{x}}_d - \mathbf{x}_0\big)d^2.$$

*and Lemma 5.11 yields*

$$0 \leq v^* - \widetilde{\mathbf{v}}_d \leq \mathbf{v}_d - \widetilde{\mathbf{v}}_d \leq \overline{\mathbf{x}}_d - \widetilde{\mathbf{x}}_d^* \leq \tau\big(\overline{\mathbf{x}}_d - \mathbf{x}_0\big)d \ .$$

*The latter gives (up to a constant factor) the bound in [28] and the first bound yields a $\mathcal{O}(d^2)$ error instead of a rate $\mathcal{O}(d)$ in [28].*

# Bibliography


[1] F. AHMED, M. DÜR, AND G. STILL, *Copositive programming via semi-infinite optimization*, Journal of Optimization Theory and Applications, 159 (2013), pp. 322--340.

[2] F. AHMED AND G. STILL, *A note on set-semidefinite relaxations of nonconvex quadratic programs*, Journal of Global Optimization, 57 (2013), pp. 1139--1146.

[3] F. AHMED AND G. J. STILL, *Quadratic maximization on the unit simplex: structure, stability, genericity and application in biology*, Memorandum 2034, Department of Applied Mathematics, University of Twente, Enschede, February 2014.

[4] F. A. AL-KHAYYAL, C. LARSEN, AND T. VAN VOORHIS, *A relaxation method for nonconvex quadratically constrained quadratic programs*, Journal of Global Optimization, 6 (1995), pp. 215--230.

[5] F. ALIZADEH, *Optimization over the positive semi-definite cone: interior-point methods and combinatorial applications*, in Advances in optimization and parallel computing, P. M. Pardalos, ed., North-Holland Publishing Co., Amsterdam, 1992, pp. 1--25.

[6] ———, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM Journal on Optimization, 5 (1995), pp. 13--51.

[7] E. J. ANDERSON AND A. S. LEWIS, *An extension of the simplex algorithm for semi-infinite linear programming*, Mathematical Programming, 44 (1989), pp. 247--269.

[8] M. F. ANJOS, *An improved semidefinite programming relaxation for the satisfiability problem*, Mathematical Programming, 102 (2005), pp. 589--608.

[9] K. M. ANSTREICHER, *Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming*, Journal of Global Optimization, 43 (2009), pp. 471--484.

[10] N. ARIMA, S. KIM, AND M. KOJIMA, *A quadratically constrained quadratic optimization model for completely positive cone programming*, SIAM Journal on Optimization, 23 (2013), pp. 2320--2340.

[11] A. BEN-TAL AND A. NEMIROVSKI, *Lectures on modern convex optimization*, MPS/SIAM Series on Optimization, Society for Industrial and Applied

Mathematics (SIAM), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications.

[12] A. Berman, *Cones, matrices and mathematical programming*, vol. 79, Springer, 1973.

[13] A. Berman and N. Shaked-Monderer, *Completely positive matrices*, World Scientific Publishing Co. Inc., River Edge, NJ, 2003.

[14] D. Bishop and C. Cannings, *A generalized war of attrition*, Journal of Theoretical Biology, 70 (1978), pp. 85--124.

[15] D. T. Bishop and C. Cannings, *Models of animal conflict*, Advances in Applied Probability, 8 (1976), pp. 616--621.

[16] M. J. Bomhoff, *Bipartite graphs and the decomposition of systems of equations*, PhD thesis, University of Twente, Enschede, January 2013.

[17] I. Bomze, *Non-cooperative two-person games in biology: A classification*, International Journal of Game Theory, 15 (1986), pp. 31--57.

[18] I. Bomze, *Copositive optimization--recent developments and applications*, European Journal of Operational Research, 216 (2012), pp. 509--520.

[19] I. Bomze, F. Jarre, and F. Rendl, *Quadratic factorization heuristics for copositive programming*, Mathematical Programming Computation, 3 (2011), pp. 37--57.

[20] I. M. Bomze, *Copositivity conditions for global optimality in indefinite quadratic programming problems*, Czechoslovak Journal of Operations Research, 1 (1992), pp. 7--19.

[21] I. M. Bomze, *Detecting all evolutionarily stable strategies*, Journal of Optimization Theory and Applications, 75 (1992), pp. 313--329.

[22] I. M. Bomze, *Evolution towards the maximum clique*, Journal of Global Optimization, 10 (1997), pp. 143--164.

[23] I. M. Bomze, *Global escape strategies for maximizing quadratic forms over a simplex*, Journal of Global Optimization, 11 (1997), pp. 325--338.

[24] ———, *On standard quadratic optimization problems*, Journal of Global Optimization, 13 (1998), pp. 369--387.

[25] ———, *Standard quadratic optimization problems: Algorithm, applications, theory*, in Encyclopedia of optimization, C. A. Floudas and P. M. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, 2009, pp. 3688--3693.

[26] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo, *The maximum clique problem*, in Handbook of combinatorial optimization, Supplement Vol. A, D. Du and P. M. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, 1999, pp. 1--74.

[27] I. M. Bomze and G. Danninger, *A finite algorithm for solving general quadratic problems*, Journal of Global Optimization, 4 (1994), pp. 1--16.

[28] I. M. Bomze and E. De Klerk, *Solving standard quadratic optimization problems via linear, semidefinite and copositive programming*, Journal of Global Optimization, 24 (2002), pp. 163--185.

[29] I. M. Bomze, M. Locatelli, and F. Tardella, *New and old bounds for standard quadratic optimization: dominance, equivalence and incomparability*, Mathematical Programming, 115 (2008), pp. 31--64.

[30] I. M. Bomze, W. Schachinger, and G. Uchida, *Think co (mpletely) positive! matrix properties, examples and a clustered bibliography on copositive optimization*, Journal of Global Optimization, 52 (2012), pp. 423--445.

[31] I. M. Bomze, W. Schachinger, and R. Ullrich, *From seven to eleven: completely positive matrices with high cp-rank*, Preprint NI14007-POP, Isaac Newton Institute for Mathematical Sciences, University of Cambridge UK, 2014.

[32] J. Borwein, *Necessary and sufficient conditions for quadratic minimality*, Numerical Functional Analysis and Optimization, 5 (1982), pp. 127--140.

[33] J. Borwein and H. Wolkowicz, *Regularizing the abstract convex program*, Journal of Mathematical Analysis and Applications, 83 (1981), pp. 495--530.

[34] M. Broom, *Bounds on the number of ESSs of a matrix game*, Mathematical Biosciences, 167 (2000), pp. 163--175.

[35] M. Broom, C. Cannings, and G. T. Vickers, *On the number of local maxima of a constrained quadratic form*, Proceedings of the Royal Society. London. Series A., 443 (1993), pp. 573--584.

[36] T. C. Brown, *A proof of Sperner's lemma via Hall's theorem*, Mathematical Proceedings of the Cambridge Philosophical Society, 78 (1975), p. 387.

[37] S. Bundfuss, *Copositive matrices,Copositive programming and applications*, PhD thesis, TU Damstadt, 2009.

[38] S. Bundfuss and M. Dür, *An adaptive linear approximation algorithm for copositive programs*, SIAM Journal on Optimization, 20 (2009), pp. 30--53.

[39] S. Burer, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Mathematical Programming, 120 (2009), pp. 479--495.

[40] S. Burer, *Copositive programming*, in Handbook of Conic and Polynomial Optimization, M. Anjos and J. Lasserre, eds., Springer, 2011.

[41] C. Cannings and G. Vickers, *Patterns and invasions of evolutionary stable strategies*, Applied Mathematics and Computation, 32 (1989), pp. 227--253.

[42] C. Cannings and G. T. Vickers, *Patterns of ESSs. II*, Journal of Theoretical Biology, 132 (1988), pp. 409--420.

[43] ———, *The geneaology of patterns of ESSs*, in Selected Proceedings of the Sheffield Symposium on Applied Probability (Sheffield), I. V. Basawa and R. L. Taylor, eds., vol. 18 of IMS Lecture Notes Monograph Series, Institute of Mathematical Statistic, 1991, pp. 193--204.

[44] A. Charnes, W. Cooper, and K. Kortanek, *Duality in semi-infinite programs and some works of Haar and Caratheodory*, Management Science, 9 (1963), pp. 209--228.

[45] R. W. Cottle, G. J. Habetler, and C. E. Lemke, *On classes of copositive matrices*, Linear Algebra and Applications, 3 (1970), pp. 295--310.

[46] ———, *Quadratic forms semi-definite over convex cones*, in Proceedings of the Princeton Symposium on Mathematical Programming (Princeton Univiversity, 1967), H. W. Kuhn, ed., Princeton, N.J., 1970, Princeton University Press, pp. 551--565.

[47] E. de Klerk and D. V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM Journal on Optimization, 12 (2002), pp. 875--892.

[48] ———, *A linear programming reformulation of the standard quadratic optimization problem*, Journal of Global Optimization, 37 (2007), pp. 75--84.

[49] E. de Klerk, D. V. Pasechnik, and R. Sotirov, *On semidefinite programming relaxations of the traveling salesman problem*, SIAM Journal on Optimization, 19 (2008), pp. 1559--1573.

[50] P. H. Diananda, *On non-negative forms in real variables some or all of which are non-negative*, Mathematical Proceedings of the Cambridge Philosophical Society, 58 (1962), pp. 17--25.

[51] P. Dickinson and L. Gijben, *On the computational complexity of membership problems for the completely positive cone and its dual*, Computational Optimization and Applications, (2013), pp. 1--13.

[52] P. J. Dickinson, *Geometry of the copositive and completely positive cones*, Journal of Mathematical Analysis and Applications, 380 (2011), pp. 377--395.

[53] P. J. C. Dickinson, *An improved characterisation of the interior of the completely positive cone*, Electronic Journal of Linear Algebra, 20 (2010), pp. 723--729.

[54] P. J. C. Dickinson and M. Dür, *Linear-time complete positivity detection and decomposition of sparse matrices*, SIAM Journal on Matrix Analysis and Applications, 33 (2012), pp. 701--720.

[55] J. H. Drew, C. R. Johnson, and R. Loewy, *Completely positive matrices associated with m-matrices*, Linear and Multilinear Algebra, 37 (1994), pp. 303--310.

[56] I. Dukanovic and F. Rendl, *Semidefinite programming relaxations for graph coloring and maximal clique problems*, Mathematical Programming, 109 (2007), pp. 345--365.

[57] ———, *A semidefinite programming-based heuristic for graph coloring*, Discrete Applied Mathematics, 156 (2008), pp. 180--189.

[58] M. Dür, *Copositive programming--a survey*, in Recent advances in optimization and its applications in engineering, M. Diehl, F. Glineur, E. Jarlebring, and W. Michiels, eds., Springer, 2010, pp. 3--20.

[59] M. DÜR AND G. STILL, *Interior points of the completely positive cone*, Electronic Journal of Linear Algebra, 17 (2008), pp. 48--53.

[60] G. EICHFELDER AND J. JAHN, *Set-semidefinite optimization*, Journal of Convex Anal, 15 (2008), pp. 767--801.

[61] H. A. EISELT AND C.-L. SANDBLOM, *Operations research: a model-based approach*, Springer, 2012.

[62] K. ETESSAMI AND A. LOCHBIHLER, *The computational complexity of evolutionarily stable strategies*, International Journal of Game Theory, 37 (2008), pp. 93--113.

[63] U. FAIGLE, W. KERN, AND G. STILL, *Algorithmic principles of mathematical programming*, vol. 24 of Kluwer Texts in the Mathematical Sciences, Kluwer Academic Publishers, Dordrecht, 2002.

[64] A. V. FIACCO, *Introduction to sensitivity and stability analysis in nonlinear programming*, vol. 226, Academic Press New York, 1983.

[65] T. FUJIE AND M. KOJIMA, *Semidefinite programming relaxation for nonconvex quadratic programs*, Journal of Global Optimization, 10 (1997), pp. 367--380.

[66] J. GADDUM, *Linear inequalities and quadratic forms.*, Pacific Journal of Mathematics, 8 (1958), pp. 411--414.

[67] B. GÄRTNER AND J. MATOUŠEK, *Approximation algorithms and semidefinite programming*, Springer, Heidelberg, 2012.

[68] S. I. GASS, *Linear programming: methods and applications*, Courier Dover Publications, 2003.

[69] B. GHADDAR, M. F. ANJOS, AND F. LIERS, *A branch-and-cut algorithm based on semidefinite programming for the minimum k -partition problem*. To appear.

[70] L. E. GIBBONS, D. W. HEARN, P. M. PARDALOS, AND M. V. RAMANA, *Continuous characterizations of the maximum clique problem*, Mathematics of Operations Research, 22 (1996), pp. 754--768.

[71] M. A. GOBERNA AND V. JORNET, *On Haar's dual problem*, OR Spektrum, 18 (1996), pp. 209--217.

[72] M. A. GOBERNA AND M. A. LÓPEZ, *Linear semi-infinite optimization*, vol. 2 of Wiley Series in Mathematical Methods in Practice, John Wiley & Sons Ltd., Chichester, 1998.

[73] M. Á. GOBERNA AND M. A. LÓPEZ, eds., *Semi-infinite programming*, vol. 57 of Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, 2001. Recent advances, Papers from the conference (SIP'99) held in Alicante, September 1999.

[74] M. X. GOEMANS AND D. P. WILLIAMSON, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, Journal of the Association for Computing Machinery, 42 (1995), pp. 1115--1145.

[75] G. H. Golub and C. F. Van Loan, *Matrix computations*, vol. 3, JHU Press, 2012.

[76] M. S. Gowda, R. Sznajder, and J. Tao, *The automorphism group of a completely positive cone and its lie algebra*, Linear Algebra and its Applications, (2011).

[77] M. Grötschel, ed., *Optimization Stories*, 21st ISMP, Berlin August 2012, 2012.

[78] K. Hadeler, *On copositive matrices*, Linear Algebra and its Applications, 49 (1983), pp. 79--89.

[79] J. Haigh, *The existance of evolutionary stable strategies*, Journal of Theoretical Biology, 47 (1974), pp. 209--221. Appendix to "The theory of games and the evolution of animal conflicts".

[80] J. Haigh, *The distribution of evolutionarily stable strategies*, Journal of Applied Probability, (1988), pp. 233--246.

[81] ———, *How large is the support of an ESS?*, Journal of Applied Probability, (1989), pp. 164--170.

[82] M. Hall, Jr. and M. Newman, *Copositive and completely positive quadratic forms*, Mathematical Proceedings of the Cambridge Philosophical Society, 59 (1963), pp. 329--339.

[83] E. Halperin and U. Zwick, *Approximation algorithms for MAX 4-SAT and rounding procedures for semidefinite programs*, Journal of Algorithms, 40 (2001), pp. 184--211.

[84] J. Hannah and T. J. Laffey, *Nonnegative factorization of completely positive matrices*, Linear Algebra and its Applications, 55 (1983), pp. 1--9.

[85] S. Hart, Y. Rinott, and B. Weiss, *Evolutionarily stable strategies of random games, and the vertices of random polygons*, The Annals of Applied Probability, (2008), pp. 259--287.

[86] C. Helmberg, F. Rendl, and R. Weismantel, *A semidefinite programming approach to the quadratic knapsack problem*, Journal of Combinatorial Optimization, 4 (2000), pp. 197--215.

[87] R. Hettich, *A comparison of some numerical methods for semi-infinite programming*, in Semi-infinite programming, R. Hettich, ed., vol. 15 of Lecture Notes in Control and Information Sci., Springer, Berlin, 1979, pp. 112--125.

[88] R. P. Hettich and K. O. Kortanek, *Semi-infinite programming: theory, methods, and applications*, SIAM Review, 35 (1993), pp. 380--429.

[89] R. Hildebrand, *The extreme rays of the* $5 \times 5$ *copositive cone*, Linear Algebra and its Applications, 437 (2012), pp. 1538--1547.

[90] C. J. Hillar, S. Lin, and A. Wibisono, *Inverses of symmetric, diagonally dominant positive matrices and applications*, `arXiv:1203.6812v1`, (2012).

[91] W. Hines, *Evolutionary stable strategies: A review of basic theory*, Theoretical Population Biology, 31 (1987), pp. 195 -- 272.

[92] J.-B. Hiriart-Urruty and A. Seeger, *A variational approach to copositive matrices*, SIAM Review, 52 (2010), pp. 593--629.

[93] A. Hoffman and F. Pereira, *On copositive matrices with -1, 0, 1 entries*, Journal of Combinatorial Theory - Series A, 14 (1973), pp. 302--309.

[94] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 2012.

[95] T. Ibaraki and N. Katoh, *Resource allocation problems: Algorithmic approaches*, MIT Press, 1998.

[96] K. D. Ikramov and N. V. Savel′eva, *Conditionally definite matrices*, Journal of Mathematical Sciences, 98 (2000), pp. 1--50.

[97] F. Jarre and K. Schmallowsky, *On the computation of $C^*$ certificates*, Journal of Global Optimization, 45 (2009), pp. 281--296.

[98] K. Jittorntrum, *Solution point differentiability without strict complementarity in nonlinear programming*, Mathematical Programming Study, (1984), pp. 127--138.

[99] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373--395.

[100] M. Kaykobad, *On nonnegative factorization of matrices*, Linear Algebra and its Applications, 96 (1987), pp. 27--33.

[101] L. G. Khachiyan, *A polynomial algorithm in linear programming*, Doklady Akademii Nauk SSSR, 244 (1979).

[102] J. Kingman, *Typical polymorphisms maintained by selection at a single locus*, Journal of Applied Probability, (1988), pp. 113--125.

[103] J. F. C. Kingman, *On an inequality in partial averages*, The Quarterly Journal of Mathematics. Oxford. Second Series, 12 (1961), pp. 78--80.

[104] V. Klee and G. J. Minty, *How good is the simplex algorithm?*, in Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin), O. Shisha, ed., Academic Press, New York, 1972, pp. 159--175.

[105] M. Kojima and L. Tunçel, *Cones of matrices and successive convex relaxations of nonconvex sets*, SIAM Journal on Optimization, 10 (2000), pp. 750--778.

[106] J. B. Lasserre, *An explicit exact SDP relaxation for nonlinear 0-1 programs*, in Integer Programming and Combinatorial Optimization, K. Aardal and B. Gerards, eds., vol. Volume 2081/2001 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2001, pp. 293--303.

[107] A. N. Letchford and M. M. Sorensen, *Binary positive semidefinite matrices and associated integer polytopes*, in Integer Programming and Combinatorial Optimization, A. Lodi, A. Panconesi, and G. Rinaldi, eds., vol. Volume 5035/2008 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2008, pp. 125--139.

[108] M. López and G. Still, *Semi-infinite programming (invited review)*, European Journal of Operational Research, 180 (2007), pp. 491--518.

[109] V. Losert and E. Akin, *Dynamics of games and genes: Discrete versus continuous time*, Journal of Mathematical Biology, 17 (1983), pp. 241--251.

[110] L. Lovász, *On the Shannon capacity of a graph*, IEEE Transactions on Information Theory, 25 (1979), pp. 1--7.

[111] H. Markowitz, *Portfolio selection*, Journal of Finance, 7 (1952), pp. 77--91.

[112] J. Maynard Smith, *The theory of games and the evolution of animal conflicts*, Journal of Theoretical Biology, 47 (1974), pp. 209--221.

[113] J. Moon and L. Moser, *On cliques in graphs*, Israel Journal of Mathematics, 3 (1965), pp. 23--28.

[114] T. Motzkin, *Copositive quadratic forms*, Tech. Report 1818, National Bureau of Standards Report, 1952.

[115] T. Motzkin and E. Straus, *Maxima for graphs and a new proof of a theorem of Turan*, Canadian Journal of Mathematics, 17 (1965), pp. 533--540.

[116] H. P. Mulholland and C. A. B. Smith, *An inequality arising in genetical theory*, The American Mathematical Monthly, 66 (1959), pp. 673--683.

[117] K. G. Murty and S. N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Mathematical Programming, 39 (1987), pp. 117--129.

[118] N. Nisan, *A note on the computational hardness of evolutionary stable strategies*, Electronic Colloquium on Computational Complexity (ECCC), 13 (2006).

[119] M. Pelillo, *Replicator dynamics in combinatorial optimization*, in Encyclopedia of optimization, Springer, 2001, pp. 2197--2209.

[120] J. Peña, J. Vera, and L. F. Zuluaga, *Computing the stability number of a graph via linear and semidefinite programming*, SIAM Journal on Optimization, 18 (2007), pp. 87--105.

[121] L. Ping and Y. Y. Feng, *Criteria for copositive matrices of order four*, Linear Algebra and its Applications, 194 (1993), pp. 109--124.

[122] J. Povh, *Application of semidefinite and copositive programming in combinatorial optimization*, PhD thesis, Department of Mathematics, Faculty of Mathematics and Physics,University of Ljubljana, 2006.

[123] J. Povh and F. Rendl, *A copositive programming approach to graph partitioning*, SIAM Journal on Optimization, 18 (2007), pp. 223--241 (electronic).

[124] ———, *Copositive and semidefinite relaxations of the quadratic assignment problem*, Discrete Optimization. The Journal of Combinatorial Operations Research, 6 (2009), pp. 231--241.

[125] J. C. Preisig, *Copositivity and the minimization of quadratic functions with nonnegativity and quadratic equality constraints*, SIAM Journal on Control and Optimization, 34 (1996), pp. 1135--1150.

[126] A. J. Quist, E. De Klerk, C. Roos, and T. Terlaky, *Copositive relaxation for general quadratic programming*, Optimization Methods and Software, 9 (1998), pp. 185--208.

[127] M. V. Ramana, *An exact duality theory for semidefinite programming and its complexity implications*, Mathematical Programming, 77 (1997), pp. 129--162.

[128] M. V. Ramana, L. Tunçel, and H. Wolkowicz, *Strong duality for semidefinite programming*, SIAM Journal on Optimization, 7 (1997), pp. 641--662.

[129] R. Reemtsen and S. Görner, *Numerical methods for semi-infinite programming: a survey*, in Semi-infinite programming, R. Reemtsen and J.-J. Rückmann, eds., vol. 25 of Nonconvex Optimization and Applications, Kluwer Academic Publishers, Boston, MA, 1998, pp. 195--275.

[130] F. Rendl, G. Rinaldi, and A. Wiegele, *Solving max-cut to optimality by intersecting semidefinite and polyhedral relaxations*, technical report, Alpen-Adria-Universit at Klagenfurt, May 2007.

[131] R. T. Rockafellar, *Convex analysis*, vol. 28, Princeton University Press, 1997.

[132] C. Roos, T. Terlaky, and J.-P. Vial, *Theory and algorithms for linear optimization: an interior point approach*, Wiley Chichester, 1997.

[133] W. Rudin, *Principles of mathematical analysis*, vol. 3, McGraw-Hill New York, 1964.

[134] ———, *Real and complex analysis*, McGraw-Hill Education, 1987.

[135] A. Schaeffer and W. M. Greenlee, *Iterative methods for eigenvalues of symmetric matrices as fixed point theorems*, December 2007. Available at: `http://math.arizona.edu/~ura-reports/074/Schaeffer.Amanda/Final.pdf`.

[136] C. Schellewald and C. Schn"orr, *Subgraph matching with semidefinite programming*, Electronic Notes in Discrete Mathematics, 12 (2003), pp. 279 --289. 9th International Workshop on Combinatorial Image Analysis.

[137] P. Scheuer and S. Mandel, *An inequality in population genetics*, Heredity, 13 (1959), pp. 519--524.

[138] N. Shaked-Monderer, *A note on upper bounds on the cp-rank*, Linear Algebra and its Applications, 431 (2009), pp. 2407--2413.

[139] N. Shaked-Monderer, A. Berman, I. M. Bomze, F. Jarre, and W. Schachinger, *New results on the cp rank and related properties o co(mpletely)positive matrices*, Preprint available at: `http://arxiv.org/abs/1305.0737v2`, (2013).

[140] N. Shaked-Monderer, I. Bomze, F. Jarre, and W. Schachinger, *On the cp-rank and minimal cp factorizations of a completely positive matrix*, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 355--368.

[141] J. Smith, *Evolution and the Theory of Games*, Cambridge University Press, 1982.

[142] J. M. Smith and G. Price, *The logic of animal conflict*, Nature, 246 (1973), p. 15.

[143] J. Sponsel, S. Bundfuss, and M. Dür, *An improved algorithm to test copositivity*, Journal of Global Optimization, 52 (2012), pp. 537--551.

[144] G. Still, *Discretization in semi-infinite programming: the rate of convergence*, Mathematical Programming, 91 (2001), pp. 53--69.

[145] J. F. Sturm and S. Zhang, *On cones of nonnegative quadratic functions*, Mathematics of Operations Research, 28 (2003), pp. 246--267.

[146] H. Väliaho, *Criteria for copositive matrices*, Linear Algebra and its Applications, 81 (1986), pp. 19--34.

[147] E. Van Damme, *Stability and perfection of Nash equilibria*, Springer, 1991.

[148] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Review, 38 (1996), pp. 49--95.

[149] ———, *Connections between semi-infinite and semidefinite programming*, in Semi-infinite programming, vol. 25 of Nonconvex Optim. Appl., Kluwer Acad. Publ., Boston, MA, 1998, pp. 277--294.

[150] S. A. Vavasis, *Nonlinear optimization*, vol. 8 of International Series of Monographs on Computer Science, The Clarendon Press Oxford University Press, New York, 1991.

[151] G. T. Vickers and C. Cannings, *Patterns of ESSs. I*, Journal of Theoretical Biology, 132 (1988), pp. 387--408.

[152] H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., *Handbook of semidefinite programming*, International Series in Operations Research & Management Science, 27, Kluwer Academic Publishers, Boston, MA, 2000. Theory, algorithms, and applications.

[153] H. Wolkowicz and Q. Zhao, *Semidefinite programming relaxations for the graph partitioning problem*, Discrete Applied Mathematics, 96/97 (1999), pp. 461--479.

[154] E. A. Yildirim, *On the accuracy of uniform polyhedral approximations of the copositive cone*, Optimization Methods and Software, 27 (2012), pp. 155--173.

[155] Q. Zhao, *Semidefinite programming for Assignment and Partition Problem*, PhD thesis, University of Waterloo, Canada, 1996.

[156] A. Zhou and J. Fan, *Interiors of completely positive cones*, January 2014. Available at: `http://arxiv.org/abs/1401.1255`.

[157] J. Žilinskas, *Copositive programming by simplicial partition*, Informatica, 22 (2011), pp. 601--614.

[158] J. Žilinskas and M. Dür, *Depth-first simplicial partition for copositivity detection, with an application to maxclique*, Optimization Methods and Software, 26 (2011), pp. 499--510.

# Summary

In this thesis, copositive programming and problems associated with copositive programming are studied. Copositive programming refers to the following:

$$(COP_P) \qquad \max_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \qquad s.t. \quad B - \sum_{i=1}^{n} x_i A_i \in \mathcal{C}_m$$

$$(COP_D) \qquad \min_{Y\in\mathcal{S}_m} \langle Y, B\rangle \qquad s.t. \quad \langle Y, A_i\rangle = c_i \;\; \forall i = 1,\cdots,n, \;\; Y \in \mathcal{C}_m^*,$$

where $\mathcal{C}_m$ and $\mathcal{C}_m^*$ are, respectively, the cone of copositive and completely positive matrices defined below,

$$\mathcal{C}_m := \{A \in \mathcal{S}_m : \mathbf{v}^T A\mathbf{v} \geq 0 \text{ for all } \mathbf{v} \in \mathbb{R}_+^m\},$$

$$\mathcal{C}_m^* := \left\{A \in \mathcal{S}_m : A = \sum_{k=1}^{N} \mathbf{b}_k\mathbf{b}_k^T \text{ with } \mathbf{b}_k \in \mathbb{R}_+^m, N \in \mathbb{N}\right\}.$$

In the last decade copositive programming has gained much attention. A main contribution is the result of Burer [39], saying that mixed binary continuous optimization problems can be reformulated, exactly, as a copositive program.

Associated with the feasibility problem of copositive programming is the standard quadratic program (StQP). We have given a particular attention to this problem. We have provided a characterization for a KKT point to be a strict local maximizer of StQP. We have also analysed the effect of small perturbations, in the matrix involved, to strict local maximizers of StQP.

Strict local maximizers of StQP are related to the notion of evolutionarily stable strategy (ESS). In fact, for a symmetric matrix a point is a strict local maximzier of StQP if and only if it is an ESS. We have shown that for a symmetric matrix, with each principal submatrix nonsingular, there always exists an ESS. Moreover the existence of an ESS in symmetric matrices is a generic property.

A matrix $Q \in \mathcal{S}_m$ is said to be set-semidefinite if $\mathbf{v}^T Q\mathbf{v} \geq 0$ holds for all $\mathbf{v} \in K \subseteq \mathbb{R}^m$. The set of all set-semidefinite matrices forms a cone called the set-semidefinite cone. Cone programming problems over a cone of set-semidefinite

matrices are called set-semidefinite programs.

Hard optimization problems can be (approximately) reformulated by cone programming relaxations. This reformulation provides bounds for the original problem. In this thesis we have analysed the sharpness of set-semidefinite programming relaxations for quadratically constrained quadratic program (QCQP). The result we have obtained is somewhat negative. It roughly speaking says that without adding extra restrictions into the relaxation we cannot expect the set-semidefinite relaxation of (nonconvex) quadratic programs to be sharp.

Mathematical programming can be classified into finite and infinite problems. A special case of infinite problems is given by semi-infinite programming, where the number of constraints are infinite while the number of variables are finite. In this thesis we have considered the following primal linear semi-infinite programming problem,

$$(\mathrm{SIP}_P) \qquad \max_{\mathbf{x}\in\mathbb{R}^n} \ \mathbf{c}^T\mathbf{x} \qquad \textit{s.t.} \quad b(\mathbf{z}) - a(\mathbf{z})^T\mathbf{x} \geq 0 \quad \forall\, \mathbf{z} \in Z,$$

with an infinite compact index set $Z \subset \mathbb{R}^m$ and continuous functions $a : Z \to \mathbb{R}^n$ and $b : Z \to \mathbb{R}$.

An alternative condition for a matrix $Q$ to be copositive is that $\mathbf{v}^T Q\mathbf{v} \geq 0$ holds for all $\mathbf{v} \in \Delta_m$, where $\Delta_m$ is the standard simplex. By using this condition one can reformulate copositive programming as semi-infinite programming. We have used this reformulation to analyse copositive programming from the viewpoint of SIP.

A discritization of the simplex defines a simplicial partition. By using such partitions an approximation method for copositive programming is presented. This approximation method can be seen as a special case of a discritization method for semi-infinite programming. We have analysed the behaviour of the approximation error in dependence of the discretization meshsize $d$. We have shown that the error for the optimal values of the schemes in [38] behave like $\mathcal{O}(d^2)$ for $d \to 0$. Another scheme $(\widehat{P}_d)$ shows a convergence rate $\mathcal{O}(d)$. The concept of order of maximizers allows to analyse the behaviour of the error for the maximizers in the approximation schemes. It also has been shown that maximizer of arbitrary large order may appear in copositive programming.

# خلاصہ


مخلوط مثبتی کاملیت اور متعلقہ مسائل

ریاضیاتی کاملیت سے مراد ایسے ریاضیاتی مسائل ہیں جن میں دستیاب متبادلات کے مجموعہ سے بہترین رکن کا انتخاب کرنا ہوتا ہے ۔ اس مقالے میں ریاضیاتی کاملیت کی ایک ذیلی شاخ مخلوط مثبتی کاملیت اور اس سے متعلقہ مسائل پر بحث کی گئی ہے ۔ مخلوط مثبتی کاملیت سے ہماری مراد درج ذیل ریاضیاتی مسائل ہیں :

$$\max_{\mathbf{x}\in\mathbb{R}^n} \ \mathbf{c}^T\mathbf{x} \qquad s.t. \quad B-\sum_{i=1}^{n} x_iA_i \in \mathcal{C}_m$$

$$\min_{Y\in\mathcal{S}_m} \ \langle Y,B\rangle \qquad s.t. \quad \langle Y,A_i\rangle = c_i \ \forall i=1,\cdots,n, \ \ Y\in\mathcal{C}_m^*,$$

یہاں $\mathcal{C}_m$ اور $\mathcal{C}_m^*$ بالترتیب مخلوط مثبتی قالب اور مکمل مثبتی قالب سے تشکیل شدہ مخروطہ کو ظاہر کرتے ہیں جن کی تعریف ذیل میں دی گئی ہے :

$$\mathcal{C}_m := \{A\in\mathcal{S}_m : \mathbf{v}^TA\mathbf{v}\geq 0 \ \ \forall \ \mathbf{v}\in\mathbb{R}^m_+\},$$

$$\mathcal{C}_m^* := \left\{A\in\mathcal{S}_m : A=\sum_{k=1}^{N}\mathbf{b}_k\mathbf{b}_k^T \text{ with } \mathbf{b}_k\in\mathbb{R}^m_+, N\in\mathbb{N}\right\}.$$

گزشتہ دہائی میں مخلوط مثبتی کاملیت کو انتہائی اہمیت حاصل رہی ہے کیونکہ بیشتر مشکل مسائل کی تشکیل نو بطور مخلوط مثبتی کاملیت کے کی جا سکتی ہے ۔
کسی قالب کے مخلوط مثبتی ہونے کی تعریف سے یہ اخذ کیا جا سکتا ہے کہ مخلوط مثبتی کاملیت کی امکان پذیری دراصل ایک مربعی کاملیاتی مسئلہ ہے ۔ اس مسئلہ کو معیاری مربعی کاملیت کا نام دیا جاتا ہے جس کی ریاضیاتی شکل ذیل میں دی گئی ہے ۔

$$\max \quad q(\mathbf{v}) := \frac{1}{2}\mathbf{v}^TQ\mathbf{v} \quad s.t. \quad \mathbf{v}\in\Delta_m := \left\{\mathbf{v}\in\mathbb{R}^m_+ : \sum_{i=1}^{m} v_i = 1\right\}.$$

اس مقالہ میں ہم نے معیاری مربعی کاملیت کو خصوصی اہمیت دی ہے ۔ ہم نے ایسی کاملیاتی شرائط پر بحث کی ہے جو منفرد نکتہ ہونے کی ضمانت ہوتی ہیں ۔ ہم نے اس بات کا بھی جائزہ لیا ہے کہ قالب کے ارکان میں ہونے والی نامعلوم تبدیلی منفرد نکتہ پر کیسے اثر پذیر ہوتی ہے ۔

کسی دیے ہوئے قالب $Q$ کے لیے ایسا نکتہ $\mathbf{v} \in \Delta_m$ جو کہ درج ذیل شرائط پوری کرتا ہو ارتقائی مستحکم حکمت عملی کہلاتا ہے ۔

ا. $\mathbf{v}^T Q\mathbf{v} \geq \mathbf{u}^T Q\mathbf{v}$ تمام $\mathbf{u} \in \Delta_m$

ي. کسی $\mathbf{u} \in \Delta_m$ کیلیے اگر $\mathbf{v} \neq \mathbf{u}, \mathbf{v}^T Q\mathbf{v} = \mathbf{u}^T Q\mathbf{v}$ ہو تو $\mathbf{v}^T Q\mathbf{u} > \mathbf{u}^T Q\mathbf{u}$

معیاری مربعی کاملیت کے منفرد نکتہ کا تعلق ارتقائی مستحکم حکمت عملی سے ہے ۔ درحقیقت متناظر قالب کی صورت میں کوئی نکتہ منفرد نکتہ تب ہی ہو سکتا ہے جب وہ ارتقائی مستحکم حکمت عملی ہو ۔ اس کے علاوہ ہم نے یہ بھی ثابت کیا ہے کہ ایسے قالب جن کے ذیلی قالب غیر یکتائی ہوں کے لیے ہمیشہ کم از کم ایک ارتقائی مستحکم حکمت عملی ہوتی ہے ۔

کوئی قالب $Q$ طاقم نیم قطعی کہلاتا ہے اگر تمام $\mathbf{v} \in K \subseteq \mathbb{R}^m$ کے لیے $\mathbf{v}^T Q\mathbf{v}$ کی قیمت غیر منفی ہو ۔ طاقم نیم قطعی قالبوں کے طاقم سے تشکیل پانے والے مخروطہ کو طاقم نیم قطعی مخروطہ کہتے ہیں ۔ ایسے مخروطی کاملیاتی مسائل جن میں بندشیں نیم قطعی مخروطہ پر مشتمل ہوں، کو طاقم نیم قطعی کاملیت کہتے ہیں ۔

انتہائی پیچیدہ کاملیاتی مسائل کو مخروطی کاملیاتی مسائل کے طور پر پیش کیا جا سکتا ہے ۔ اس مقالہ میں ہم نے ایسے مربعی کاملیاتی مسائل جن میں بندشی تفاعل مربعی اصطلاحات پر مشتمل ہو اور ہدف تفاعل لکیری ہو کی طاقم نیم قطعی رعایتوں پر بحث کی ہے ۔ ہمارے پیش کردہ نتائج سے یہ اخذ کیا جا سکتا ہے کہ مربعی بندشوں پر مشتمل مربعی کاملیتی مسائل کو مکمل طور پر طاقم نیم قطعی رعایتوں کی صورت میں پیش کرنے کے لیے اضافی اصطلاحات کی ضرورت ہوتی ہے ۔

ریاضیاتی کاملیت کے مسائل کو متناہی اور لا متناہی مسائل میں تقسیم کیا جا سکتا ہے ۔ لا متناہی کاملیاتی مسائل کی ایک ذیلی قسم نیم لا متناہی کا ملیت کہلاتی ہے جس میں متغیرات کی تعداد متناہی ہوتی ہے جبکہ بندشوں کی تعداد لا متناہی ہوتی ہے ۔ اس مقالہ میں ہم نے درج ذیل لکیری نیم لا متناہی کاملیت کا استعمال کیا ہے ۔

$$\max_{\mathbf{x} \in \mathbb{R}^n} \ \mathbf{c}^T \mathbf{x} \qquad s.t. \quad b(\mathbf{z}) - a(\mathbf{z})^T \mathbf{x} \geq 0 \quad \forall \mathbf{z} \in Z,$$

مندرجہ بالا کلیہ میں $Z \subseteq \mathbb{R}^m$ اشاری طاقم کو ظاہر کرتا ہے جبکہ $a : Z \to \mathbb{R}^n$ اور $b : Z \to \mathbb{R}$ مسلسل تفاعل ہیں ۔

کسی قالب $Q$ کے مخلوط مثبتی قالب ہونے کی ایک متبادل شرط یہ ہے کہ تمام $\mathbf{v} \in \Delta m$ کے لیی، $\mathbf{v}^T Q\mathbf{v}$ کی قیمت غیر منفی ہو ۔ اس متبادل شرط کو استعمال کرتے ہوئے مخلوط مثبتی کاملیت کو لکیری نیم لا متناہی کاملیت کے طور پر پیش کیا جا سکتا ہے ۔ ہم نے اس تعلق کو مخلوط مثبتی کاملیت کی کاملیاتی شرائط اور دوہریت سے متعلقہ نتائج اخذ کرنے کے لیے استعمال کیا ہے ۔

اس تعلق کا ایک اور فائدہ یہ ہے کہ نیم متناہی لکیری کاملیت کو حل کرنے کے کسی بھی طریقہ کار کو مخلوط مثبتی کاملیت کے مسائل کو حل کرنے کے لیے استعمال کی جا سکتا ہے ۔ اس مقالہ میں مخلوط مثبتی کاملیت کے مسائل کو حل کرنے کے لئے موجود ایک طریقہ کار کا تجزیہ لکیری نیم لا متناہی کاملیت کے نکتہ نظر سے کیا گیا ہے ۔

# Acknowledgment

Writing these pages seems to me harder than I thought because last five years have been such an eventful journey that cannot be recapped in just two pages. I came to the Netherlands carrying the biggest dream of my life to get a PHD and now here I am, ready to turn it into a reality. Looking back to the whole journey, this dissertation would have not been possible without the guidance, support and encouragement of many individuals. Below I would like to mention a few who have contributed in many ways and those which I forgot to mention deserve a same level of gratitude.

Words cannot express how thankful and grateful, I am to my supervisor Dr. Georg Still for his patience and guidance, he provided, in last five years. He took great pain in correcting the manuscripts, I have submitted time to time. He is not only a good advisor for my PhD but his advices in my personal life also contributed a great deal of comfort. Special thanks goes to Prof. Dr. Marc Uetz for being the Promoter of this research study. I would like to extend my gratitude to the members of the DWMP group for providing a nice working environment.

I am also thankful to Jan Schut for his kindness and prompt support that simplified significantly my early stay in the Netherlands.

I am very grateful to friends which I made during my stay in Enschede (Saba, Ram, Niam, katja, Amir, Somayah, Matin, Andre and Hayhan), the memories of get togethers and Bar-BQ's are still fresh in my mind. I am also very thankful to Sulaiman and Sidra for all the 7:00 pm tea and chit chat. I am also thankful to Mohsin Raza and Irfan Zafar for their support and encouragement. I owe a great deal of gratitude to Zia-Ud-Din for providing guarantee to government of Pakistan, which was a mandatory requirement for the scholarship.

Then, I mention the people who are always very close to my heart. First among them, my heart-filled thank goes to my dear mother and father for being my inner spirit at every moment of my life and for the understanding that it is in best interest of my future to be away from them and stay in the Netherlands. Few words of thanks for anti and uncle (Mrs. Ishrat & Mr. Allah Bukhsh) for all the prayers and silent support. Then I write a sentence for my loving sisters

Guria who always there to guide me, Chanda as a fighter to refresh my mind. Special sorry/thanks for Saba, Zeeshan, Bilal, Jaree and Zaharah for bearing all the hard words, anger and BAYSTI. Thanks also goes to my elder brothers (Imran Bhai, Nauman Bhai, Adnan Bhai, Rehan Bhai) and sister-in-laws for all the encouragement and support.

Finally, special thanks goes to the three F (Faiza, Farzan and Faheema) of my family for understanding that I am at home and at work at the same time. They are the biggest strength of my life, for Faiza I have chosen a special way of saying thanks, which you can realize if you read first letter of each chapter of this dissertation. Some words for our ever loving little son- Farzan, who always says "Baba wokking, going offshe" and for the little doll- Faheema who lie beside my chair and wait for my attention. Every hard time, every tired day and unbearable pressure, are vanished with a single smile of these little faces.

**Faizan Ahmed**

# List of Notations

**Acronyms**

| | |
|---|---|
| COP | Copositive Programming, page 10 |
| ESS | Evolutionarily Stable Strategy, page 55 |
| LP | Linear Programming, page 7 |
| LSIP | Linear Semi-infinite Program, page 11 |
| SDP | Semidefinite Programming, page 8 |
| StQP | Standard Quadratic Programming, page 40 |

**Cone**

| | |
|---|---|
| $\mathcal{C}_m^*$ | the completely positive cone, page 5 |
| $\mathcal{C}_m$ | the copositive cone, page 4 |
| $\mathcal{C}_m(K)$ | the cone of set-semidefinite matrices, page 20 |
| $\mathcal{C}_m^*(K)$ | the dual of the cone of set-semidefinite matrices, page 20 |
| $\mathcal{N}_m$ | the cone of $m \times m$ symmetric, nonnegative matrices, page 6 |
| cone$(V)$ | the convex cone generated by $V \subseteq \mathbb{R}^m$, page 11 |
| $\mathcal{S}_m^+$ | the positive semidefinite cone , page 4 |
| $\mathcal{S}_m^{++}$ | the positive definite cone, page 4 |
| $\mathcal{S}_m$ | the cone of symmetric $m \times m$ matrices, page 4 |
| Ext$(K)$ | the set of elements of $K$ which generate extreme rays., page 28 |
| $DNN_m$ | the set of doubly nonnegative matrices, $DNN_m := \mathcal{S}_m^+ \cap \mathcal{N}_m$, page 31 |

**Matrix**

| | |
|---|---|
| | for the matrix $A$, $\mathbf{a}_i$ will denote the $i^{th}$ column of $A$, page 26 |
| adj$(A)$ | the adjoint of the matrix $A$, page 26 |

| | |
|---|---|
| det$(A)$ | the determinant of the matrix $A$, page 25 |
| Diag$(\mathbf{u})$ | is the matrix with **u** on the main diagonal while all other elements are zero, page 22 |
| diag$(A)$ | is the vector of the diagonal elements of the matrix $A$, page 22 |
| ker$(A)$ | the kernel or null space of the matrix, i.e., ker$(A) := \{\mathbf{v} : A\mathbf{v} = \mathbf{o}\}$, page 48 |
| $\|\|A\|\|$ | the Frobenius norm, i.e., $\|\|A\|\| := \sqrt{\text{tr}(AA^T)}$, page 51 |
| tr$(A)$ | for $A \in \mathbb{R}^{m\times m}$, tr$(A) := \sum_{i=1}^{m} a_{ii}$ , page 4 |
| $A^T$ | for $A \in \mathbb{R}^{m\times n}$, $A^T$ is the transpose of $A$ , page 4 |
| $A^{-1}$ | the inverse of the matrix $A$, page 26 |
| $A^{ij}$ | the matrix obtained from $A$ after deleting the $i^{th}$ row of $A$ and the $j^{th}$ column of $A$, page 25 |
| $a_{ij}$ | for $A \in \mathbb{R}^{m\times n}$, $a_{ij}$ is the element in $i^{th}$ row and $j^{th}$ column , page 4 |
| $E$ | usually the $m$-dimensional matrix of all ones, page 24 |
| $I$ | usually the $m$-dimensional identity matrix, page 24 |
| $O$ | the matrix of all zeros, the order of $O$ will be clear from the context, page 28 |
| $Q_J$ | the principal submatrix obtained after deleting the rows and the columns of the matrix $Q$ not corresponding to the elements of the index set $J \subseteq \mathcal{U}$ i.e. $(Q_J)_{ij} = q_{ij}$ for all $i \in J, j \in J$, page 26 |
| $rank(A)$ | rank of the matrix $A$, page 32 |

**Mathematical Programming**

| | |
|---|---|
| $(Cone_D)$ | the dual cone program, page 5 |
| $(Cone_P)$ | Primal cone program, page 5 |
| $(COP_D)$ | the dual copositive program, page 10 |
| $(COP_P)$ | the primal copositive program, page 10 |
| $(SIP_D)$ | the Haar dual of the linear semi-infinite program $(SIP_P)$, page 11 |
| $(SIP_P)$ | the linear semi-infinite program, page 11 |
| $\mathcal{F}(P)$ | set of feasible points of the program $(P)$, page 6 |
| $val(P)$ | value of the program $(P)$, page 6 |
| S(P) | the set of maximizers of the program $(P)$, page 12 |

**Miscellaneous**

| | |
|---|---|
| $\langle ., . \rangle$ | standard inner product, i.e., $\langle U, V \rangle = \text{tr}(U^T V)$ for $U, V \in \mathbb{R}^{m \times n}$, page 4 |
| $N_\epsilon(\bar{\mathbf{v}})$ | the $\varepsilon$-neighbourhood of $\bar{\mathbf{v}} \in \mathbb{R}^m$, i.e., $N_\varepsilon(\bar{\mathbf{v}}) := \{\mathbf{v} \in \mathbb{R}^m : \lvert\lvert\mathbf{v} - \bar{\mathbf{v}}\rvert\rvert \leq \varepsilon\}$, for $\varepsilon > 0$, page 42 |

**Set**

| | |
|---|---|
| conv $(S)$ | the convex hull of the set $S$, page 4 |
| $\mathcal{U}$ | $\mathcal{U} := \{1, 2, \cdots, m\}$, page 24 |
| int$(S)$ | interior of the set $S$, page 7 |
| $\lfloor a \rfloor$ | the largest integer less then or equal to $a$, page 65 |
| $\mathbb{R}$ | the real space, page 3 |
| $\mathbb{R}^m$ | the $m$-dimensional real space, page 3 |
| $\mathbb{R}^m_+$ | the nonnegative orthant, page 3 |
| $\mathbb{R}^m_{++}$ | $\mathbb{R}^m_{++} := \{\mathbf{b} \in \mathbb{R}^m : b_i > 0, \; \forall \, i = 1, \ldots, m\}$, page 34 |
| $\mathbb{R}^{m \times n}$ | the space of $m \times n$ real matrices, page 3 |
| $rint(S)$ | the relative interior of the set $S$, page 45 |
| aff$(S)$ | the affine hull of the set $S$, page 45 |

**Vector**

| | |
|---|---|
| **e** | the vector of all ones, usually $\mathbf{e} \in \mathbb{R}^m$, otherwise the dimension of **e** is clear from the context, page 24 |
| **o** | the vector of all-zeros, the dimension of **o** will be clear from the context, page 4 |
| $\mathbf{v}_J$ | the sub vector corresponding to the elements of the index set $J$, i.e, $(\mathbf{v}_J)_i = v_i$ for all $i \in J$, page 28 |
| $\lvert\lvert\mathbf{u}\rvert\rvert$ | the Euclidean norm of the vector $\mathbf{u} \in \mathbb{R}^m$ i.e. $\lvert\lvert\mathbf{u}\rvert\rvert := \sqrt{\sum_i^m u_i^2}$, page 42 |
| $e_i$ | the unit vectors of either length $m$, or the length is clear from the context, page 24 |
| $R(\mathbf{v})$ | support of the vector $\mathbf{v} \in \mathbb{R}^m$, page 41 |

# Index