Dissertation

Copositivity in Infinite Dimension

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaft (Dr. rer. nat.).

Dem Fachbereich IV - Mathematik
der Universität Trier vorgelegt

von

M.Sc. Claudia Adams

Trier, 2018

Eingereicht am:	17.12.2018
Disputation am:	11.02.2019
Gutachter:	Prof. Dr. Mirjam Dür
	Prof. Dr. Leonhard Frerick

Zusammenfassung

Viele kombinatorische Optimierungsprobleme können als konvexe Probleme über einem Kegel formuliert werden, beispielsweise das stabile Mengen Problem, das Cliquenproblem oder das Problem des maximalen Schnitts. Insbesondere NP-schwere Probleme können als copositive Optimierungsprobleme formuliert werden. Hierbei liegt die Schwierigkeit des neuen Problems vollständig in der Copositivitätsbedingung.
Copositive Optimierung ist ein vergleichsweise neues Thema in der Optimierung. Es behandelt die Optimierung über dem sogenannten copositiven Kegel, welcher eine Obermenge des positiv semidefiniten Kegels darstellt. Sein Dualkegel ist der Kegel der vollständig positiven Matrizen, in welchem alle Matrizen enthalten sind, die als Summe von nichtnegativen, symmetrischen Vektor-Vektor-Produkten zerlegt werden können. Die zugehörigen Optimierungsprobleme haben eine lineare Zielfunktion, lineare Nebenbedingungen in der Matrixvariable und eine Kegelbedingung. Manche Optimierungsprobleme können als kombinatorische Probleme über unendlich dimensionalen Graphen formuliert werden, wie zum Beispiel die Berechnung der sogennanten Kusszahl, welche als stabile Mengen Problem über der Sphäre beschrieben werden kann. In der vorliegenden Arbeit werden wir diskutieren, inwieweit das Konzept der Copositivität in einem unendlichdimensionalen Raum verallgemeinert werden kann. Für Spezialfälle werden wir Anwendungen in der kombinatorischen Optimierung präsentieren.

In Kapitel 2 geben wir zunächst eine Einführung in die Thematik der copositiven Optimierung im Endlichdimensionalen und präzisieren einige Anwendungen in der kombinatorischen Optimierung. Um diese Theorie in einem unendlichendimensionalen Raum zu verallgemeinern, benötigen wir Werkzeuge aus der Funktionalanalysis, denn anstelle von beispielsweise Matrizen wollen wir Operatoren in Hilberträumen betrachten. Die wichtigsten Aspekte hierzu werden wir in Kapitel 3 aufzeigen. Eine wichtige und verbreitete Norm in der Funktionalanalysis ist die (sym-

metrische) projektive Norm. In Kapitel 4 werden wir diese und einige ihrer Eigenschaften näher betrachten sowie ihre Bedeutung für unser Ziel, die Verallgemeinerung des Konzepts von Copositivität bzw. vollständiger Positivität im Unendlichdimensionalen, herausarbeiten. Mit Hilfe dieser Norm werden wir zwei besondere Fälle, welche für eine Verallgemeinerung von copositiver Optimierung in Frage kommen, betrachten. In Kapitel 5 analysieren wir den ersten dieser beiden Fälle. Dieser basiert auf [24], wo die Autoren eine Verallgemeinerung des copositiven Kegels als Kegel der copositiven Kerne über einem kompakten metrischen Raum vorstellten. Darüber hinaus wurden der Dualkegel, die zugehörigen Optimierungsprobleme sowie Aussagen zur Dualität formuliert und bewiesen. Außerdem wurden Anwendungen in der kombinatorischen Optimierung vorgestellt. Der zweite Fall, den die symmetrische projektive Norm liefert, führt zu einer anderen Verallgemeinerung von Copositivität, hier bezüglich eines Hilbertraumes. Diese neue Theorie wird in Kapitel 6 eingeführt. Im Rahmen dieser Thematik werden wir Eigenschaften diskutieren, die analog zum endlichdimensionalen Fall gelten, und Unterschiede zu diesem aufzeigen. Wir werden Verallgemeinerungen des copositiven und des vollständig positiven Kegels im L^2 sowie deren besondere Eigenschaften betrachten. Als zentraler Punkt wird hier die Darstellung des vollständig positiven Kegels mit Hilfe seiner Extremalstrahlen bewiesen.

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die zu der Entstehung dieser Arbeit beigetragen haben.

Mein erster Dank gilt meinen beiden Betreuern Prof. Dr. Mirjam Dür und Prof. Dr. Leonhard Frerick, nicht nur für die Möglichkeit diese Arbeit zu schreiben, sondern auch für ihre umfassende und fachlichliche Betreuung in den letzten Jahren. Prof. Dr. Mirjam Dür brachte mir die copositive Optimierung nahe und betreute sowohl meine Bachelor- als auch meine Masterarbeit. Trotz ihres Wechsels an die Universität Augsburg während meiner Promotionszeit war die sehr gute Betreuung und ein reger Austausch weiterin gegeben. Darüber hinaus, gab sie mir die Möglichkeit meine Forschung auf vielen Konferenzen im In- und Ausland zu präsentieren. Prof. Dr. Leonhard Frerick war stets offen für Fragen und Diskussionen jeglicher Form. Inbesondere in der zweiten Hälfte der Promotion, war er mein wichtiger Ansprechpartner vor Ort. Seine Kompetenz und Intuition für interessante Fälle und Eigenschaften in der Analysis waren ausschlaggebend für den Verlauf dieser Arbeit.

Weiter möchte ich der Deutschen Forschungsgemeinschaft und den Verantwortlichen des Graduiertenkollegs 2126 *Algorithmische Optimierung* (ALOP) an der Universität Trier für die finanzielle Unterstützung und die vielfältigen Möglichkeiten zur fachlichen Weiterbildung während meiner Promotionszeit danken.

Außerdem möchte ich mich bei meinen Kollegen und dem SIAM Student Chapter für die angenehme Arbeitsatmosphäre und hilfreiche Diskussionen bedanken. Mein Dank gilt insbesondere Oliver Hauke, der mir fachlich wie auch privat während unseres Studiums und der Promotion zur Seite stand. Ferner danke ich ihm, für seine hilfreichen Hinweise bei der Entstehung dieser Arbeit. Ein weiterer Dank geht an Dr. Duy Van Nguyen für die gemeinsame Zusammenarbeit sowie für seine Hilfe während meines Studiums und der Promotion, aber vor allem auch für die angenehme Zeit im Büro. Ich danke ihm auch für die Korrektur dieser Arbeit. Außerdem danke ich Bernd Perscheid für gute Gespräche, hilfreiche Ratschläge und offenen Ohren sowie für seine Korrektur.

Und natürlich gilt ein besonderer Dank meiner Familie. Meine Eltern begleiteten und unterstützten mich in jeglicher Form auf meinem gesamten Weg. Ein großer Dank geht an meine Schwester Anne, nicht nur für die Korrektur dieser Arbeit, sondern auch für ihre vielfältige und stetige Unterstützung in jeder Lebenslage. Ich danke Jonas, für seine Ermutigungen und engagierte Unterstützung.

Contents

Chapter 1

Introduction

Many combinatorial optimization problems on finite graphs can be formulated as conic convex programs, e.g. the stable set problem, the maximum clique problem or the maximum cut problem. Especially NP-hard problems can be written as copositive programs. In this case the complexity is moved entirely into the copositivity constraint. Copositive programming is a quite new topic in optimization. It deals with optimization over the so-called copositive cone, a superset of the positive semidefinite cone, where the quadratic form $x^\top Ax$ has to be nonnegative for only the nonnegative vectors x. Its dual cone is the cone of completely positive matrices, which includes all matrices that can be decomposed as a sum of nonnegative symmetric vector-vector-products. The related optimization problems are linear programs with matrix variables and cone constraints.
However, some optimization problems can be formulated as combinatorial problems on infinite graphs. For example, the kissing number problem can be formulated as a stable set problem on a circle. In this thesis we will discuss how the theory of copositive optimization can be lifted up to infinite dimension. For some special cases we will give applications in combinatorial optimization.

In Chapter 2 we give an introduction to the topic of copositive optimization in finite dimension and some applications in combinatorial optimization. To lift this theory to infinite dimension we need tools from functional analysis since we will study e.g. operators in Hilbert spaces instead of matrices. These tools will be described in Chapter 3.
An essential and popular norm in functional analysis is the (symmetric) projective norm. In Chapter 4 we will have a closer look at it and explain its importance for our theory in infinite dimension. With this norm we can

examine two special cases for a generalization of copositive optimization in infinite dimension.
The first case will be discussed in Chapter 5. This case is based on [24], where the authors gave an approach how to generalize the copositive cone as the cone of copositive kernels over a compact metric space. They also formulated the dual cone, the related optimization problems and duality statements. Beyond that they presented some applications in combinatorial optimization.
The second case that results from the symmetric projective norm leads to a different approach to generalize copositive optimization to an infinite dimensional space, in particular to a selfdual Hilbert space. This new theory will be presented in Chapter 6. In this context we will discuss some properties which are equivalent to the ones in finite dimension and we will also point out differences to finite dimension. We will consider the generalization of the copositive and completely positive cone, respectively, in the setting of L^2 and their special characteristics. As an important instrument we will prove the representation of the completely positive cone by its extreme rays.

Chapter 2

Basic facts on copositive optimization

In this chapter we will introduce the fundamental knowledge about copositive optimization. At first we will give a short motivation of this special kind of cone optimization and embed it in the theory. Furthermore we will give an overview of the current research in copositive optimization. Last but not least we will consider some applications of copositive optimization in finite dimension.

2.1 Motivation

Copositive optimization is a quite new topic in mathematical optimization. It considers linear optimization problems over the cone of the so-called copositive matrices.
Many combinatorial optimization problems on finite graphs can be formulated as conic convex programs (e.g. the stable set problem, the maximum clique problem, the maximum cut problem). Especially NP-hard problems can be written as copositive programs. In this case the complexity is moved entirely into the cone constraint. Copositive programs are useful not only in combinatorial but also in quadratic optimization. Since checking copositivity is NP-hard, we have to mention that there exist different approaches to approximate the copositive cone, for example by sum-of-squares decomposition [42] or nonnegativity of the coefficients of a special polynomial [16]. These approximations lead to linear or semidefinite programs. More particularly with an algorithm via inner and outer approximations [11] a

copositive program can be approximated by a sequence of linear programs. A comprehensive overview over the topic of copositive optimization can be found in [7, 13, 25].

2.2 Preliminaries

In this section we will introduce the copositive and the completely positive cone and some of their properties which will be relevant later in this thesis. In this context, we will also have a look at the cone of positive semidefinite matrices. Furthermore, we will consider the related optimization problems over these cones.

But first we will discuss the notation. Although in this chapter we consider the theory in finite dimension, the notation will already now be geared to the infinite dimensional case. Therefore we introduce now the notation of tensors and tensor products from an algebraic viewpoint. For two vector spaces X, Y the *tensor product* can be constructed as a space of linear functionals on the vector space $B(X \times Y)$ of bilinear maps on $X \times Y$. For $x \in X$ and $y \in Y$ we denote by $x \otimes y$ the functional given by evaluation at the point (x, y), i.e.

$$(x \otimes y)(A) = \langle A, x \otimes y \rangle = A(x, y),$$

for each bilinear form A on $X \times Y$. The tensor product $X \otimes Y$ is a subspace of the dual $B(X \times Y)'$, which includes all linear functionals on $B(X \times Y)$, spanned by their elements. Therefore a typical *tensor*, i.e. an element of the tensor product $X \otimes Y$, has the form

$$u = \sum_{i=1}^{k} \lambda_i x_i \otimes y_i,$$

where k is a positive integer, $\lambda_i \in \mathbb{R}$, $x_i \in X$ and $y_i \in Y$ for $i = 1, \ldots, k$. Note that the representation of u is in general not unique. This description and more details about it can be found in [48, Chapter 1].

Moreover for a positive integer n we will denote by $\mathbb{R}^n$ the n-dimensional real space and by $\mathbb{R}^n_+$ the nonnegative orthant in $\mathbb{R}^n$. An element $a \in \mathbb{R}^n$ is a map $a\colon \{1, \ldots, n\} \to \mathbb{R}$. By $a(i)$ we will mark the i-th entry of the vector a. For two positive integers m, n we will denote the space of $m \times n$ matrices by $\mathbb{R}^{m\times n}$, which is equal to $\mathbb{R}^m \otimes \mathbb{R}^n$ in the tensor notation. A matrix A with

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

is a map $A\colon \{1,\dots,m\}\times\{1,\dots,n\} \to \mathbb{R}$. By $A(i,j)$ we will denote the (i,j)-th entry of A. For $a,b \in \mathbb{R}^n$ the product $ab^\top \in \mathbb{R}^{n\times n}$ is in tensor notation equal to $a\otimes b \in \mathbb{R}^n\otimes\mathbb{R}^n$. This product describes a map $a\otimes b\colon \{1,\dots,n\}\times\{1,\dots,n\} \to \mathbb{R}$ with $(a\otimes b)(i,j) = a(i)b(j)$. If it fits with the context, we will use the tensor notation already in this chapter.

In this thesis we will only consider symmetric matrices or operators. Let $\mathcal{S}_n$ be the set of symmetric $n\times n$ matrices. Then we have the following definition:

Definition 2.1. The *copositive cone* $\mathcal{COP}_n$ is defined as

$$\mathcal{COP}_n := \{A \in \mathcal{S}_n \colon x^\top A x \geq 0 \text{ for all } x \in \mathbb{R}^n_+\}.$$

In the context of tensors the condition $x^\top Ax \geq 0$ is equivalent to the condition $\langle A, x\otimes x\rangle \geq 0$, where

$$\langle A, B\rangle := \operatorname{trace}(B^\top A) = \sum_{i,j=1}^{n} A(i,j)B(i,j)$$

denotes the inner product of two matrices. Since we only consider symmetric matrices in this thesis, the inner product is equal to $\langle A, B\rangle := \operatorname{trace}(BA)$.

Definition 2.2. For an arbitrary given cone $\mathcal{K} \subseteq \mathcal{S}_n$ the *dual cone* $\mathcal{K}^*$ is defined as follows:

$$\mathcal{K}^* := \{A \in \mathcal{S}_n \colon \langle A, B\rangle \geq 0 \text{ for all } B \in \mathcal{K}\}.$$

With this definition it can be shown, cf. [5, Theorem 2.3], that the dual cone of the copositive cone is the cone $\mathcal{CP}_n$ of *completely positive matrices*:

$$\mathcal{COP}_n^* = \mathcal{CP}_n := \operatorname{conv}\{x\otimes x \colon x \in \mathbb{R}^n_+\}, \tag{2.1}$$

where $\operatorname{conv}(M)$ denotes the convex hull of a set $M \subseteq \mathbb{R}^n$. Furthermore, the dual cone of $\mathcal{CP}_n$ is again $\mathcal{COP}_n$, i.e.

$$\mathcal{COP}_n^{**} = \mathcal{CP}_n^* = \mathcal{COP}_n,$$

cf. [5, Theorem 2.3].

Proposition 2.3. *$\mathcal{COP}_n$ and $\mathcal{CP}_n$ are both pointed closed convex cones with nonempty interior.*

For a proof of Proposition 2.3 see [5, Proposition 1.24, Theorem 2.2].

In this context we also have to mention that testing, whether a matrix is copositive or not, is a co-NP-complete decision problem. This means that checking copositivity is NP-hard, but if the to be examined matrix is not copositive then there exists a certificate for this that can be checked in polynomial time, cf. [21, 41]. Intuitively this should also hold for testing if a matrix is completely positive, but a formal proof for this assumption does not exist so far. These issues are discussed in [1], [22, Theorem 5.5] and [41, Theorem 3].

Another important cone in this context is the *positive semidefinite cone*:

$$\begin{aligned} \mathcal{S}_n^+ &:= \{A \in \mathcal{S}_n \colon x^\top Ax \geq 0 \ \text{ for all } x \in \mathbb{R}^n\} \\ &= \operatorname{conv}\{x \otimes x \colon x \in \mathbb{R}^n\} \\ &= (\mathcal{S}_n^+)^* . \end{aligned}$$

Indeed the equations hold by the selfduality of $\mathcal{S}_n^+$, cf. [32, Lemma 1.2.6].

Furthermore, the following relations hold true:

$$\mathcal{CP}_n \subseteq \mathcal{S}_n^+ \quad \text{and} \quad \mathcal{S}_n^+ \subseteq \mathcal{COP}_n,$$

cf. [5, Remark 1.10, Remark 2.4].

Like the copositive and the completely positive cone, the positive semidefinite cone is full-dimensional, closed, convex, pointed and has a nonempty interior, cf. [5, Proposition 1.21]. If we consider in addition the cone of entrywise nonnegative matrices $\mathcal{N}_n$, the following inclusions hold true:

$$\mathcal{CP}_n \subseteq \mathcal{S}_n^+ \cap \mathcal{N}_n \quad \text{and} \quad \mathcal{S}_n^+ + \mathcal{N}_n \subseteq \mathcal{COP}_n. \tag{2.2}$$

Matrices in $\mathcal{S}_n^+ \cap \mathcal{N}_n$ are called *doubly nonnegative*. For $n \leq 4$ equality holds in both inclusions and for $n \geq 5$ the inclusions are strict, i.e. not every copositive matrix can be written as the sum of a positive semidefinite and a symmetric nonnegative matrix, cf. [5, Remark 1.10, Remark 2.4]. A special example for this is the so-called 5×5 Horn-matrix, which is copositive, but neither nonnegative, nor positive semidefinite and it cannot be represented as a sum of a positive semidefinite and a symmetric nonnegative matrix. A proof of this can be found in [5, Example 1.30].

An essential tool in conic optimization is the following well-known result:

Lemma 2.4. *A closed convex cone $\mathcal{K} \subseteq \mathbb{R}^n$ is pointed if and only if its dual cone $\mathcal{K}^*$ has a nonempty interior.*

Proof. At first we show the sufficient part: Let $\bar{y} \in \operatorname{int}(\mathcal{K}^*)$, then $\langle x, \bar{y} \rangle > 0$ for all $0 \neq x \in \mathcal{K}$ because the interior of $\mathcal{K}^*$ can be characterized as follows:

$$\begin{aligned}\operatorname{int}(\mathcal{K}^*) &= \mathbb{R}^n \setminus \operatorname{cl}\left((\mathcal{K}^*)^c\right)\\ &= \mathbb{R}^n \setminus \operatorname{cl}\left(\{y \in \mathbb{R}^n \colon \exists x \in \mathcal{K} \colon \langle x, y \rangle < 0\}\right)\\ &= \mathbb{R}^n \setminus \{y \in \mathbb{R}^n \colon \exists\, 0 \neq x \in \mathcal{K} \colon \langle x, y \rangle \leq 0\}\\ &= \{y \in \mathbb{R}^n \colon \langle x, y \rangle > 0 \ \forall 0 \neq x \in \mathcal{K}\}.\end{aligned}$$

Assume now that $\mathcal{K}$ is not pointed. Then there exists $0 \neq \bar{x} \in \mathcal{K} \cap (-\mathcal{K})$ with $\bar{x} \in \mathcal{K}$ and $-\bar{x} \in \mathcal{K}$. Together we have:

$$\begin{aligned}&\langle \bar{x}, \bar{y} \rangle > 0, \text{ since } \bar{x} \in \mathcal{K} \text{ and } \bar{y} \in \operatorname{int}(\mathcal{K}^*),\\ &\langle -\bar{x}, \bar{y} \rangle > 0, \text{ since } -\bar{x} \in \mathcal{K} \text{ and } \bar{y} \in \operatorname{int}(\mathcal{K}^*).\end{aligned}$$

This is a contradiction and therefore $\mathcal{K}$ is pointed.
Now we prove the necessity: Let $\mathcal{K}$ be pointed. Since $\mathcal{K}$ is closed and convex, there exists a supporting hyperplane $H = \{x \in \mathbb{R}^n \colon \langle x, \bar{y} \rangle = 0\}$ with $H \cap \mathcal{K} = \{0\}$ and $\mathcal{K} \subseteq H^+ = \{x \in \mathbb{R}^n \colon \langle x, \bar{y} \rangle \geq 0\}$. Therefore we have $\langle x, \bar{y} \rangle \geq 0$ for all $x \in \mathcal{K}$ and hence $\bar{y} \in \mathcal{K}^*$. Since $H \cap \mathcal{K} = \{0\}$, it follows that $\langle x, \bar{y} \rangle = 0 \Leftrightarrow x = 0$. Hence, we get $\langle x, \bar{y} \rangle > 0$ for all $0 \neq x \in \mathcal{K}$. Therefore $\bar{y} \in \operatorname{int}(\mathcal{K}^*)$ and hence $\operatorname{int}(\mathcal{K}^*) \neq \emptyset$. ◻

It is well-known that the interior of the copositive cone $\mathcal{COP}_n$ is the *cone of strictly copositive matrices*:

$$\begin{aligned}\operatorname{int}(\mathcal{COP}_n) &= \{A \in \mathcal{S}_n \colon x^\top A x > 0 \ \text{ for all } x \geq 0, x \neq 0\}\\ &= \{A \in \mathcal{S}_n \colon \langle A, x \otimes x \rangle > 0 \ \text{ for all } x \geq 0, x \neq 0\},\end{aligned}$$

cf. [10, Lemma 2.3].

The following characterization of the interior of $\mathcal{CP}_n$ has been proven in [20, Theorem 7.4]:

$$\operatorname{int}(\mathcal{CP}_n) = \{BB^\top \colon \operatorname{rank}(B) = n, B > 0\}.$$

The interior points of $\mathcal{COP}_n$ and $\mathcal{CP}_n$ are important instruments to show strong duality for optimization problems. In these statements a strictly feasible point is necessary, cf. Theorem 2.8.

Other important tools in this context are the extreme rays of $\mathcal{COP}_n$ and $\mathcal{CP}_n$.

Definition 2.5. Let $\mathcal{K}$ be a convex cone and $x \in \mathcal{K}$. Then x is called *extreme* in $\mathcal{K}$ if for every decomposition $x = y+z$ with $y, z \in \mathcal{K}$ and $y \neq 0 \neq z$ holds: $y = \alpha x$ and $z = \beta x$ with $\alpha, \beta \geq 0$. The ray $\{\alpha x \colon \alpha \geq 0\}$ generated by x is called *extreme ray* of $\mathcal{K}$. We denote the set of extreme rays of $\mathcal{K}$ by $\mathrm{Ext}(\mathcal{K})$.

For $n \leq 4$ the extreme rays of $\mathcal{COP}_n$ are the extreme rays of $\mathcal{S}_n^+ + \mathcal{N}_n$, which are generated by $e_i \otimes e_j$ for $i, j = 1, \dots, n$ and by $a \otimes a$ with $a \in \mathbb{R}^n$ having positive and negative entries, cf. [31, Theorem 3.2]. For $n = 5$ for example the Horn-matrix is extremal for $\mathcal{COP}_n$. Unfortunately for $n > 5$ a full characterization of the extreme rays of the copositive cone $\mathcal{COP}_n$ is still an open problem. Some examples of extreme rays are listed in the following theorem:

Theorem 2.6. *[20, Theorem 8.20] For $n \geq 2$ the following results concerning the extreme rays of the copositive cone $\mathcal{COP}_n$ hold:*

1. *$\alpha E_{ij} \in \mathrm{Ext}\,(\mathcal{COP}_n)$, where E_{ij} denotes the matrix with all entries equal to 0, except $E(i, j) = E(j, i) = 1$ for $i, j = 1, \dots, n$ and $\alpha > 0$.*
2. *$x \otimes x \in \mathrm{Ext}\,(\mathcal{COP}_n)$, where $x \in \mathbb{R}^n \setminus (\mathbb{R}^n_+ \cup (-\mathbb{R}^n_+))$.*
3. *Properties, when $A \in \mathrm{Ext}\,(\mathcal{COP}_n)$ if $A(i, j) \in \{-1, 0, +1\}$ and $A(i, i) = +1$ for all $i, j = 1, \dots, n$ can be found in [33].*
4. *$A \in \mathrm{Ext}\,(\mathcal{COP}_n)$ if and only if $PDADP^\top \in \mathrm{Ext}\,(\mathcal{COP}_n)$, where P is a permutation matrix and D is a diagonal matrix such that $D(i, i) > 0$ for all $i = 1, \dots, n$.*
5. *For $M \in \mathcal{COP}_n \setminus \{0\}$, $B \in \mathbb{R}^{n \times m}$ we have $\begin{pmatrix} M & B \\ B^\top & 0 \end{pmatrix} \in \mathrm{Ext}\,(\mathcal{COP}_{n+m})$ if and only if $B = 0$ and $M \in \mathrm{Ext}\,(\mathcal{COP}_n)$.*

For a comprehensive overview of the extreme rays of $\mathcal{COP}_n$ we refer to [20, Chapter 8.3] and the references given therein.

In contrast to $\mathcal{COP}_n$, the sets $\mathrm{Ext}(\mathcal{S}_n^+)$ and $\mathrm{Ext}(\mathcal{CP}_n)$ are exactly known. The extreme rays of the positive semidefinite cone $\mathcal{S}_n^+$ are generated by the symmetric rank-one matrices:

$$\mathrm{Ext}\,(\mathcal{S}_n^+) = \{x \otimes x \colon x \in \mathbb{R}^n\}\,,$$

cf. [5, Proposition 1.21].

The extreme rays of the completely positive cone are generated by the nonnegative symmetric rank-one matrices:

$$\operatorname{Ext}(\mathcal{CP}_n) = \{x \otimes x \colon x \geq 0\}\,. \tag{2.3}$$

A proof of this can be found in [5, Proposition 2.1 and Remark 2.3]. Note here that the extreme rays of $\mathcal{CP}_n$ are not only pleasant objects in theory, they will be also very useful for the reformulation of the maximum clique problem, which we will discuss in the next subsection.

If we consider the copositive program as the primal optimization problem, it has the following form:

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

Its dual is the following completely positive program, which results from the usual Lagrangian approach.

$$\begin{aligned} \max\ & \sum_{i=1}^m b_i y_i \\ \text{s.t. }\ & C - \sum_{i=1}^m y_i A_i \in \mathcal{CP}_n \\ & y_i \in \mathbb{R} \quad (i = 1, \dots, m). \end{aligned} \tag{D}$$

The following property is known as weak duality:

Theorem 2.7 (Weak Duality)**.** *[3, IV. (6.2) Theorem 1.] Let X be feasible for (P) and $y = (y_1, \dots, y_m)$ be feasible for (D). Then we have:*

$$\sum_{i=1}^m b_i y_i \leq \langle C, X\rangle.$$

For X feasible for (P) and y feasible for (D) the difference $\langle C, X\rangle - \sum_{i=1}^m b_i y_i$ is called *duality gap*. It is easy to see that if $\langle C, \bar{X}\rangle - \sum_{i=1}^m b_i \bar{y}_i = 0$ and $\bar{X}$ feasible for (P) and $\bar{y}$ feasible for (D), then $\bar{X}$ is optimal for (P) and $\bar{y}$ is optimal for (D). Furthermore we call X *strictly feasible* for (P) if X is feasible for (P) and $X \in \operatorname{int}(\mathcal{COP}_n)$. Analogously, y is called *strictly feasible* for (D) if y is feasible for (D) and $C - \sum\limits_{i=1}^m y_i A_i \in \operatorname{int}(\mathcal{CP}_n)$.

A stronger duality statement can be found in [16, Theorem 1.1] for general conic optimization problems. We will transfer it to our copositive (primal) and completely positive (dual) programs:

Theorem 2.8 (Strong Duality)**.** *Let*

$$p^* := \inf\{\langle C, X\rangle : \langle A_i, X\rangle = b_i \,(i = 1, \ldots, m), X \in \mathcal{COP}_n\}$$
$$d^* := \sup\left\{\sum_{i=1}^m b_i y_i : C - \sum_{i=1}^m y_i A_i \in \mathcal{CP}_n\right\}.$$

If there exist a strictly feasible solution X^0 of (P) and a feasible solution of (D), then $p^* = d^*$ and the maximum in (D) is attained. Similarly, if there exist a strictly feasible y^0 for (D) and a feasible solution of (P), then $p^* = d^*$ and the minimum in (P) is attained.

A last thing that we have to mention in this chapter, is the relation between copositive and quadratic programs, cf. [9]. For this, we first consider the standard quadratic problem:

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

where $Q \in \mathcal{S}_n$ and e denotes the all-ones vector. The main difficulty of this kind of problems is that the quadratic objective function needs not be convex. Note that the standard quadratic problem (STQP) is NP-hard since many combinatorial optimization problems can be reduced to it, e.g. the maximum clique problem.

By some transformations we can reformulate the (STQP) as a completely positive program with a linear objective. The objective function can be rewritten as: $x^\top Q x = \langle Q, x \otimes x\rangle =: \langle Q, X\rangle$. With $e \otimes e =: E$ the constraint $e^\top x = 1$ results in the constraint $\langle E, x \otimes x\rangle = \langle E, X\rangle$. With these transformations we get the following relaxation of the standard quadratic problem:

$$\begin{aligned} \min\ & \langle Q, X\rangle \\ \text{s.t. } & \langle E, X\rangle = 1 \\ & X \in \mathcal{CP}_n. \end{aligned} \tag{2.4}$$

Now we have a linear objective function, therefore an optimal solution has to be attained in an extreme point of the convex feasible set. As mentioned in (2.3), the extreme points of $\mathcal{CP}_n$ are exactly the rank-one matrices $x \otimes x$ with $x \geq 0$. With the second condition $e^\top x = 1$ it follows that the reformulation (2.4) is equivalent to (STQP). Since the objective function in (2.4) is linear, an optimal solution of this optimization problem is attained in an extreme point of its feasible set. Hence in the next lemma we will have a closer look at the extreme rays of the feasible set of (2.4). The lemma and also its proof can be found in [9, Lemma 5].

Lemma 2.9. *The extremal points of the feasible set of (2.4) are exactly the rank-one matrices $X = x \otimes x$ with $x \in \Delta := \{x \in \mathbb{R}^n_+ : e^\top x = 1\}$.*

Proof. We will prove the sufficient direction first. For this purpose we denote $\mathcal{M} := \{X \in \mathcal{CP}_n : \langle E, X\rangle = 1\}$. Then every $X := x \otimes x$ with $x \in \Delta$ is a member of $\mathcal{M}$. Now we suppose that $x \otimes x$ with $x \in \Delta$ can be represented as a proper convex combination of two points of $\mathcal{M}$, i.e. there exist $Z, U \in \mathcal{M}$ and $\lambda \in (0, 1)$ such that $x \otimes x = (1 - \lambda)U + \lambda Z$. Furthermore we choose an orthogonal basis $\{x_1, x_2, \dots, x_n\}$ of $\mathbb{R}^n$ with $x = x_n$. Since $\mathcal{M} \subset \mathcal{CP}_n \subset \mathcal{S}^+_n$ we have $Z, U \in \mathcal{S}^+_n$. Consider further

$$\begin{aligned} 0 = (x_i^\top x)^2 &= (x_i^\top x)(x^\top x_i) \\ &= x_i^\top \left((1-\lambda)U + \lambda Z\right) x_i \\ &= (1-\lambda) x_i^\top U x_i + \lambda x_i^\top Z x_i. \end{aligned}$$

Then $x_i^\top U x_i = x_i^\top Z x_i = 0$ for all $i < n$ since $U, Z \in \mathcal{S}^+_n$. Thus U and Z have rank one. Since $U, Z \in \mathcal{M} \subset \mathcal{CP}_n$, there exists a rank-one decomposition $U = u \otimes u$ and $Z = z \otimes z$ for some $u, z \in \mathbb{R}^n_+$. Accordingly $x_i^\top u = x_i^\top z = 0$ for all $i < n$ and so U and Z have to be positive multiples of $x \otimes x$, since x_i is an element of the orthogonal basis. Moreover since $U, Z \in \mathcal{M}$ the condition $\langle E, U\rangle = \langle E, Z\rangle = 1$ presents that $U = Z = x \otimes x$. Hence $X = x \otimes x$ can only be represented as a combination of itself and so it is an extreme point.
Now we will prove the necessity. For this we assume that X is an extremal point of $\mathcal{M} \subset \mathcal{CP}_n$. Then $X = \sum_{i=1}^{d+1} \lambda_i x_i \otimes x_i$ with $x_i \in \mathbb{R}^n_+ \setminus \{0\}$ and $\lambda_i \geq 0$ for all i with $\sum_{i=1}^{d+1} \lambda_i = 1$ and $d = \binom{n+1}{2}$. Since $X \in \mathcal{M}$, we get

$$\begin{aligned} 1 &= \langle E, X\rangle \\ &= \left\langle e \otimes e, \sum_{i=1}^{d+1} \lambda_i x_i \otimes x_i \right\rangle \\ &= \sum_{i=1}^{d+1} \lambda_i (e^\top x_i)^2, \end{aligned}$$

where $e^\top x_i > 0$ for all i since $x_i \neq 0$. Now put $u_i := (e^\top x_i)^{-1} x_i \in \Delta$ such that $U_i := u_i \otimes u_i \in \mathcal{M}$ for all i. Then

$$X = \sum_{i=1}^{d+1} \lambda_i (e^\top x_i)^2 U_i,$$

which is of course a convex combination of matrices U_i in $\mathcal{M}$. From the assumption that X is an extremal point of $\mathcal{M}$ we get that $X = U_1$ is of the desired form. □

With the property from Lemma 2.9 it is possible to construct a solution of (STQP) from a solution of (2.4) with the same optimal value.

2.3 Applications

In this section we will discuss some applications of copositive optimization, in particular in combinatorial optimization. First we will consider a generalization of (2.4). It is a pleasant result that not only standard quadratic problems can be formulated as completely positive programs, but also quadratic programs of a more general form.

Binary quadratic problems

In [14, Chapter 2] Burer showed this generalization: Every quadratic program with linear and binary constraints can be reformulated as a completely positive optimization problem. In particular he showed that the program

$$\begin{aligned} \min\ & \langle Q, x \otimes x \rangle + 2c^\top x \\ \text{s.t. } & a_i^\top x = b_i \quad (i = 1, \ldots, m) \\ & x \geq 0 \\ & x(j) \in \{0, 1\} \quad (j \in J) \end{aligned} \tag{2.5}$$

where $x \in \mathbb{R}^n_+$ and $J \subseteq \{1, \ldots, n\}$ can be transformed into the following completely positive program:

$$\begin{aligned}
&\min \langle Q, X\rangle + 2c^\top x \\
&\text{s.t. } a_i^\top x = b_i \quad (i = 1, \dots, m) \\
&\qquad \langle a_i \otimes a_i, X\rangle = b_i^2 \quad (i = 1, \dots, m) \\
&\qquad x(j) = X(j, j) \quad (j \in J) \\
&\qquad \begin{pmatrix} 1 & x^\top \\ x & X \end{pmatrix} \in \mathcal{CP}_{n+1}.
\end{aligned}$$

This reformulation works only if system (2.5) fulfills the so-called *key condition*, i.e. $a_i^\top x = b_i$ for $i = 1, \dots, m$ and $x \geq 0$ implies $x(j) \leq 1$ for all $j \in J$. Burer also showed that the key condition can be enforced without loss of generality.

Next we will describe how a special NP-hard combinatorial optimization problem can be transferred into a copositive program. In this context we will consider in particular the maximum stable set problem and the maximum clique problem.

The maximum stable set problem

In [16] de Klerk and Pasechnik showed that the maximum stable set problem can be transformed into a conic convex optimization problem over the copositive cone. Unfortunately the copositive program remains NP-hard.

Definition 2.10. Let $G = (V, E)$ be a finite, undirected and simple graph with set of vertices V and set of edges $E \subseteq V \times V$. A set of vertices $S \subseteq V$ is called a *stable set* if for every pair $i, j \in S$ holds $\{i, j\} \notin E$, where $\{i, j\}$ denotes the edge between i and j.

With this definition we can formulate the *maximum stable set problem*:

$$\begin{aligned}
\alpha(G) = \max |S| \\
\text{s.t. } S \text{ is a stable set in } G,
\end{aligned}$$

i.e. finding a stable set S in G of maximal cardinality $|S|$. $\alpha(G)$ is called the *stability number* of G.

If we consider the following graph $\tilde{G}$

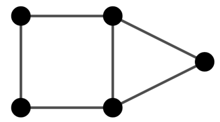


Figure 2.1: Example graph $\tilde{G}$

then a maximum stable set in $\tilde{G}$ is given as follows:

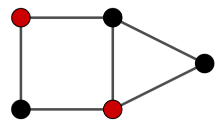


Figure 2.2: Maximum stable set of $\tilde{G}$

De Klerk and Pasechnik proved in [16, Theorem 2.2] that the stability number can be computed by a completely positive program:

Theorem 2.11. *Let $G = (V, E)$ be given with $|V| = n$. Then the stability number of G is given by*

$$\begin{aligned} \alpha(G) = \max\, & \langle e \otimes e, X \rangle \\ \textit{s.t. }\; & \langle I, X \rangle = 1 \\ & X(i, j) = 0 \;\; \textit{for all} \;\; \{i, j\} \in E,\; i \neq j \\ & X \in \mathcal{CP}_n, \end{aligned} \tag{2.6}$$

where e denotes again the all-ones vector and I is the identity matrix.

Proof. Define $C_G := \{X \in \mathcal{CP}_n : X(i, j) = 0, \{i, j\} \in E\}$, which is a convex cone. As written in (2.3) the extreme rays of $\mathcal{CP}_n$ are of the form $x \otimes x$ for nonnegative $x \in \mathbb{R}^n$. Therefore the extreme rays of C_G are of the form $x \otimes x$, where $x \in \mathbb{R}^n$ is also nonnegative and where its support corresponds to a stable set of G. Hence the extreme rays of the feasible set in (2.6) are given by the intersection of the extreme rays of C_G with the hyperplane defined by $\langle I, X \rangle = 1$.
Since $\langle e \otimes e, X \rangle$ is a linear map, the optimal value of this map is attained in an extreme point of the feasible set. Therefore there is an optimal solution of the form:

$$X^* = x^* \otimes x^*, \quad x^* \in \mathbb{R}^n, \quad x^* \geq 0, \quad \|x^*\|_2 = 1,$$

and where the support of x^* corresponds to a stable set of G, denoted by S^*. We denote now the optimal value of the objective function in (2.6) by $\lambda := \max \langle e \otimes e, X^* \rangle$. Then:

$$\lambda = \max \left\{ (e^\top x)^2 : \; \|x\|_2 = 1, x \geq 0, \mathrm{supp}(x) = \mathrm{supp}(x^*) \right\}.$$

The optimality conditions of this problem imply

$$x^* = \frac{1}{\sqrt{|S^*|}} x_{S^*},$$

where x_{S^*} denotes the incidence vector of the stable set S^*, i.e. $(x_{S^*})(i) = 1$ if $i \in S^*$ and $(x_{S^*})(i) = 0$ otherwise. Therefore we have

$$\lambda = (e^\top x^*)^2 = \frac{|S^*|^2}{|S^*|} = |S^*|.$$

Hence S^* must be a maximum stable set and hence $\lambda = \alpha(G)$. □

Remark 2.12. An essential tool in the proof of Theorem 2.11 is the set of extreme rays of $\mathcal{CP}_n$. Without their characterization the proof of the reformulation (2.6) of the maximum stable set problem as completely positive program would not work.
Another relevant aspect in this context is that there exists an optimal solution X^* with $X^* = x^* \otimes x^*$ and that $\mathrm{supp}(x^*)$ corresponds to the vertices of the maximum stable set. This follows from the proof of Theorem 2.11.

The dual (copositive) program of (2.6) is the following:

$$\begin{aligned} &\min t \\ &\text{s.t. } t \in \mathbb{R},\, K \in \mathcal{COP}_n \\ &\qquad K(i,i) = t - 1 \text{ for all } i \in V \\ &\qquad K(i,j) = -1 \text{ for all } \{i,j\} \notin E. \end{aligned} \tag{2.7}$$

Although the copositive formulation (2.7) of the maximum stable set problem remains NP-hard, this formulation has the advantage that the complexity is moved entirely into the copositivity constraint. The new formulation is a convex optimization problem with linear objective, linear constraints and one cone constraint.

The maximum clique problem

Another well-known NP-hard problem in combinatorial optimization is the maximum clique problem. This problem is equivalent to the maximum stable set problem in the complementary graph. The complementary graph $\bar{G}$ of a graph G is defined by the same set of vertices as G and two different vertices in $\bar{G}$ are connected by an edge if and only if they were not connected in G. For a survey of the maximum clique problem see [8]. In this section we will show that this problem can be reformulated as a copositive and completely positive program, respectively.

Definition 2.13. Let $G = (V, E)$ be a finite, undirected and simple graph with set of vertices V and set of edges $E \subseteq V \times V$. A *clique* of G is a subset C of vertices such that every pair of vertices in C is joined by an edge.

With this definition the *maximum clique problem* is given by:

$$\begin{aligned} \omega(G) = \max\, & |C| \\ \text{s.t. } & C \text{ is a clique in } G, \end{aligned}$$

i.e. finding a clique set C in G of maximal cardinality $|C|$. $\omega(G)$ is called the *clique number* of G.

A maximum clique of the graph $\tilde{G}$ in Figure 2.1 is the following:

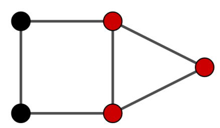


Figure 2.3: Maximum clique of $\tilde{G}$

We denote now by A_G the adjacency matrix of a graph G. Motzkin and Straus [40] showed the following equality:

$$\frac{1}{\omega(G)} = \min\left\{x^\top (e \otimes e - A_G)x \colon e^\top x = 1, x \geq 0\right\}, \tag{2.8}$$

where e denotes the all-ones vector as before. If we combine the Motzkin-Straus formulation (2.8) with the completely positive formulation (2.4), we get:

$$\frac{1}{\omega(G)} = \min\left\{\langle e \otimes e - A_G, X\rangle : \langle e \otimes e, X\rangle = 1, X \in \mathcal{CP}_n\right\}.$$

Then its dual is:

$$\frac{1}{\omega(G)} = \max\left\{\lambda\colon \lambda(e\otimes e - A_G) - e\otimes e \in \mathcal{COP}_n\right\},$$

cf. [25, Chapter 2]. In particular strong duality holds. As in the maximum stable set reformulation, the complexity of the reformulation is entirely in the cone constraint since we have a convex program with a linear objective, linear constraints and the cone constraint. Furthermore a discrete program changed to a continuous one.

2.4 Approximation hierarchies

In the last part of this chapter we will discuss some approximation hierarchies of the copositive cone. Since checking whether a matrix is copositive is NP-hard it is important to think about an easier way to check copositivity by approximating the copositive cone. The easiest way to approximating the copositive cone is of course to replace it by the positive semidefinite cone. Here we will present three popular approximation techniques, which lead to approximation hierarchies.

The first approximation is based on the following idea: By definition a matrix $A \in \mathbb{R}^{n\times n}$ is copositive if and only if the quadratic form $z^\top Az$ is nonnegative for all nonnegative arguments z. Equivalently for a given symmetric matrix A and $x \in \mathbb{R}^n$ the following polynomial can be considered:

$$P_A(x) := \sum_{i=1}^{n}\sum_{j=1}^{n} A(i,j)x(i)^2x(j)^2.$$

In conclusion A is copositive if and only if $P_A(x) \geq 0$ for all $x \in \mathbb{R}^n$. A sufficient condition for this nonnegativity is that $P_A(x)$ has a representation as a sum of squares of polynomials, i.e.

$$P_A(x) = \sum_{i=1}^{l} f_i(x)^2 \text{ for all } x \in \mathbb{R}^n$$

for some polynomial functions $f_1(x), \ldots, f_l(x)$. Then it is easy to see that $P_A(x) \geq 0$ for all $x \in \mathbb{R}^n$. Parrilo showed in [42] that $P_A(x)$ has such a representation as a sum of squares if and only if $A \in \mathcal{S}_n^+ + \mathcal{N}_n$. For a given matrix $A \in \mathbb{R}^{n\times n}$ we consider the following family of polynomials:

$$P_r(x) := \left(\sum_{k=1}^{n} x(k)^2\right)^r \cdot P_A(x).$$

It is clear that $P_0(x) = P_A(x)$. Furthermore if for some r we have $P_r(x) \geq 0$ for all x, then $P_A(x) \geq 0$ for all x. Moreover if $P_r(x)$ has a sum of squares decomposition, then $P_{r+1}(x)$ also has a sum of squares decomposition. With this property Parrilo [42] defined the following hierarchy of cones for $r \in \mathbb{N}_0$:

$$\mathcal{K}_n^r := \left\{ A \in \mathcal{S}_n \colon P_A(x) \cdot \left(\sum_{i=1}^{n} x(i)^2\right)^r \text{ has a sum of squares decomposition} \right\}.$$

Parrilo also showed that $\mathcal{S}_n^+ + \mathcal{N}_n = \mathcal{K}_n^0 \subset \mathcal{K}_n^1 \subset \ldots \subset \mathcal{COP}_n$ and that $\operatorname{int}(\mathcal{COP}_n) \subseteq \bigcup_{r\in\mathbb{N}_0} \mathcal{K}_n^r$. Therefore the convex cone $\mathcal{K}_n^r$ approximates the copositive cone $\mathcal{COP}_n$ from the interior. The condition $A \in \mathcal{K}_n^r$ can be checked by solving a semidefinite feasibility problem. So the related optimization problem over $\mathcal{K}_n^r$ is equivalent to a semidefinite program.

Another approach has been introduced by de Klerk and Pasechnik [16]. They approximated the copositive cone by a system of linear inequalities. For this we consider again the polynomial $P_A(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} A(i,j)x(i)^2x(j)^2$ with $x \in \mathbb{R}^n$. Another sufficient condition for its nonnegativity is that all coefficients of $P_A(x)$ are nonnegative, i.e. $A(i,j) \geq 0$ for all i,j. This property is equivalent to $A \in \mathcal{N}_n$. For $r \in \mathbb{N}_0$ and $P_r(x) = \left(\sum_{k=1}^{n} x(k)^2\right)^r P_A(x)$ as before the authors defined the convex cone

$$\mathcal{C}_n^r := \left\{ A \in \mathcal{S}_n \colon P_A(x) \cdot \left(\sum_{i=1}^{n} x(i)^2\right)^r \text{ has nonnegative coefficients} \right\}.$$

De Klerk and Pasechnik verified that $\mathcal{N}_n = \mathcal{C}_n^0 \subset \mathcal{C}_n^1 \subset \ldots \subset \mathcal{COP}_n$ and $\operatorname{int}(\mathcal{COP}_n) \subseteq \bigcup_{r\in\mathbb{N}_0} \mathcal{C}_n^r$. Accordingly the convex cone $\mathcal{C}_n^r$ also approximates the copositive cone from the interior. Since all of these cones $\mathcal{C}_n^r$ are polyhedral cones, the related optimization problem corresponds to a linear program.

Between the two hierarchies of cones the following relation holds true:

$$\mathcal{C}_n^r \subseteq \mathcal{K}_n^r \text{ for all } r \in \mathbb{N}_0.$$

The last approximation we want to illustrate, goes back to Peña et al. [44]. The authors used the standard multiindex notation, i.e. for a given multiindex $\beta \in \mathbb{N}^n$ we have $|\beta| := \beta_1 + \ldots + \beta_n$ and $x^\beta := x_1^{\beta_1} \ldots x_n^{\beta_n}$. With this notation the following set can be defined:

$$\mathcal{E}_n^r := \left\{ \sum_{\substack{\beta \in \mathbb{N}^n, \\ |\beta| = r}} x^\beta x^\top (S_\beta + N_\beta) x \colon S_\beta \in \mathcal{S}_n^+, N_\beta \in \mathcal{N}_n \right\}.$$

This approach is based on the fact that a matrix, which can be represented as the sum of a positive semidefinite matrix and an entrywise nonnegative matrix, is copositive, cf. (2.2). Furthermore Peña et al. defined the following cone:

$$\mathcal{Q}_n^r := \left\{ A \in \mathcal{S}_n \colon x^\top A x \left(\sum_{i=1}^n x(i)^2 \right)^r \in \mathcal{E}_n^r \right\}.$$

They were also able to show that $\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r \subseteq \mathcal{K}_n^r$ for all $r \in \mathbb{N}_0$. In particular $\mathcal{Q}_n^r = \mathcal{K}_n^r$ for $r \in \{0, 1\}$. Therefore also $\mathcal{Q}_n^r$ grows with r. The condition $A \in \mathcal{Q}_n^r$ can be written as a system of linear matrix inequalities, hence if $\mathcal{COP}_n$ is replaced by the cone $\mathcal{Q}_n^r$, the resulting optimization problem is a positive semidefinite program.

Note that all convex cones $\mathcal{K}_n^r$, $\mathcal{C}_n^r$ and $\mathcal{Q}_n^r$ have pleasant theoretical properties, but a fundamental aspect is that they approximate the copositive cone uniformly. There are other techniques to approximate the copositive cone in a certain direction more precisely, e.g. approximate $\mathcal{COP}_n$ via partitions of the standard simplex into smaller simplices [12]. By using the objective function to figure out which simplices have to be splitted further, this method can approximate $\mathcal{COP}_n$ in the needed direction more precisely. This technique also leads to linear programs.

In this chapter we formulated and proved some basics of copositive optimization in finite dimension. We defined the copositive as well as the completely positive cone and their properties and furthermore the related optimization problems. These results can be used among other things to reformulate the standard quadratic problem (STQP) and also a more general form of a quadratic program (2.5) into a completely positive program. Moreover we presented some applications of copositive and completely positive programs, respectively, especially in combinatorial optimization. Besides the stability number and the clique number, many other quantities in combinatorial optimization can be calculated by solving a copositive program. An overview of the applications of copositive optimization can be found in [7, 13, 25]. Finally we gave some approximation hierarchies of the copositive cone.

Chapter 3

Relevant tools from functional analysis

In this chapter we will introduce some important aspects from functional analysis. We will also define three special norms: the (symmetric) projective norm, the (symmetric) injective norm and the (symmetric) Hilbert-Schmidt norm. Especially the symmetric projective norm gives us helpful hints for our goal to generalize the principle of copositivity and its duality in a suitable infinite dimensional space.

3.1 Preliminaries

Our notation concerning Banach spaces is the standard one, like e.g. in [46]. A *Banach space* is a normed space which is complete in the metric defined by its norm. Or in other words, every Cauchy sequence is required to converge. Many of the well-known function spaces are Banach spaces. For example the spaces of continuous functions on compact sets, the L^p−spaces (i.e. the space of classes of p-integrable functions), Hilbert spaces, or the spaces of continuous linear maps from one Banach space into another are Banach spaces.

3.2 The tensor product of two vector spaces

We consider in this thesis real vector spaces only, i.e. vector spaces over the field $\mathbb{R}$ of real numbers. First we will consider two fundamental examples:

Example 3.1. For $m, n \in \mathbb{N}$ as well as $u \in \mathbb{R}^m = \mathbb{R}^{\{1,\ldots,m\}}$ and $v \in \mathbb{R}^n = \mathbb{R}^{\{1,\ldots,n\}}$, let $u \otimes v \in \mathbb{R}^{m \times n} = \mathbb{R}^{\{1,\ldots,m\} \times \{1,\ldots,n\}}$ be defined by

$$u \otimes v\,(i, j) := u(i)\, v(j), \quad \text{for } 1 \le i \le m, 1 \le j \le n.$$

We set

$$\mathbb{R}^m \otimes \mathbb{R}^n := \operatorname{span}\{u \otimes v \colon u \in \mathbb{R}^m, v \in \mathbb{R}^n\}$$

and calculate

$$\mathbb{R}^m \otimes \mathbb{R}^n = \mathbb{R}^{m \times n},$$

cf. [19, Section 2.4 (1)]. This vector space has the following *universal property*:
Given any bilinear map $\psi \colon \mathbb{R}^m \times \mathbb{R}^n \to G$ into a vector space G. Then there is a unique linear map $L_\psi \colon \mathbb{R}^m \otimes \mathbb{R}^n \to G$ such that $\psi(u, v) = L_\psi(u \otimes v)$ for all u, v.

$$\begin{tikzcd}
\mathbb{R}^m \times \mathbb{R}^n \arrow[r, "\psi"] \arrow[d, "\otimes"'] & G \\
\mathbb{R}^m \otimes \mathbb{R}^n \arrow[ur, dashed, "L_\psi"'] &
\end{tikzcd}$$

In the infinite dimensional setting, the situation is more complicated:

Example 3.2. Consider compact spaces K and L (e.g. $K = L = [-1, 1]$) as well as continuous functions $f \colon K \to \mathbb{R}$ and $g \colon L \to \mathbb{R}$. Then we define in the same way as before the function $f \otimes g \colon K \times L \to \mathbb{R}$ by

$$f \otimes g\,(x, y) := f(x)g(y).$$

Set

$$C(K) \otimes C(L) := \operatorname{span}\{f \otimes g \colon f \in C(K), g \in C(L)\} \subseteq C(K \times L),$$

where $C(K)$ and $C(L)$ denote the spaces of continuous maps $f \colon K \to \mathbb{R}$ and $g \colon L \to \mathbb{R}$, respectively. In general (i.e. if K, L are no finite sets), we have $C(K) \otimes C(L) \neq C(K \times L)$.

$C(K) \otimes C(L)$ has the same *universal property* as above:
Given any bilinear map $\psi \colon C(K) \times C(L) \to G$ into a vector space G, then there is a unique linear map $L_\psi \colon C(K) \otimes C(L) \to G$ such that $\psi(f, g) = L_\psi(f \otimes g)$ for all $f \in C(K), g \in C(L)$.

$$\begin{tikzcd} C(K) \times C(L) \arrow[r, "\psi"] \arrow[d, "\otimes"'] & G \\ C(K) \otimes C(L) \arrow[ur, dashed, "L_\psi"'] & \end{tikzcd}$$

There is a general principle behind these two examples, cf. [19, Section 2.2]: If E, F are vector spaces, a tensor product of E and F is a pair $(E \otimes F, \otimes)$, where $E \otimes F$ is a vector space and $\otimes\colon E \times F \to E \otimes F$ is a bilinear map such that for all bilinear maps $\psi\colon E \times F \to G$ into a vector space G, there is a unique linear map $L_\psi\colon E \otimes F \to G$ with $\psi = L_\psi \circ \otimes$.

$$\begin{tikzcd} E \times F \arrow[r, "\psi"] \arrow[d, "\otimes"'] & G \\ E \otimes F \arrow[ur, dashed, "L_\psi"'] & \end{tikzcd}$$

Remark 3.3. As usual, we write $x \otimes y$ instead of $\otimes(x, y)$. Clearly, if the maps $T\colon E_2 \to E_1$ and $S\colon F_2 \to F_1$ are isomorphisms and $(E_1 \otimes F_1, \otimes)$ is a tensor product for E_1 and F_1, then $(E_2 \tilde{\otimes} F_2, \tilde{\otimes})$, where $\tilde{\otimes}\colon E_2 \times F_2 \to E_2 \tilde{\otimes} F_2$ with $(x_2, y_2) \mapsto T(x_2) \otimes S(y_2)$, is a tensor product for E_2 and F_2.

Moreover, a tensor product $(E \otimes F, \otimes)$ is uniquely determined up to canonical isomorphisms: If $(E \otimes' F, \otimes')$ is another tensor product for E and F, then there are, due to the universal property, canonical linear maps $S\colon E \otimes F \to E \otimes' F$ and $T\colon E \otimes' F \to E \otimes F$ with $\otimes' = S \circ \otimes$ and $\otimes = T \circ \otimes'$. The uniqueness of the maps L_ψ implies $E \otimes F = \operatorname{span}\{x \otimes y\colon x \in E, y \in F\}$ and analogously, $E \otimes' F = \operatorname{span}\{x \otimes' y\colon x \in E, y \in F\}$. With this it is easy to see that $S \circ T = \mathrm{id}_{E \otimes' F}$ and $T \circ S = \mathrm{id}_{E \otimes F}$, where id_G denotes the identity map $\mathrm{id}_G\colon G \to G$.

$$\begin{tikzcd} E \times F \arrow[r, "\otimes'"] \arrow[d, "\otimes"'] & E \otimes' F \arrow[dl, dashed, shift left, "T"] \\ E \otimes F \arrow[ur, dashed, shift left, "S"] & \end{tikzcd}$$

Due to the uniqueness of the tensor product - up to canonical isomorphisms - it is common to speak about *the* tensor product $E \otimes F$. Furthermore it is very useful to consider specific realizations like

$$\mathbb{R}^{m\times n} = \mathbb{R}^m \otimes \mathbb{R}^n$$

or

$$\operatorname{span}\{f(\cdot)g(\cdot\cdot)\colon f \in C(K), g \in C(L)\} = C(K) \otimes C(L).$$

These properties can be found in [19, Section 2.4].

Finally we prove the existence of the tensor product: Let $A \subset E$, $B \subset F$ be bases of the vector spaces E and F, respectively. Then E is isomorphic to $\mathcal{F}(A)$ and F is isomorphic to $\mathcal{F}(B)$, where we have $\mathcal{F}(C) := \{f\colon C \to \mathbb{R}\colon f \text{ has finite support}\}$. Set $\mathcal{F}(A) \otimes \mathcal{F}(B) := \mathcal{F}(A \times B)$ and $\otimes\colon \mathcal{F}(A) \times \mathcal{F}(B) \to \mathcal{F}(A) \otimes \mathcal{F}(B)$, with $(f \otimes g)(a, b) := f(a)g(b)$ for $a \in A$, $b \in B$. This is a tensor product for $\mathcal{F}(A)$ and $\mathcal{F}(B)$, so there exists a tensor product for E and F as in Remark 3.3, cf. [19, Sections 2.1 and 2.2].

Remark 3.4. Note here that the same notation of the tensor product and the well-known Kronecker product from optimization is not accidental. The Kronecker product can be considered as a special case of the tensor product that is restricted on matrices. The connection can be roughly described as follows: If we have two matrices, which can be considered as linear maps between vector spaces equipped with chosen bases, then the tensor product of the linear maps is representing the Kronecker product of the two matrices. For a more detailed description see [35, §9, Section 7 c)].

3.3 Tensor norms

We will start with additional structures on the vector spaces E and F, e.g. if E and F are normed, Banach or locally convex spaces, the tensor product should also carry the same structure. We restrict ourselves here to normed spaces, Banach spaces and Hilbert spaces. For our purpose it is sufficient to consider three canonical norms.

The projective norm

The first one, the $\pi-$norm or *projective norm* appears naturally if we change from vector spaces to normed spaces:

Definition 3.5. Let E, F be normed spaces. For $z \in E \otimes F$ set

$$\pi(z) := \pi(z; E, F) := \inf\left\{ \sum_{\nu=1}^{n} \|x_\nu\|_E \left\|y_\nu\right\|_F : n \in \mathbb{N}, z = \sum_{\nu=1}^{n} x_\nu \otimes y_\nu \right\}.$$

Then π is a norm and has the following *universal property*:
If $\psi \colon E \times F \to G$ is a **continuous** bilinear map with values in a normed space G, then there is a unique continuous linear map $L_\psi \colon (E \otimes F, \pi) \to G$ such that $\psi = L_\psi \circ \otimes$. Moreover,

$$\begin{aligned}
\left\|\psi\right\| &:= \sup_{\substack{\|x\|_E \le 1 \\ \|y\|_F \le 1}} \left\|\psi(x,y)\right\|_G \\
&= \sup_{\substack{\|x\|_E \le 1 \\ \|y\|_F \le 1}} \left\|L_\psi(x \otimes y)\right\|_G \\
&= \sup_{\pi(z) \le 1} \left\|L_\psi(z)\right\|_G \\
&=: \left\|L_\psi\right\|.
\end{aligned}$$

With $E \otimes_\pi F$ we will denote the tensor product $E \otimes F$ equipped with $\pi(\cdot; E, F)$.

In the next proposition we will formulate some properties of π, cf. [19, Section 3.2]. For normed spaces E, F and G we denote by $L(E, F)$ the set of linear maps from E to F. Moreover we denote the set of linear and continuous functions by $\mathcal{L}(E, F) := \{T \colon E \to F \colon T \text{ linear and continuous}\}$ and we denote by $\mathcal{B}(E, F; G) := \{B \colon E \times F \to G \colon B \text{ bilinear and continuous}\}$ the set of continuous and bilinear functions, where bilinear means linear in each variable. Furthermore $E' := \mathcal{L}(E, \mathbb{R})$ denotes the (topological) dual of E. Last we denote by $E \overset{1}{=} F$ that E and F are isometric.

Proposition 3.6. *Let E, F be normed spaces, then:*

1. $$\mathcal{B}(E, F; G) \overset{1}{=} \mathcal{L}(E \otimes_\pi F, G)$$

   and π is the unique seminorm on $E \otimes F$ which has this property for $G = \mathbb{R}$. In particular,

   $$\pi(z; E, F) = \max\{|\langle S, z\rangle| \colon S \in \mathcal{L}(E, F'), \|S\| \le 1\}$$

and the duality bracket given by

$$(E \otimes_\pi F)' \stackrel{1}{=} \mathcal{B}(E, F; \mathbb{R}) \stackrel{1}{=} \mathcal{L}(E, F')$$

can be calculated by the trace duality.

2. $\pi(x \otimes y; E, F) = \|x\|_E \|y\|_F$ *for all* $(x, y) \in E \times F$.

3. π *is finitely generated, i.e. for all* $z \in E \otimes F$

$$\pi(z; E, F) = \inf\Big\{\pi(z; M, N) \colon z \in M \otimes N, M \subset E, N \subset F, \\ \dim(M), \dim(N) < \infty\Big\}.$$

Special situations concerning the projective norm are stated in the next example. These examples give a concrete setting to calculate the projective norm:

Example 3.7. 1. If $E = \mathbb{R}^m$ and $F = \mathbb{R}^n$ equipped with $\|\cdot\|_1$, then on $E \otimes F = \mathbb{R}^{m \times n}$ we have

$$\pi(z) = \sum_{i=1}^{m} \sum_{j=1}^{n} |z(i, j)|,$$

where $z = (z(i, j))_{(i,j)} \in \mathbb{R}^{m \times n}$, cf. [30, Example 4.47].

2. More generally, if μ and ν are σ-finite measures on measure spaces X and Y, we have for the spaces $E = L^1(\mu)$ and $F = L^1(\nu)$ as well as for $f \in E \otimes F = \operatorname{span}\left\{g(\cdot)h(\cdot\cdot) \colon g \in L^1(\mu), h \in L^1(\nu)\right\}$ that

$$\pi(f) = \int_{X \times Y} |f(x, y)| \, d(\mu \otimes \nu)(x, y) = \|f\|_{L^1(\mu \otimes \nu)},$$

cf. [19, Section 3.3 Proposition].

3. If $E = \mathbb{R}^m$ and $F = \mathbb{R}^n$ equipped with the Euclidean norm $\|\cdot\|_2$, then

$$\pi(z) = \sum_{\nu=0}^{k} \lambda_\nu,$$

where $0 \leq \lambda_0 \leq \ldots \leq \lambda_k$ are the singular values of $z \in \mathbb{R}^m \otimes \mathbb{R}^n = \mathbb{R}^{m \times n}$, cf. [48, Example 2.10].

Remark 3.8. Example 3.7 can be generalized to Hilbert spaces E and F, cf. [19, Ex 3.28].

The injective norm

The second tensor norm we want to consider is the ε−norm or *injective norm*, which is the dual norm of the projective norm.

Definition 3.9. If E and F are normed spaces, then for $z \in E \otimes F$ we define:

$$\begin{aligned}\varepsilon(z) := \varepsilon(z; E, F) : &= \sup\left\{ |\langle x' \otimes y', z \rangle| : \|x'\|_{E'} \le 1, \left\|y'\right\|_{F'} \le 1 \right\} \\ &= \sup\left\{ \left| \sum_{\nu=1}^{n} x'(x_\nu) y'(y_\nu) \right| : \|x'\|_{E'} \le 1, \left\|y'\right\|_{F'} \le 1 \right\},\end{aligned}$$

if $z = \sum_{\nu=1}^{n} x_\nu \otimes y_\nu$. We write $E \otimes_\varepsilon F$ for $E \otimes F$ equipped with $\varepsilon(\cdot; E, F)$.

As mentioned in [19, Section 4.3] the norm ε behaves in a dual way to π concerning subspaces and quotients, i.e. ε respects subspaces isometrically and ε does not respect quotients.

Remark 3.10. The injective norm ε is the dual norm of the projective norm π in the sense of trace duality. But the converse property that π is the dual of ε, does not hold in this sense, cf. [19, Section 6.1]. This means that $(E\hat{\otimes}_\varepsilon E)' \neq E'\hat{\otimes}_\pi E'$ in general. But in the finite case this equality holds. Trace duality means as formulated in [19, Section 2.6]: Let $\phi \in (E \otimes F)'$ with associated map $T \in L(E, F')$ and $z \in E \otimes F$ with associated $S \in L(F', E)$ with finite rank. Then we have $\left\langle \phi, z \right\rangle = \mathrm{tr}_E(S \circ T) = \mathrm{tr}_{F'}(T \circ S)$.

For the next proposition we need the following definition:

Definition 3.11. A subset $A \subset B_{E'}$ is called *norming* if for all $x \in E$

$$\|x\|_E = \sup\{|\langle x', x \rangle| : x' \in A\}.$$

For a normed space E we note by B_E the closed unit ball $\{x \in E : \|x\|_E \le 1\}$.

The following proposition gives some properties of the injective norm:

Proposition 3.12. *[19, Section 4.1] Let E, F be normed spaces. Then:*

1. *If E and F are finite dimensional, then*

$$E \otimes_\varepsilon F \stackrel{1}{=} (E' \otimes_\pi F')'.$$

2. $\varepsilon(x \otimes y; E, F) = \|x\|_E \left\|y\right\|_F$ *for all* $x \in E$, $y \in F$.

3. *$\varepsilon \leq \pi$ on $E \otimes F$.*

4. *For norming subsets $A \subset B_{E'}$ and $B \subset B_{F'}$ we have*

$$\varepsilon(z; E, F) = \sup \{ | \langle x' \otimes y', z \rangle | : x' \in A, y' \in B \} .$$

The following examples present a way to calculate the injective norm in two special situations.

Example 3.13. 1. If $E = \mathbb{R}^m$ and $F = \mathbb{R}^n$ are equipped with $\|\cdot\|_\infty$, then on $E \otimes F = \mathbb{R}^{m\times n}$ we have

$$\varepsilon(z) = \sup_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} |z(i,j)|,$$

where $z = (z(i,j))_{(i,j)} \in \mathbb{R}^{m\times n}$, cf. [30, Remark 4.64].

2. More generally, if K, L are compact spaces, $f \in C(K) \otimes C(L)$, then we have

$$C(K) \otimes C(L) = \Big\{ f \in C(K \times L) : \exists n \in \mathbb{N}, g_\nu \in C(K), h_\nu \in C(L) :$$
$$f(\cdot, \cdot\cdot) = \sum_{\nu=1}^{n} g_\nu(\cdot) h_\nu(\cdot\cdot) \Big\},$$

cf. [19, Section 2.4. (2)] and moreover

$$\varepsilon(f) = \sup \{ |f(x,y)| : (x,y) \in K \times L \} .$$

Duality of norms

Before we consider a third norm, we will discuss the definition of a dual norm in infinite dimension, cf. [48, Chapter 7.1]. In finite dimension the definition of the dual norm is clear: If E and F are finite dimensional normed spaces and α is a tensor norm, then $E \otimes F$ is algebraically the dual space of $E' \otimes_\alpha F'$ and we set α' to be the dual norm such that

$$E \otimes_{\alpha'} F = (E' \otimes_\alpha F')' .$$

I.e. if $u \in E \otimes F$, then

$$\alpha'(u) = \sup \{ | \langle u, v \rangle | : v \in E' \otimes F', \alpha(v) \leq 1 \} . \tag{3.1}$$

In the sense of trace duality, this means that the duality between $E \otimes F$ and $E' \otimes F'$ works in the following way: If we consider u as an operator from E' to F and v from F into E', then $\langle u, v\rangle = \operatorname{tr}(v \circ u)$.

To get a definition of the dual norm in infinite dimension, Ryan [48, Chapter 7.1] uses [48, Proposition 6.3] to extend the definition in finite dimension. With that, the dual norm α' is defined as the unique tensor norm that corresponds with the dual norm on tensor products of finite dimensional spaces. Let $u \in X \otimes Y$, where X and Y have to be Banach spaces, then:

$$\alpha'(u) = \inf\left\{\alpha'_{E,F}(u) \colon u \in E \otimes F, E \subseteq X, F \subseteq Y, \dim E, \dim F < \infty\right\},$$

where $\alpha'_{E,F}(u) = \alpha'(u; E \otimes F)$.

The Hilbert-Schmidt norm

Now we have a look at the third norm. If the vector spaces G and H are equipped with scalar products $\langle\cdot,\cdot\rangle_G$ and $\langle\cdot,\cdot\rangle_H$, there is a uniquely determined scalar product $\langle\cdot,\cdot\rangle_{G\otimes H}$ on $G \otimes H$ with the property $\langle x_1 \otimes y_1, x_2 \otimes y_2\rangle_{G\otimes H} = \langle x_1, x_2\rangle_G \langle y_1, y_2\rangle_H$ for all $x_1, x_2 \in G$, $y_1, y_2 \in H$, cf. [19, Section 26.7]. The norm $\sigma := \sigma(\cdot; G, H)$, with

$$\sigma(z) := \sqrt{\langle z, z\rangle_{G\otimes H}}$$

is called the *Hilbert-Schmidt norm* on $G \otimes H$ and we write $G \otimes_\sigma H$ for the normed space $(G \otimes H, \sigma)$. Note that the Hilbert-Schmidt norm is only defined for vector spaces that are equipped with a scalar product.

Remark 3.14. Note that the dual norm of the Hilbert-Schmidt norm is again the Hilbert-Schmidt norm, i.e. $\sigma' = \sigma$, cf. [19, Section 26.7, Corollary 2].

Last we will give a special example for the Hilbert-Schmidt norm:

Example 3.15. If $G = \mathbb{R}^m$, $H = \mathbb{R}^n$, then on $G \otimes H = \mathbb{R}^{m\times n}$ we have

$$\langle w, z\rangle_{G\otimes H} = \sum_{i=1}^{m}\sum_{j=1}^{n} w(i,j)z(i,j)$$

for $w = (w(i,j))_{(i,j)}$, $z = (z(i,j))_{(i,j)} \in \mathbb{R}^{m\times n}$ and $\sigma(w) = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |w(i,j)|^2\right)^{1/2}$, cf. [19, Section 26.7].

The completions in Banach spaces

If we consider the context of Banach spaces (or particularly in the context of Hilbert spaces), the „tensor products “ should be of the same form. There we consider the completions

$$\begin{aligned} &E\hat{\otimes}_\pi F \text{ of } E \otimes_\pi F,\\ &E\hat{\otimes}_\varepsilon F \text{ of } E \otimes_\varepsilon F,\\ \text{and } &G\hat{\otimes}_\sigma H \text{ of } G \otimes_\sigma H. \end{aligned}$$

By definition the first two spaces are Banach spaces and the third one is even a Hilbert space. It can be proven that

$$\begin{aligned} L^1(\mu)\hat{\otimes}_\pi L^1(\nu) &\stackrel{1}{=} L^1(\mu \otimes \nu),\\ C(K)\hat{\otimes}_\varepsilon C(L) &\stackrel{1}{=} C(K \times L),\\ L^2(\mu)\hat{\otimes}_\sigma L^2(\nu) &\stackrel{1}{=} L^2(\mu \otimes \nu), \end{aligned}$$

cf. [19, Ex 3.27 and Section 4.2 (3)] and [34, 2.6.11 Example]. In case of two Hilbert spaces G and H we have

- $G'\hat{\otimes}_\pi H$ is canonically isomorphic to the space of nuclear operators from G to H,
- $G'\hat{\otimes}_\varepsilon H$ is canonically isomorphic to the space of compact operators from G to H,
- $G'\hat{\otimes}_\sigma H$ is canonically isomorphic to the space of Hilbert-Schmidt operators from G to H,

with $\sum_{\nu=1}^n x'_\nu \otimes y_\nu \mapsto \sum_{\nu=1}^n x'_\nu(\cdot) y_\nu$.

3.4 Algebraic theory of symmetric tensor products

In this section we are going to define a symmetric tensor product, which has the same universal property only for symmetric bilinear maps instead of all bilinear maps, as it is in the case in the full tensor product $E \otimes F$.

Definition 3.16. If E is a real vector space, a pair $(E \otimes^s E, \otimes^s)$, where $E \otimes^s E$ is a real vector space and $\otimes^s \colon E \times E \to E \otimes^s E$ is a symmetric bilinear map, is called a *symmetric tensor product* of E (formally a 2nd-symmetric tensor product) if for all vector spaces F and all symmetric bilinear maps $\psi \colon E \times E \to F$, there is a unique linearization $S_\psi \colon E \otimes^s E \to F$, i.e. there is a unique linear S_ψ with $\psi = S_\psi \circ \otimes^s$. In other words, the following diagram holds:

$$\begin{tikzcd}
E \times E \arrow[r, "\psi"] \arrow[d, "\otimes^s"'] & F \\
E \otimes^s E \arrow[ur, dashed, "S_\psi"'] &
\end{tikzcd}$$

Remark 3.17. A generalization of the 2nd-symmetric tensor product is the n-fold symmetric tensor product $(\otimes^{n,s} E, \otimes^s)$, where $\otimes^{n,s} E = E \otimes^s E \otimes^s \ldots \otimes^s E$. For a closer look at this generalization see [28, Section 1.2 et seqq.].

Since $E \otimes^s E$ is unique up to canonical isomorphisms, we are talking about **the** symmetric tensor product $E \otimes^s E$ of E. Later we will give a concrete description of this abstract object in certain examples. Let us now prove that the symmetric tensor product always exists. For this, we define

$$\begin{aligned} E \otimes^s E :&= \operatorname{span}\left\{\frac{1}{2}(x \otimes y + y \otimes x) \colon x, y \in E\right\} \\ &= \operatorname{span}\{z \otimes z \colon z \in E\} \\ &\subset E \otimes E. \end{aligned}$$

Recall here that $\frac{1}{4}\left[(x+y) \otimes (x+y) - (x-y) \otimes (x-y)\right] = \frac{1}{2}\left[x \otimes y + y \otimes x\right]$. Furthermore we define $\otimes^s \colon E \times E \to E \otimes^s E$ with $(x, y) \mapsto \frac{1}{2}(x \otimes y + y \otimes x)$. Then $E \otimes^s E$ is the image of the linear projection $P \colon E \otimes E \to E \otimes E$ with $\sum_{\nu=1}^n x_\nu \otimes y_\nu \mapsto \sum_{\nu=1}^n \frac{1}{2}(x_\nu \otimes y_\nu + y_\nu \otimes x_\nu)$, which is the linearization of the map $E \times E \to E \otimes^s E$.

Let $\psi \colon E \times E \to F$ be symmetric and bilinear. Then there is a unique factorization $L_\psi \colon E \otimes E \to F$ with $L_\psi(x \otimes y) = \psi(x, y)$. Since $\psi(x, y) = \psi(y, x)$, we get $L_\psi(x \otimes y) = L_\psi(y \otimes x)$ and $L_\psi\left(\frac{1}{2}(x \otimes y + y \otimes x)\right) = \psi(x, y)$. Thus the map $S_\psi \colon E \otimes^s E \to F$ with

$$S_\psi\left(\sum_{\nu=1}^n \frac{1}{2}(x_\nu \otimes y_\nu + y_\nu \otimes x_\nu)\right) = \sum_{\nu=1}^n \psi(x_\nu, y_\nu)$$

is well-defined, linear and satisfies

$$\psi(x,y) = S_\psi\left(\frac{1}{2}(x \otimes y + y \otimes x)\right).$$

If $S\colon E \otimes^s E \to F$ is linear, we write

$$S\left(\frac{1}{2}(x \otimes y + y \otimes x)\right) = \psi(x,y) = S_\psi\left(\frac{1}{2}(x \otimes y + y \otimes x)\right).$$

Then by linearity we have $S = S_\psi$ on $E \otimes^s E$.

3.5 Symmetric tensor norms

As in the case of the tensor product of normed spaces (or spaces with scalar product) in Section 3.3, we consider here in an analogous way the symmetric projective norm, the symmetric injective norm and the symmetric Hilbert-Schmidt norm on the symmetric tensor product.

Therefore we consider from now on only symmetric tensor products of the form $E \otimes^s E$, where E is either a normed space or a Banach space and the index s denotes that it is symmetric.

The symmetric projective norm

We will start with the analogue of the π-norm for a symmetric tensor product, which is the so-called *symmetric projective norm π_s*:

Definition 3.18. Let E be a normed space. For $z \in E \otimes^s E$ we set

$$\pi_s(z) := \pi_s(z; E, E) := \inf\left\{\sum_{\nu=1}^m \|x_\nu\|_E^2 : m \in \mathbb{N}, z = \sum_{\nu=1}^m \pm x_\nu \otimes x_\nu\right\}.$$

Note that $\pm$ is meant summandwise, i.e. it depends on the summation index ν.

Remark 3.19. [28, Section 2.3] The following relation between the projective norm and the symmetric projective norm holds:

$$\pi(z) \leq \pi_s(z) \leq 2\pi(z),$$

for all $z \in E \otimes^s E$.

Concerning this remark we will consider a concrete example:

Example 3.20. Let $\mathbb{R}^n$ be equipped with $\|\cdot\|_1$. Then let $z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^2 \otimes^s_{\pi_s} \mathbb{R}^2$. We have $z = e_1 \otimes e_2 + e_2 \otimes e_1$, where e_1, e_2 denote the first and second unit vector in $\mathbb{R}^2$. Then:

$$\begin{aligned}
\pi(z) &= \inf_{\sum_{\nu=1}^n x_\nu \otimes y_\nu = z} \sum_{\nu=1}^n \|x_\nu\|_1 \left\|y_\nu\right\|_1 \\
&\leq \|e_1\|_1 \|e_2\|_1 + \|e_2\|_1 \|e_1\|_1 \\
&= 2.
\end{aligned}$$

This estimate is already optimal. Because of the first result in Example 3.7, we have that $\pi(z) = \sum_{i=1}^2 \sum_{j=1}^2 |z(i,j)| = 2$. Furthermore with $z = \frac{1}{2}\left((e_1 + e_2) \otimes (e_1 + e_2) - (e_1 - e_2) \otimes (e_1 - e_2)\right)$ we have

$$\begin{aligned}
\pi_s(z) &= \inf_{\sum_{\iota=1}^m x_\iota \otimes x_\iota = z} \sum_{\iota=1}^m \|x_\iota\|_1^2 \\
&\leq \frac{1}{2}\left(\|e_1 + e_2\|_1^2 + \|e_1 - e_2\|_1^2\right) \\
&= 4.
\end{aligned}$$

Moreover we can show that $\pi_s(z) \geq 4$. For the proof of this inequality we need a result from Example 3.26, so we refer the reader to Remark 3.30 for the second part of the proof. Hence $\pi_s(z) = 4$ and so for this example we have $\pi_s(z) = 4 = 2\pi(z)$. This shows that the estimate $\pi_s(z) \leq 2\pi(z)$ for all z is sharp.

Note here that the symmetric projective norm is not the projective norm restricted to $E \otimes^s E$, cf. [28, Section 2.1]. If $\pi|_{E\otimes^s E}$ denotes the restriction to $E \otimes^s E$ of the projective norm on $E \otimes E$, then we have $\pi|_{E\otimes^s E} \leq \pi_s$ and $\pi_s \neq \pi|_{E\otimes^s E}$ in general but $\pi_s = \pi|_{E\otimes^s E}$ for Hilbert spaces, cf. [28, Section 2.3].

The symmetric projective norm has the following universal property: If $\phi\colon E \times E \to F$ is a continuous bilinear map with values in a normed space F, then there is a unique continuous linear map $L_\phi\colon (E \otimes^s E, \pi_s) \to F$ such that $\phi = L_\phi \circ \otimes^s$. Moreover

$$\begin{aligned}
\left\|\phi\right\| &= \sup_{\|x\|_E \leq 1} \left\|\phi(x,x)\right\|_F \\
&= \sup_{\pi(z) \leq 1} \left\|L_\psi(z)\right\|_F \\
&= \left\|L_\phi\right\|.
\end{aligned}$$

With $E \otimes^s_{\pi_s} E$ we will denote the tensor product $E \otimes^s E$ equipped with $\pi_s(\cdot; E, E)$.

A more general definition of the symmetric projective norm is the following:

$$\pi_s(z; \otimes^{n,s} E) := \inf \left\{ \sum_{\nu=1}^{m} \|x_\nu\|_E^n : m \in \mathbb{N}, z = \sum_{\nu=1}^{m} \pm \otimes^n x_\nu \right\},$$

where $\otimes^{n,s} E := E \otimes^s \ldots \otimes^s E$ and $\otimes^n x := x \otimes \ldots \otimes x$ n-times respectively. The notation will be analogous to the special case: $\otimes^{n,s}_{\pi_s} E$ and $\hat{\otimes}^{n,s}_{\varepsilon_s} E$ for the completion.
For our purpose Definition 3.18 will be sufficient. Hence we will formulate the properties in next proposition concerning Definition 3.18.

For normed spaces E, F let $\mathcal{P}^n$ be the set of continuous n-homogeneous polynomials, where $q \colon E \to F$ is an n-homogeneous polynomial if there is a map $\phi \colon E \times \ldots \times E \to F$ with $q(x) = \phi(x, \ldots, x)$ for all $x \in E$, cf. [28, Sections 1.12 and 1.13]. Next we will give some properties of the symmetric projective norm from Definition 3.18. A general formulation of these properties can be found in [28, Section 2.2]:

Proposition 3.21. 1. *For all normed spaces F we have*

$$\mathcal{P}^2(E; F) \stackrel{1}{=} \mathcal{L}\left(E \otimes^s_{\pi_s} E; F\right);$$

in particular: $\mathcal{P}^2(E; F)$ is complete if F is complete, and we have

$$\mathcal{P}^2(E; F) \stackrel{1}{=} \mathcal{L}\left(E \hat{\otimes}^s_{\pi_s} E; F\right)$$

in this case.

2. *π_s is the unique seminorm α on $(E \otimes^s E, \alpha)$ which satisfies*

$$(E \otimes^s E)' \stackrel{1}{=} \mathcal{P}^2(E).$$

3. *$\pi_s(x \otimes x; E \otimes^s E) = \|x\|_E^2$ for all $x \in E$.*

4. *π_s is finitely generated in the sense that*

$$\pi_s\left(z; E \otimes^s E\right) = \inf\left\{\pi_s\left(z; M \otimes^s M\right) : M \subset E, \dim(M) < \infty, z \in M \otimes^s M\right\}.$$

Similar to [28, Sections 1.3, 1.4, 2.1] we set

$$\theta_E^2(z) := \frac{1}{2}\left(z_1 \otimes z_2 + z_2 \otimes z_1\right) \in E \otimes E.$$

Let $\operatorname{im}(T) := \{T(x) \in F \colon x \in E\}$ be the range of $T \colon E \to F$ linear and continuous, then we denote by ι_E^2 the embedding $\operatorname{im}\left(\theta_E^2\right) \to E \otimes E$. Moreover the 2*nd polarization constant* of a normed space E is defined by

$$c(2, E) := \sup\left\{\left\|\bar{q}\right\|_{\mathcal{L}(E \times E)} : q \in B_{\mathcal{P}^2(E)}\right\},$$

where $\bar{q} \in \mathcal{L}_{\mathrm{sym}}(E \times E; F)$ stands for the unique linearization of the 2-homogeneous polynomial $q \in \mathcal{P}^2(E; F)$.

In general we have

$$c(2, E) \leq \frac{2^2}{2!} = 2.$$

For $\ell_p := \ell_p(\mathbb{R}, \|\cdot\|_p)$ it is well-known that $c(2, \ell_1) = \frac{2^2}{2!} = 2$ and $c(2, \ell_2) = 1$ and in particular $c(2, H) = 1$ for all Hilbert spaces H. Properties of the n-th polarization constant can be found in [28, Chapter 2].

With these notations the following proposition holds:

Proposition 3.22. *[28, Sections 2.1 and 2.3] Let E be a normed space. Then:*

1. $\left\|\iota_E^2 \colon E \otimes_{\pi_s}^s E \to E \otimes_\pi E\right\| = 1$ *if* $E \neq \{0\}$
2. $\left\|\theta_E^2 \colon E \otimes_\pi E \to E \otimes_{\pi|_{E \otimes^s E}}^s E\right\| = 1$
3. $\left\|\theta_E^2 \colon E \otimes_\pi E \to E \otimes_{\pi_s}^s E\right\| = c(2, E)$
4. $E \otimes_{\pi_s}^s E$ *is a topologically complemented subspace of* $E \otimes_\pi E$.

Note that the upper index 2 denotes just the case $n = 2$ and not the squared function.

Remark 3.23. Unfortunately there is no easy way as in Example 3.7, to calculate the symmetric projective norm π_s directly in special cases like in $\mathbb{R}^{n \times n}$, where $\mathbb{R}^n$ is equipped with $\|\cdot\|_1$ or $\|\cdot\|_2$.

The symmetric injective norm

The second tensor norm we want to consider, analogous to the ε-norm, is the *symmetric injective norm* ε_s, which is the dual norm of the symmetric projective norm.

Definition 3.24. If E is a normed space, then for $z \in E \otimes^s E$ we define

$$\begin{aligned}\varepsilon_s(z) := \varepsilon_s(z; E, E) &:= \sup \{ |\langle x' \otimes x', z\rangle| : \ \|x'\|_{E'} \le 1\} \\ &= \sup \left\{ \Big| \sum_{\nu=1}^{n} x'(x_\nu) \cdot x'(x_\nu) \Big| : \ \|x'\|_{E'} \le 1 \right\}\end{aligned}$$

if $z = \sum_{\nu=1}^{n} x_\nu \otimes x_\nu$. We write $E \otimes^s_{\varepsilon_s} E$ for $E \otimes^s E$ equipped with $\varepsilon_s(\cdot; E, E)$.

Remark 3.25. For the (symmetric) injective norm the following inequalities are true:

$$\varepsilon_s(z) \le \varepsilon(z) \le 2\varepsilon_s(z)$$

for all $z \in E \otimes^s E$, cf. [28, Section 3.1].

The next example illustrates the inequality stated the last remark:

Example 3.26. Let $\mathbb{R}^n$ be equipped with $\|\cdot\|_\infty$. Then let $z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^2 \otimes^s_{\varepsilon_s} \mathbb{R}^2$. We have $z = e_1 \otimes e_2 + e_2 \otimes e_1$, where e_1 and e_2 denote the first and second unit vector in $\mathbb{R}^2$. Then:

$$\begin{aligned}\varepsilon(z) &= \sup_{\substack{\|x'\|_1 \le 1 \\ \|y'\|_1 \le 1}} |\langle x' \otimes y', z\rangle| \\ &= \sup_{\substack{\|x'\|_1 \le 1 \\ \|y'\|_1 \le 1}} |\langle x' \otimes y', e_1 \otimes e_2 + e_2 \otimes e_1\rangle| \\ &= \sup_{\substack{|x'(1)|+|x'(2)| \le 1 \\ |y'(1)|+|y'(2)| \le 1}} |x'(1) \cdot y'(2) + x'(2) \cdot y'(1)| \\ &= 1.\end{aligned}$$

And furthermore:

$$\begin{aligned}
\varepsilon_s(z) &= \sup_{\|x'\|_1 \le 1} |\langle x' \otimes x', z \rangle| \\
&= \sup_{\|x'\|_1 \le 1} |\langle x' \otimes x', e_1 \otimes e_2 + e_2 \otimes e_1 \rangle| \\
&= \sup_{|x'(1)|+|x'(2)| \le 1} |x'(1) \cdot x'(2) + x'(2) \cdot x'(1)| \\
&= 1/2.
\end{aligned}$$

Indeed the second equality in each case holds by Definition 3.9 and Definition 3.24. This shows that the estimate $\varepsilon(z) \le 2\varepsilon_s(z)$ for all z is sharp.

Note here that also the symmetric injective norm is not the injective norm restricted to $E \otimes^s E$ (see [28, Section 3.1]). If $\varepsilon|_{E \otimes^s E}$ denotes the restriction to $E \otimes^s E$ of the injective norm, then $\varepsilon_s \le \varepsilon|_{E \otimes^s E}$ is valid and in particular $\varepsilon_s = \varepsilon|_{E \otimes^s E}$ for Hilbert spaces.

Definition 3.24 is just the special definition for $E \otimes^s E$. A more general one is the following:

$$\begin{aligned}
\varepsilon_s(z; E, E) :&= \sup \{ |\langle \otimes^n x', z \rangle| : \|x'\|_{E'} \le 1 \} \\
&= \sup \left\{ \Big| \sum_{k=1}^{m} \langle x', x_k \rangle^n \Big| : \|x'\|_{E'} \le 1 \right\}
\end{aligned}$$

if $z = \sum_{k=1}^m \otimes^n x_k$. The notation will be analogous to the one we used before: $\otimes_{\varepsilon_s}^{n,s} E$ and $\hat{\otimes}_{\varepsilon_s}^{n,s} E$ for the completion, cf. [28, Section 3.1]. For our purpose Definition 3.24 is sufficient. For this case we will give some properties of the symmetric injective norm from Definition 3.24. The general statements can be found in [28, Section 3.2].

Proposition 3.27. *1. $\varepsilon_s(x \otimes x; E \otimes^s E) = \|x\|_E^2$ for all $x \in E$*

2. $\varepsilon_s \le \pi_s$ on $E \otimes^s E$

3. ε_s is finitely generated, i.e.

$$\varepsilon_s(z; E \otimes^s E) = \inf \{ \varepsilon_s(z; M \otimes^s M) : M \subset E, \dim(M) < \infty, z \in M \otimes^s M \}.$$

The general formulation of following proposition can be found in [28, Section 3.1]:

Proposition 3.28. *Let E be a normed space. Then:*

1. $\left\|\iota_E^2 \colon E \otimes_{\varepsilon_s}^s E \to E \otimes_\varepsilon E\right\| \leq c(2, E')$
2. $\left\|\iota_{E'}^2 \colon E' \otimes_{\varepsilon_s}^s E' \to E' \otimes_\varepsilon E'\right\| \leq c(2, E)$
3. $\left\|\theta_E^2 \colon E \otimes_\varepsilon E \to E \otimes_{\varepsilon_s}^s E\right\| = 1$ *if* $E \neq \{0\}$
4. $E \otimes_{\varepsilon_s}^s E$ *is the topologically complemented subspace of* $E \otimes_\varepsilon E$.

Again, the upper index 2, denotes the special case for our purpose and not the squared function.

Remark 3.29. Unfortunately there is no simple formula known to calculate the symmetric projective norm ε_s concerning $\mathbb{R}^n$ equipped with $\|\cdot\|_\infty$ or in $C(K) \otimes C(K)$, where K is a compact space, analogous to Example 3.13.

Duality

Now we will formulate the duality between π_s and ε_s for the 2nd symmetric tensor product, cf. [28, Chapter 4]. If E is a normed space, then the maps

$$E \hat{\otimes}_{\varepsilon_s}^s E \to \left(E' \otimes_{\pi_s}^s E'\right)' \stackrel{1}{=} \mathcal{P}^2(E')$$
$$E' \hat{\otimes}_{\varepsilon_s}^s E' \to \left(E \otimes_{\pi_s}^s E\right)' \stackrel{1}{=} \mathcal{P}^2(E)$$

are metric injections, i.e. $\|Ix\|_F = \|x\|_E$ for $I \in \mathcal{L}(E, F)$, cf. [19, A1].

Furthermore just like in (3.1) the duality between π_s and ε_s holds, i.e.

$$\pi_s(u) = \sup_{\varepsilon_s(w) \leq 1} |\langle u, w \rangle|$$

for $u \in E \otimes_{\pi_s}^s E$ and $w \in E' \otimes_{\varepsilon_s}^s E'$. With this property and the results from Example 3.26 we can now prove the inequality from Example 3.20, which was still to be done.

Remark 3.30. In Example 3.20 we used the fact that $\pi_s(z) \geq 4$, which we will prove now. In the situation of Example 3.26 we know that $\varepsilon_s(z) = 1/2$. Moreover with the duality of π_s and ε_s in the situation of Example 3.20 the following holds for $w \in \mathbb{R}^2 \otimes_{\varepsilon_s}^s \mathbb{R}^2$:

$$\begin{aligned}\pi_s(z) &= \sup_{\varepsilon_s(w) \leq 1} |\langle z, w \rangle| \\ &\geq \langle e_1 \otimes e_2 + e_2 \otimes e_1, 2\,(e_1 \otimes e_2 + e_2 \otimes e_1) \rangle \\ &= 4.\end{aligned}$$

Consequently $\pi_s(z) \geq 4$ and the still pending inequality is proved.

The symmetric Hilbert-Schmidt norm

Last but not least we consider a third symmetric norm, analogous to the Hilbert-Schmidt norm. If a vector space F is equipped with a scalar product $\langle \cdot, \cdot \rangle_F$ there is a uniquely determined scalar product $\langle \cdot, \cdot \rangle_{F \otimes^s F}$ with $\langle x_1 \otimes x_1, x_2 \otimes x_2 \rangle_{F \otimes^s F} = \langle x_1, x_2 \rangle_F \langle x_1, x_2 \rangle_F$ for $x_1, x_2 \in F$, cf. [19, Section 26.7]. The norm $\sigma_s := \sigma_s(\cdot; F, F)$ with

$$\sigma_s(z) := \sqrt{\langle z, z \rangle_{F \otimes^s F}}$$

is called the *symmetric Hilbert-Schmidt norm* on $F \otimes^s F$ and we write $F \otimes^s_{\sigma_s} F$ for the normed space $(F \otimes^s F, \sigma_s)$.

Remark 3.31. Note here that in contrast to the symmetric projective and the symmetric injective norms, the symmetric Hilbert-Schmidt norm is equal to the Hilbert-Schmidt norm restricted to $F \otimes^s F$, which is denoted by $\sigma|_{F \otimes^s F}$, i.e. $\sigma_s(z) = \sigma|_{F \otimes^s F}(z)$ for all $z \in F \otimes^s F$. This results from the uniqueness of the scalar product.

In the present chapter concerning functional analysis we gave an overview over the most important tools for our further research. Of particular importance were the algebraic theory of the symmetric tensor products. It will be essential for our further research since we will consider only symmetric operators. An important instrument will be the symmetric projective norm, which is not only the restriction of the projective norm to symmetric tensor products. Rather it arises via the symmetric decompositions of the considered tensor. With the help of this norm we will find suitable spaces for our goal to generalize the concept of copositivity in an infinite dimensional space.

Chapter 4

Discussion of special cases

In this chapter we will have a closer look at the symmetric projective norm and its special properties. We will analyze how the value of the symmetric projective norm depends on the different p-norms. Furthermore we will see that there are two special cases depending on the choice of p in which the symmetric projective norm is easier to calculate and that in the other cases the calculation is not so easy. Moreover we will give different examples to illustrate these special cases. In the second part of this chapter we will give some numerical experiments for a special kind of decomposition. Last we will discuss the importance of the symmetric projective norm for our goal to generalize the topic of copositivity and completely positivity in an infinite dimensional space.

4.1 Properties of the symmetric projective norm

First we will repeat the definition of the symmetric projective norm particularly for $\ell_p^n \otimes^s \ell_p^n$:

Definition 4.1. The *symmetric projective norm* $\pi_{s,p}$ on the symmetric tensor space $\ell_p^n \otimes^s \ell_p^n$, with $\ell_p^n := (\mathbb{R}^n, \|\cdot\|_p)$ depending on p is for $z \in \ell_p^n \otimes^s \ell_p^n$ defined as follows:

$$\pi_{s,p}(z) := \pi_s\left(z; \ell_p^n, \ell_p^n\right) := \inf\left\{\sum_{\nu=1}^{N} \|x_\nu\|_p^2 : N \in \mathbb{N}, z = \sum_{\nu=1}^{N} \pm\, x_\nu \otimes x_\nu\right\}.$$

For the sake of simplicity we will drop p in the index of $\pi_{s,p}$ if it is clear which p is used.

The symmetric projective norm is generally very impractical to calculate because of the many possibilities for decompositions of a tensor and beyond that the different choices of values of p. Therefore we thought about special cases depending on the value of p, in which the symmetric projective norm would be easier to calculate. Furthermore as a special issue, it would be less complicated if the set of decompositions would be smaller. With regard to the generalization of the completely positive cone, we were able to prove the following result that deals only with nonnegative decompositions:

Theorem 4.2. *Let $z = \sum_{\nu=1}^{N} u_\nu \otimes u_\nu$ with $u_\nu \geq 0$. If $p = 1$ or $p = 2$, then:*

$$\pi_{s,p}(z) = \sum_{\nu=1}^{N} \|u_\nu\|_p^2 .$$

Proof. First we prove the case $p = 1$: Consider $z = \sum_{\nu=1}^{N} u_\nu \otimes u_\nu \in \ell_1^n \otimes^s \ell_1^n$ and $u_\nu \geq 0$. With $z(i, j)$ we denote the (i, j)-th entry of z. Then:

$$\begin{aligned}
\pi_s\left(z; \ell_1^n, \ell_1^n\right) &\geq \pi\left(z; \ell_1^n, \ell_1^n\right) \\
&= \sum_{i,j=1}^{n} \left|z(i, j)\right| \\
&= \sum_{i,j=1}^{n} z(i, j) \\
&= \sum_{\nu=1}^{N} \left(\sum_{i,j=1}^{n} u_\nu(i) \cdot u_\nu(j) \right) \\
&= \sum_{\nu=1}^{N} \left(\sum_{i=1}^{n} u_\nu(i) \right)^2 \\
&= \sum_{\nu=1}^{N} \|u_\nu\|_1^2 \\
&\geq \pi_s\left(z; \ell_1^n, \ell_1^n\right).
\end{aligned}$$

This chain of inequalities shows that

$$\pi_s\left(z; \ell_1^n, \ell_1^n\right) \geq \sum_{\nu=1}^{N} \|u_\nu\|_1^2 \geq \pi_s\left(z; \ell_1^n, \ell_1^n\right)$$

and therefore the equality is proved:

$$\pi_s\left(z;\ell_1^n,\ell_1^n\right)=\sum_{\nu=1}^{N}\|u_\nu\|_1^2 .$$

Now we show the property for $p=2$: Let $z=\sum_{\nu=1}^{N}u_\nu\otimes u_\nu\in\ell_2^n\otimes^s\ell_2^n$ and $u_\nu\geq 0$. With $z(i,j)$ we denote again the (i,j)-th entry of z. Furthermore let $\lambda_\nu(z)$ with $\nu=1,\ldots,K$ be the singular values of z. Then:

$$\begin{aligned}
\pi_s\left(z;\ell_2^n,\ell_2^n\right)&\geq\pi\left(z;\ell_2^n,\ell_2^n\right)\\
&=\sum_{\nu=1}^{K}|\lambda_\nu(z)|\\
&=\sum_{\nu=1}^{K}\lambda_\nu(z)\\
&=\operatorname{tr}(z)\\
&=\sum_{i=1}^{n}z(i,i)\\
&=\sum_{\nu=1}^{N}\left(\sum_{i=1}^{n}u_\nu(i)\cdot u_\nu(i)\right)\\
&=\sum_{\nu=1}^{N}\sum_{i=1}^{n}(u_\nu(i))^2\\
&=\sum_{\nu=1}^{N}\|u_\nu\|_2^2\\
&\geq\pi_s\left(z;\ell_2^n,\ell_2^n\right).
\end{aligned}$$

Again this chain of inequalities shows that

$$\pi_s\left(z;\ell_2^n,\ell_2^n\right)\geq\sum_{\nu=1}^{N}\|u_\nu\|_2^2\geq\pi_s\left(z;\ell_2^n,\ell_2^n\right)$$

and therefore equality holds:

$$\pi_s\left(z;\ell_2^n,\ell_2^n\right)=\sum_{\nu=1}^{N}\|u_\nu\|_2^2 .$$

□

Remark 4.3. Note that in the case $p = 1$ the restriction $u_\nu \geq 0$ for all ν is essential for the proof of Theorem 4.2. But in the case $p = 2$ this restriction is not necessary for the proof. So this property holds for any positive semidefinite z, and not only for completely positive z. But nonetheless it is a useful result that we need just a single - and particularly nonnegative - decomposition to determine $\pi_{s,p}$.

In the next example we will show that in the case $p \in (1, 2)$ the equality from Theorem 4.2 does not hold in general:

Example 4.4. Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. We have one decomposition $A = \sum_{i=1}^{2} b_i \otimes b_i$ with $b_1 = (\sqrt{3/2}, 0)^\top$ and $b_2 = (1/\sqrt{2}, \sqrt{2})^\top$ and a second $A = \sum_{j=1}^{3} c_j \otimes c_j$ with $c_1 = (0, 1)^\top$, $c_2 = (1, 1)^\top$ and $c_3 = (1, 0)^\top$. First for $p \in [1, 2]$ we determine:

$$\begin{aligned}
\sum_{i=1}^{2} \|b_i\|_p^2 &= \left(\left(\left(\sqrt{\frac{3}{2}} \right)^p + 0^p \right)^{1/p} \right)^2 + \left(\left(\left(\frac{1}{\sqrt{2}} \right)^p + \sqrt{2}^p \right)^{1/p} \right)^2 \\
&= \frac{3}{2} + \left(\frac{1}{2^{p/2}} + 2^{p/2} \right)^{2/p} \\
&= \frac{3}{2} + \frac{1}{2} (1 + 2^p)^{2/p}
\end{aligned}$$

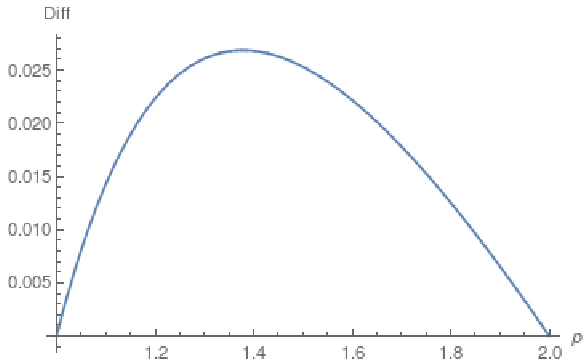
whereas

$$\begin{aligned}
\sum_{j=1}^{3} \left\| c_j \right\|_p^2 &= \left((0^p + 1^p)^{1/p} \right)^2 + \left((1^p + 1^p)^{1/p} \right)^2 + \left((1^p + 0^p)^{1/p} \right)^2 \\
&= 1 + 2^{2/p} + 1 \\
&= 2 + 2^{2/p}.
\end{aligned}$$

Next we show that for $p \in [1, 2]$ we have

$$\sum_{j=1}^{3} \left\| c_j \right\|_p^2 \geq \sum_{i=1}^{2} \|b_i\|_p^2 .$$

Especially for $p \in (1, 2)$ the inequality holds strictly. To illustrate this inequality, we consider the graph of the difference $\sum_{j=1}^{3} \left\| c_j \right\|_p^2 - \sum_{i=1}^{2} \|b_i\|_p^2$:



Let $p \in [1,2]$, then:

$$\begin{aligned} 2 + 2^{2/p} \quad &\geq \frac{3}{2} + \left(\frac{1}{2^{p/2}} + 2^{p/2}\right)^{2/p} \\ \Leftrightarrow \frac{1}{2} + 2^{2/p} \quad &\geq \left(\frac{1}{2^{p/2}} + 2^{p/2}\right)^{2/p} \\ \Leftrightarrow 1 + 2^{2/p+1} &\geq 2\left(\frac{1}{2^{p/2}} + 2^{p/2}\right)^{2/p} \\ \Leftrightarrow 1 + 2^{2/p+1} &\geq (1 + 2^p)^{2/p}. \end{aligned}$$

To prove the inequality

$$(1 + 2^p)^{2/p} \leq 1 + 2^{2/p+1} \tag{4.1}$$

we look for a function $g(p)$ such that

$$(1 + 2^p)^{2/p} \leq g(p) \leq 1 + 2^{2/p+1}.$$

For this purpose we use the property that the map $t \mapsto t^{2/p}$ is convex. For $\lambda \in (0,1)$ we have:

$$\begin{aligned} (1 + 2^p)^{2/p} &= \left(\lambda \frac{1}{\lambda} + (1-\lambda)\frac{2^p}{(1-\lambda)}\right)^{2/p} \\ &\leq \lambda \left(\frac{1}{\lambda}\right)^{2/p} + (1-\lambda)\,\frac{(2^p)^{2/p}}{(1-\lambda)^{2/p}} \\ &= \left(\frac{1}{\lambda}\right)^{2/p-1} + 4\frac{1}{(1-\lambda)^{2/p-1}} =: g(p). \end{aligned}$$

For $p = 1$ by (4.1) and Theorem 4.2 we have $(1+2^p)^{2/p} = 1 + 2^{2/p+1}$, hence $(1+2^p)^{2/p} = g(p) = 1 + 2^{2/p+1}$ and we can find a solution for $\lambda \in (0,1)$:

$$\begin{aligned}
\left(1+2^1\right)^2 = \left(\frac{1}{\lambda}\right)^{2/1-1} + 4\left(\frac{1}{1-\lambda}\right)^{2/1-1} &= 1 + 2^{2/1+1} \\
\Leftrightarrow \frac{1}{\lambda} + 4\frac{1}{1-\lambda} &= 9 \\
\Leftrightarrow (1-\lambda) + 4\lambda &= 9\lambda(1-\lambda) \\
\Leftrightarrow 9\lambda^2 - 6\lambda + 1 &= 0 \\
\Leftrightarrow \lambda^2 - \frac{2}{3}\lambda + \frac{1}{9} &= 0 \\
\Leftrightarrow \left(\lambda - \frac{1}{3}\right)^2 &= 0 \\
\Leftrightarrow \lambda &= \frac{1}{3}.
\end{aligned}$$

With $\lambda = \frac{1}{3}$ we get $g(p) = 3^{2/p-1} + 4\left(\frac{3}{2}\right)^{2/p-1}$. To simplify this expression we substitute $2/p - 1 =: x$ and hence $x \in [0,1]$. Then

$$g(x) = 3^x + 4\left(\frac{3}{2}\right)^x$$

and the right-hand side of (4.1) is equal to $1 + 2^{x+2} = 1 + 4 \cdot 2^x$. Instead of proving (4.1) we prove the following inequality for $x \in [0,1]$:

$$\begin{aligned}
3^x + 4\left(\frac{3}{2}\right)^x &\leq 1 + 4 \cdot 2^x \\
\Leftrightarrow 1 + 4 \cdot 2^x - 3^x - 4\left(\frac{3}{2}\right)^x &\geq 0 \\
\Leftrightarrow \left(\frac{1}{2}\right)^x + 4 \cdot 2^x \cdot \left(\frac{1}{2}\right)^x - \left(\frac{3}{2}\right)^x - 4\left(\frac{3}{4}\right)^x &\geq 0 \\
\Leftrightarrow \underbrace{\left(\frac{1}{2}\right)^x + 4 - \left(\frac{3}{2}\right)^x - 4\left(\frac{3}{4}\right)^x}_{=:\, f(x)} &\geq 0.
\end{aligned}$$

We have $f(0) = 1 + 4 - 1 - 4 = 0$ and $f(1) = \frac{1}{2} + 4 - \frac{3}{2} - 3 = 0$. To show that the inequality above holds strictly for $x \in (0, 1)$, it is enough to verify that $f\colon [0, 1] \to \mathbb{R}$ is strictly concave. To prove this we calculate the second derivative and show that $f''(x) < 0$ for $x \in (0, 1)$:

$$f'(x) = -2^{-x} \ln(2) - \left(\frac{3}{2}\right)^x \ln\left(\frac{3}{2}\right) + 4 \cdot \left(\frac{3}{4}\right)^x \ln\left(\frac{4}{3}\right),$$

where ln denotes the natural logarithm that has e as its base. Thus for $x \in (0, 1)$ we have:

$$\begin{aligned}
f''(x) &= -\ln(2)\left(-2^{-x}\right)\ln(2) - \ln\left(\frac{3}{2}\right)\left(\frac{3}{2}\right)^x \ln\left(\frac{3}{2}\right) + 4\ln\left(\frac{4}{3}\right)\left(\frac{3}{4}\right)^x \ln\left(\frac{4}{3}\right) \\
&= \ln^2(2)\left(\frac{1}{2}\right)^x - \ln^2\left(\frac{3}{2}\right)\left(\frac{3}{2}\right)^x - 4\ln^2\left(\frac{4}{3}\right)\left(\frac{3}{4}\right)^x \\
&\le \ln^2(2)\left(1 - \frac{x}{2}\right) - \ln^2\left(\frac{3}{2}\right)\left(1 + \ln\left(\frac{3}{2}\right)x\right) - 4\ln^2\left(\frac{4}{3}\right)\left(1 + \ln\left(\frac{3}{4}\right)x\right) \\
&= \left(\frac{\ln^2(2)}{-2} - \ln^3\left(\frac{3}{2}\right) - 4\ln^3\left(\frac{4}{3}\right)\right)x + \ln^2(2) - \ln^2\left(\frac{3}{2}\right) - 4\ln^2\left(\frac{4}{3}\right) \\
&< -0.40212x - 0.014992 \\
&< 0.
\end{aligned}$$

The inequality in the third row follows with the Taylor series of $\left(\frac{3}{2}\right)^x$ and $\left(\frac{3}{4}\right)^x$. The inequality $\left(\frac{1}{2}\right)^x \le 1 - \frac{x}{2}$ is valid because $\left(\frac{1}{2}\right)^x$ is a convex function and $1 - \frac{x}{2}$ is a linear function and they are equal in $x = 0$ and $x = 1$. Hence f is strictly concave on $(0, 1)$. Therefore the inequality (4.1) is correct and holds strictly if $p \in (0, 1)$.

With this example we proved that if $p \in (1, 2)$ then $\sum_{i=1}^{2} \|b_i\|_p^2 \neq \sum_{j=1}^{3} \left\|c_j\right\|_p^2$ for two different decompositions of A. So in these cases Theorem 4.2 does not apply, and it is more complicated to determine the symmetric projective norm. Here we have to find a decomposition for which the infimum is attained.

In a second example we will see that even the quotient of the sum of squared norms of two decompositions can be arbitrarily large and that there is no way to estimate them against each other. In this example we have a large representation with many different tensors.

Example 4.5. Let $1 < p < 2$ and let $P_n^k := \{S \subset \{1, \ldots, n\} : |S| = k\}$. Then $\left|P_n^k\right| = \binom{n}{k}$. Set

$$A = (A(i,j))_{(i,j)} := \sum_{S \in P_n^k} \left(\sum_{i \in S} e_i \right) \otimes \left(\sum_{i \in S} e_i \right) \in \mathbb{R}^{n \times n},$$

where e_i denotes again the i-th unit vector in $\mathbb{R}^n$. Since there exists a decomposition consisting of nonnegative elements, A is completely positive.

If $(i,j) \in \{1, \ldots, n\}^2$ then:

$$\left|\left\{S \in P_n^k : (i,j) \in S^2\right\}\right| = \begin{cases} \binom{n-1}{k-1} & : i = j \\ \binom{n-2}{k-2} & : i \neq j \end{cases}$$

and consequently

$$a_{ij} = \begin{cases} \binom{n-1}{k-1} & : i = j \\ \binom{n-2}{k-2} & : i \neq j. \end{cases}$$

With this property we get a second decomposition of A:

$$A = \binom{n-2}{k-2} \left(\sum_{i=1}^{n} e_i \right) \otimes \left(\sum_{j=1}^{n} e_j \right) + \left(\binom{n-1}{k-1} - \binom{n-2}{k-2} \right) \sum_{i=1}^{n} e_i \otimes e_i.$$

For the first representation of A we set:

$$\sum_{S \in P_n^k} \left\| \sum_{i \in S} e_i \right\|_p^2 =: \alpha_n.$$

And for the second decomposition of A we define:

$$\binom{n-2}{k-2} \left\| \left(\sum_{i=1}^{n} e_i \right) \right\|_p^2 + \left(\binom{n-1}{k-1} - \binom{n-2}{k-2} \right) \sum_{i=1}^{n} \|e_i\|_p^2 =: \beta_n.$$

The comparison of both for $p \in (1,2)$ leads to

$$\frac{\beta_n}{\alpha_n} = \frac{\binom{n-2}{k-2}\overbrace{\left\|\left(\sum_{i=1}^{n} e_i\right)\right\|_p^2}^{=n^{2/p}} + \left(\binom{n-1}{k-1} - \binom{n-2}{k-2}\right)\overbrace{\sum_{i=1}^{n}\|e_i\|_p^2}^{=n}}{\underbrace{\sum_{S\in P_n^k}\left\|\sum_{i\in S} e_i\right\|_p^2}_{=\binom{n}{k}k^{2/p}}}$$

$$= \frac{\frac{(n-2)!}{(k-2)!(n-k)!}n^{2/p} + \left(\frac{(n-1)!}{(k-1)!(n-k)!} - \frac{(n-2)!}{(k-2)!(n-k)!}\right)n}{\frac{n!}{k!(n-k)!}k^{2/p}}$$

$$= \left(\frac{(n-2)!n^{2/p}}{(k-2)!(n-k)!} + \frac{(n-1)!n}{(k-1)!(n-k)!} - \frac{(n-2)!n}{(k-2)!(n-k)!}\right)\frac{k!(n-k)!}{n!k^{2/p}}$$

$$= \frac{(k-1)kn^{2/p}}{(n-1)nk^{2/p}} + \frac{kn}{nk^{2/p}} - \frac{(k-1)kn}{(n-1)nk^{2/p}}$$

$$= \frac{k-1}{n-1}k^{1-2/p}n^{2/p-1} + k^{1-2/p} - \frac{k-1}{n-1}k^{1-2/p}$$

$$= k^{1-2/p}\Bigg(\underbrace{\frac{k-1}{n-1}n^{2/p-1}}_{\substack{\to 0\\(n\to\infty)}} + 1 - \underbrace{\frac{k-1}{n-1}}_{\substack{\to 0\\(n\to\infty)}}\Bigg),$$

for k fixed. It is also possible to choose k depending on n, due to the convergence k is only allowed to grow as fast as $\sqrt{n}$. This shows that we can choose $k = k_n$ such that

$$\frac{\alpha_n}{\beta_n} = \frac{1}{k^{1-2/p}\left(\frac{k-1}{n-1}n^{2/p-1} + 1 - \frac{k-1}{n-1}\right)} \to \infty \ (n \to \infty).$$

This second example shows much more that for two decompositions the values α_n and β_n can not be estimated against each other because the quotient of both gets arbitrarily large. Therefore it is very difficult to calculate $\pi_{s,p}$ if $p \in (1,2)$.

The last case we will discuss is the case $p > 2$:

Proposition 4.6. *Let $z = \sum_{\iota=1}^{M} u_\iota \otimes u_\iota$ with $u_\iota \geq 0$. If $p > 2$ then*

$$\pi_{s,p}(z) \neq \inf \left\{ \sum_{\nu=1}^{N} \|x_\nu\|_p^2 : z = \sum_{\nu=1}^{N} x_\nu \otimes x_\nu,\ x_\nu \geq 0, N \in \mathbb{N} \right\}.$$

Proof. To prove the result, we consider the $n \times n$ identity matrix I_n and a special decomposition of it:

$$I_n = \sum_{i=1}^{n} e_i \otimes e_i,$$

where e_i denotes again the i-th unit vector in $\mathbb{R}^n$. This is up to rearrangements or multiples the unique nonnegative and symmetric decomposition of I_n. Therefore we know:

$$\sum_{i=1}^{n} \|e_i\|_p^2 = n = \inf \left\{ \sum_{\nu=1}^{N} \|x_\nu\|_p^2 : I_n = \sum_{\nu=1}^{N} x_\nu \otimes x_\nu,\ x_\nu \geq 0, N \in \mathbb{N} \right\}.$$

There is no possibility to get a symmetric and nonnegative decomposition with a smaller sum of squared norms. However, with [19, Section 8.8, Corollary] it follows that

$$\pi_{s,p}\,(I_n) = \begin{cases} n^{2/p} & \text{if } p \geq 2 \\ n & \text{if } p \leq 2. \end{cases}$$

For $p > 2$ and $n \geq 2$ we get:

$$\pi_{s,p}(I_n) = n^{2/p} < n = \inf \left\{ \sum_{\nu=1}^{N} \|x_\nu\|_p^2 : I_n = \sum_{\nu=1}^{N} x_\nu \otimes x_\nu,\ x_\nu \geq 0, N \in \mathbb{N} \right\}.$$

In particular

$$\frac{n^{2/p}}{n} \to 0 \quad \text{for } n \to \infty.$$

□

Therefore if $p > 2$ it is not possible to calculate $\pi_{s,p}$ by considering only the nonnegative decompositions as Proposition 4.6 shows.

Another question that arises while thinking about special properties of the symmetric projective norm is whether there is any relation between the number of summands of a decomposition and the size of its sum of squared norms. For example whether a higher (or lower) number of summands leads to a higher (or lower) sum of squared norms or the other way around.

In a first result, we want to show that a decomposition with more summands may have the same norm.

Lemma 4.7. *Let $z = \sum_{\nu=1}^{m} u_\nu \otimes u_\nu$. Then for any $n > m$ there exists a different decomposition $z = \sum_{\iota=1}^{n} w_\iota \otimes w_\iota$ such that $\sum_{\nu=1}^{m} \|u_\nu\|_p^2 = \sum_{\iota=1}^{n} \|w_\iota\|_p^2$.*

Proof. Let $k := n - m$. We construct the decomposition as follows: For $\iota = 1, \dots, m-1$ we set $w_\iota := u_\nu$. Furthermore for $\iota = m, \dots, m+k$ we define $w_\iota := \frac{1}{\sqrt{k+1}} u_m$. Then obviously $\sum_{\nu=1}^{m} u_\nu \otimes u_\nu = \sum_{\iota=1}^{n} w_\iota \otimes w_\iota = z$, and we get the following sums of squared norms for the two decompositions:

$$\alpha := \sum_{\nu=1}^{m} \|u_\nu\|_p^2 = \sum_{\nu=1}^{m-1} \|u_\nu\|_p^2 + \|u_m\|_p^2$$
$$\beta := \sum_{\iota=1}^{n} \|w_\iota\|_p^2 = \sum_{\iota=1}^{m-1} \|u_\iota\|_p^2 + \sum_{\iota=m}^{m+k} \left\| \frac{1}{\sqrt{k+1}} u_m \right\|_p^2.$$

To prove that both sums of squared norms have the same value, we calculate their difference:

$$\begin{aligned} \beta - \alpha &= \sum_{\iota=1}^{m-1} \|u_\iota\|_p^2 + \sum_{\iota=m}^{m+k} \left\| \frac{1}{\sqrt{k+1}} u_m \right\|_p^2 - \sum_{\nu=1}^{m-1} \|u_\nu\|_p^2 - \|u_m\|_p^2 \\ &= (k+1) \left| \frac{1}{\sqrt{k+1}} \right|^2 \|u_m\|_p^2 - \|u_m\|_p^2 \\ &= 0. \end{aligned}$$

Hence it is true that $\sum_{\nu=1}^{m} \|u_\nu\|_p^2 = \sum_{\iota=1}^{n} \|w_\iota\|_p^2$. □

As the above lemma shows, a larger number of summands does not necessarily imply a larger sum of squared norms. The principle of duplicating summands, which is used in the proof of Lemma 4.7, is well-known in optimization. Although we duplicated only one summand, the result would stay the same if we would duplicate more than one summand. In optimization this principle is used for example in [26, Chapter 3].

The next example shows a stronger statement of Lemma 4.7. Here we will consider two decompositions and we will see that the one with more summands has a smaller sum of squared norms.

Example 4.8. We consider the matrix

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

and two decompositions $A = \sum_{i=1}^{6} b_i \otimes b_i = \sum_{j=1}^{5} c_i \otimes c_j$. For the sake of clarity we collect the vectors b_i ($i = 1, \dots, 6$) and c_j ($j = 1, \dots, 5$) in matrices $B = (b_1, b_2, b_3, b_4, b_5, b_6)$ and $C = (c_1, c_2, c_3, c_4, c_5)$, respectively. Then:

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{5\times 6},$$

$$C = \begin{pmatrix} 1.1145 & 0.3408 & 0.4039 & 0.6885 & 0.00679 \\ 0.0009 & 0.0553 & 0.0773 & 1.3326 & 0.4639 \\ 0 & 0.1039 & 1.2388 & 0.6743 & 0 \\ 0 & 1.2046 & 0.3646 & 0.6277 & 0.1483 \\ 0.3129 & 0.3461 & 0.6143 & 0.3011 & 1.1464 \end{pmatrix} \in \mathbb{R}^{5\times 5}.$$

We calculate now the sum of squared norms of the first decomposition:

$$\begin{aligned} \sum_{i=1}^{6} \|b_i\|_p^2 &= (1^p + 1^p + 1^p + 1^p + 1^p)^{2/p} + 5 \cdot (1^p)^{2/p} \\ &= 5^{2/p} + 5 \end{aligned}$$

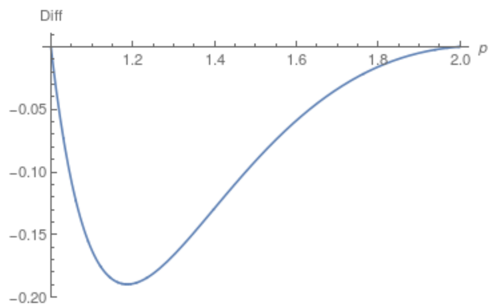
and for the second decomposition we have:

$$\begin{aligned} \sum_{j=1}^{5} \left\|c_j\right\|_p^2 = &(1.1145^p + 0.3408^p + 0.4039^p + 0.6885^p + 0.00679^p)^{2/p} \\ &+ (0.0009^p + 0.0553^p + 0.0773^p + 1.3326^p + 0.4639^p)^{2/p} \\ &+ (0^p + 0.1039^p + 1.2388^p + 0.6743^p + 0^p)^{2/p} \\ &+ (0^p + 1.2046^p + 0.3646^p + 0.6277^p + 0.1483^p)^{2/p} \\ &+ (0.3129^p + 0.3461^p + 0.6143^p + 0.3011^p + 1.1464^p)^{2/p} \end{aligned}$$

For these decompositions and $p \in (1,2)$ we have:

$$\sum_{i=1}^{6} \|b_i\|_p^2 < \sum_{j=1}^{5} \left\|c_j\right\|_p^2$$

To illustrate this inequality, we consider the graph of the difference $\sum_{i=1}^{6} \|b_i\|_p^2 - \sum_{j=1}^{5} \left\|c_j\right\|_p^2$:



Of course if $p = 1$ or $p = 2$ it results that $\sum_{i=1}^{6} \|b_i\|_p^2 - \sum_{j=1}^{5} \left\|c_j\right\|_p^2 = 0$ by Theorem 4.2.

We have to mention here that the second decomposition in Example 4.8 comes from [26, Algorithm 2]. Thus we found an example where the decomposition with the larger number of summands has a smaller value of its sum of squared norms. An illustration for decompositions where the one with the larger number of summands does also have a larger sum of squared norms is Example 4.4.

Summarizing we can say: Although it would be a desirable result since the symmetric projective norm would be easier to calculate, there is unfortunately no known relation between the number of summands of a decomposition and the value of the sum of squared norms in general.

In this section we discussed the symmetric projective norm $\pi_{s,p}$ by considering different cases of p. We were able to prove for two special cases ($p = 1$ and $p = 2$) that $\pi_{s,p}$ can be calculated by considering only a single nonnegative decomposition. These special cases also give hints for our goal to generalize the topic of copositivity and completely positivity suitably in infinite dimension. Furthermore we proved that this restriction is not possible in general, e.g. if $p > 2$, it is not possible to calculate $\pi_{s,p}$ by dealing with only the nonnegative decompositions. Moreover we were able to show that there is no direct relation between the number of summands of a decomposition and the value of its sum of squared norms.

4.2 Numerical experiments

In this section we will have a closer look at the symmetric projective norm and its properties if $p \in (1,2)$. Unfortunately we were not able to prove that the $\pi_{s,p}$-norm can be calculated by considering only the nonnegative decompositions in general. But we **conjecture** that the following equality holds for any $z \in \ell_p^n \otimes^s \ell_p^n$ that allows a symmetric nonnegative decomposition:

$$\pi_{s,p}(z) \stackrel{?}{=} \inf\left\{ \sum_{\nu=1}^{N} \|x_\nu\|_p^2 : z = \sum_{\nu=1}^{N} x_\nu \otimes x_\nu,\ x_\nu \geq 0, N \in \mathbb{N} \right\}. \tag{4.2}$$

Therefore in this section we will consider three different examples to verify conjecture (4.2) numerically. For each example we compute the sum of squared p-norms for all suitable symmetric and nonnegative decompositions and determine the numerical minimum. Furthermore we will show that there is at least one nonnegative decomposition for which the minimum is attained. In these examples we will calculate the minimum only for special completely positive matrices in $\mathbb{R}^{2\times 2}$ which can be decomposed in the sum of two symmetric and positive semidefinite vector-vector-products.

Let A be a (2×2)-completely positive matrix. Then we can represent A in a general decomposition of two positive semidefinite matrices in the following form:

$$A = \begin{pmatrix} a^2 & c^2 \\ c^2 & b^2 \end{pmatrix} = \begin{pmatrix} \alpha^2 & \alpha\beta \\ \alpha\beta & \beta^2 \end{pmatrix} + \begin{pmatrix} \gamma^2 & \gamma\delta \\ \gamma\delta & \delta^2 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\begin{pmatrix} \alpha & \beta \end{pmatrix} + \begin{pmatrix} \gamma \\ \delta \end{pmatrix}\begin{pmatrix} \gamma & \delta \end{pmatrix}. \tag{4.3}$$

Furthermore we assume that $\det(A) > 0$, which means that A has full rank. Therefore we have $a^2b^2 > c^4 \geq 0$ which is equivalent to $ab > c^2 \geq 0$. With the decomposition in (4.3) we know that

$$\begin{aligned} a^2 &= \alpha^2 + \gamma^2 \\ b^2 &= \beta^2 + \delta^2 \\ c^2 &= \alpha\beta + \gamma\delta \end{aligned}$$

and with this we get the following reformulations as functions of α:

$$\begin{aligned}
\beta &= \frac{\alpha c^2}{a^2} \pm \sqrt{\left(\frac{\alpha c^2}{a^2}\right)^2 + b^2 - \left(\frac{\alpha^2 b^2}{a^2}\right) - \frac{c^4}{a^2}} \\
\gamma &= \frac{c^2 - \alpha\beta}{\sqrt{b^2 - \beta^2}} \\
\delta &= \sqrt{b^2 - \beta^2}.
\end{aligned} \tag{4.4}$$

Note here that it is also possible to define $\gamma = -\frac{c^2-\alpha\beta}{\sqrt{b^2-\beta^2}}$ and $\delta = -\sqrt{b^2-\beta^2}$, but then, because of symmetry, the different cases we have to check change analogously. Therefore without loss of generality we consider only the cases $\gamma = \frac{c^2-\alpha\beta}{\sqrt{b^2-\beta^2}}$ and $\delta = \sqrt{b^2-\beta^2}$.

At this point we have to think about the domains of α and β. From equality (4.3) we immediately get that:

$$|\alpha| \leq a \ \text{ and } \ |\beta| \leq b. \tag{4.5}$$

If these conditions are fulfilled, then the terms under the roots in (4.4) are nonnegative and we consider only real valued solutions, as desired.

It is clear from Definition 3.18 of $\pi_{s,p}$ that

$$\begin{aligned}
\pi_{s,p}(z) &\leq \left\|\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\right\|_p^2 + \left\|\begin{pmatrix}\gamma\\ \delta\end{pmatrix}\right\|_p^2 \\
&= \left(|\alpha|^p + |\beta|^p\right)^{2/p} + \left(|\gamma|^p + |\delta|^p\right)^{2/p}.
\end{aligned} \tag{4.6}$$

Therefore, without loss of generality the interesting cases of α, β, γ and δ are the following:

(i) $\alpha, \beta, \gamma, \delta \geq 0$

(ii) $\alpha \geq 0, \beta < 0, \gamma, \delta \geq 0$.

It is easy to see that those are all cases we have to consider: There are 16 possible combinations to choose α, β, γ and δ either nonnegative or strictly negative. In the following table we discuss all combinations and show that some of them are not suitable for our decomposition and the remaining cases are equivalent to case (i) or (ii):

Combination	α	β	γ	δ	**equivalent to case**
1	≥ 0	≥ 0	≥ 0	≥ 0	(i)
2	< 0	≥ 0	≥ 0	≥ 0	(ii)
3	≥ 0	< 0	≥ 0	≥ 0	(ii)
4	≥ 0	≥ 0	< 0	≥ 0	(ii)
5	≥ 0	≥ 0	≥ 0	< 0	(ii)
6	< 0	< 0	≥ 0	≥ 0	(i)
7	< 0	≥ 0	< 0	≥ 0	/
8	< 0	≥ 0	≥ 0	< 0	/
9	≥ 0	< 0	< 0	≥ 0	/
10	≥ 0	< 0	≥ 0	< 0	/
11	≥ 0	≥ 0	< 0	< 0	(i)
12	≥ 0	< 0	< 0	< 0	(ii)
13	< 0	≥ 0	< 0	< 0	(ii)
14	< 0	< 0	≥ 0	< 0	(ii)
15	< 0	< 0	< 0	≥ 0	(ii)
16	< 0	< 0	< 0	< 0	(i)

Table 4.1: Different cases

Four of them (combinations 7, 8, 9 and 10) are not suitable for our decomposition, e.g. combination 7 ($\alpha, \gamma < 0$ and $\beta, \delta \geq 0$) leads to $c^2 = \alpha\beta + \gamma\delta < 0$, which is a contradiction to $c^2 \geq 0$.
So there are 12 combinations left: The combinations 1, 6, 11 and 16 are equivalent to case (i). The other combinations (2, 3, 4, 5, 12, 13, 14 and 15) are all equivalent to case (ii): All of them describe the case that one vector of the decomposition (4.3) has only entries with the same sign and the other vector has entries with different signs. Switching the signs in the same vector (the negative to positive and vice versa) leads to the same decomposition and the same value of the sum of squared norms. Thus, Table 4.1 lists all possible combinations and hence it is enough to consider only the cases (i) and (ii). So without loss of generality, we can assume $\alpha \geq 0$ and $\gamma \geq 0$ for the rest of our calculations.

Note that in the case (i) we use only the positive root to calculate β, i.e.

$$\beta = \frac{\alpha c^2}{a^2} + \sqrt{\left(\frac{\alpha c^2}{a^2}\right)^2 + b^2 - \left(\frac{\alpha^2 b^2}{a^2}\right) - \frac{c^4}{a^2}}$$

and in case (ii) the negative one, i.e.

$$\beta = \frac{\alpha c^2}{a^2} - \sqrt{\left(\frac{\alpha c^2}{a^2}\right)^2 + b^2 - \left(\frac{\alpha^2 b^2}{a^2}\right) - \frac{c^4}{a^2}}.$$

Furthermore we have to ensure that $b^2 - \beta^2 \neq 0$ in order to have γ well defined.

At this point we want to discuss a possible necessary condition to find a decomposition for which the right-hand side of (4.6) gets minimal. An intuitive approach would be to calculate the derivative of the right-hand side of (4.6), set it equal to zero and solve the equation. We will show now that this is not as easy as it seems. First we have to calculate the derviative of the right-hand side in (4.6), which was done by Wolfram Mathematica, Version 10.3. Note that we consider only the case

$$\beta = \frac{\alpha c^2}{a^2} + \sqrt{\left(\frac{\alpha c^2}{a^2}\right)^2 + b^2 - \left(\frac{\alpha^2 b^2}{a^2}\right) - \frac{c^4}{a^2}}.$$

The case with the negative sign before the root works analogously.

Then the derivative in case (i) is the following:

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\alpha}\left((|\alpha|^p+|\beta|^p)^{2/p}+(|\gamma|^p+|\delta|^p)^{2/p}\right)\\
&=\frac{\mathrm{d}}{\mathrm{d}\alpha}\left((\alpha^p+\beta^p)^{2/p}+(\gamma^p+\delta^p)^{2/p}\right)\\
&=\frac{2}{p}\left(p\alpha^{p-1}+p\left(\frac{c^2}{a^2}+\frac{\frac{2c^4\alpha}{a^4}-\frac{2b^2\alpha}{a^2}}{2\sqrt{t(\alpha)}}\right)\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^{p-1}\right)\\
&\left(\alpha^p+\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^p\right)^{\frac{2}{p}-1}+\frac{2}{p}\left(\left(b^2-\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^2\right)^{\frac{p}{2}}\right.\\
&+\left.\left(\frac{c^2-\alpha\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)}{\sqrt{b^2-\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^2}}\right)^p\right)^{\frac{2}{p}-1}\left(p\left(\frac{c^2-\alpha\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)}{\sqrt{b^2-\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^2}}\right)^{p-1}\right.\\
&\left(\frac{-\frac{\alpha c^2}{a^2}-\alpha\left(\frac{c^2}{a^2}+\frac{\frac{2c^4\alpha}{a^4}-\frac{2b^2\alpha}{a^2}}{2\sqrt{t(\alpha)}}\right)-\sqrt{t(\alpha)}}{\sqrt{b^2-\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^2}}\right.\\
&+\left.\left(\frac{c^2}{a^2}+\frac{\frac{2c^4\alpha}{a^4}-\frac{2b^2\alpha}{a^2}}{2\sqrt{t(\alpha)}}\right)\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)\frac{\left(c^2-\alpha\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)\right)}{\left(b^2-\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^2\right)^{\frac{3}{2}}}\right)\\
&-p\left.\left(\frac{c^2}{a^2}+\frac{\frac{2c^4\alpha}{a^4}-\frac{2b^2\alpha}{a^2}}{2\sqrt{t(\alpha)}}\right)\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)\left(b^2-\left(\frac{\alpha c^2}{a^2}+\sqrt{t(\alpha)}\right)^2\right)^{\frac{p}{2}-1}\right),
\end{aligned}
$$

where $t(\alpha)=\frac{\alpha^2c^4}{a^4}-\frac{c^4}{a^2}+b^2-\frac{b^2\alpha^2}{a^2}$.

The derivative in case (ii) results analogously when replacing β by $-\beta$, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\left((|\alpha|^p+|\beta|^p)^{2/p}+(|\gamma|^p+|\delta|^p)^{2/p}\right)=\frac{\mathrm{d}}{\mathrm{d}\alpha}\left((\alpha^p+(-\beta)^p)^{2/p}+(\gamma^p+\delta^p)^{2/p}\right).$$

To calculate the minimum of (4.6) we have to solve

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\left((|\alpha|^p + |\beta|^p)^{2/p} + (|\gamma|^p + |\delta|^p)^{2/p}\right) = 0. \tag{4.7}$$

In Mathematica there are two possibilities to solve this equation: exactly with the command *Solve* or numerically with the command *NSolve*. None of them supplies a result of (4.7). Using the first command Mathematica said by itself that it can not solve the equation and using the second command, we decided to cancel the calculation after more than an hour. The problem here might be the following: The left-hand side of (4.7) can be converted into a polynomial. But the degree of this polynomial is too high to calculate its zeros via Mathematica since there exists no general algorithm to solve such problems.

Since it was not possible to compute the minimum of the right-hand side of (4.6) in the general case, we will consider three concrete examples, for which we calculate the minimum numerically for different choices of p. The numerical results are again computed by Wolfram Mathematica, Version 10.3.

Example 4.9. Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Then the trivial estimates in (4.5) give us $|\alpha| \leq \sqrt{2}$ and $|\beta| \leq \sqrt{2}$. Note that if $p = 1$, then the minimum of the right-hand side of (4.6) is not attained in case (ii). The reason for that is the following:

$$\begin{aligned} \alpha^2 + \gamma^2 &= 2 \\ \beta^2 + \delta^2 &= 2 \\ \alpha\beta + \gamma\delta &= 1. \end{aligned}$$

In case (ii), the last equation gives with the help of the triangle inequality

$$|\alpha\beta| + |\gamma\delta| \geq |\alpha\beta + \gamma\delta| = 1.$$

Since we consider the case (ii), we can assume without loss of generality $\alpha > 0$, since the case $\alpha = 0$ with $\beta < 0$, $\gamma, \delta \geq 0$ reduces to case (i) by switching the signs. Thus it follows that $\alpha\beta < 0$ and therefore $|\alpha\beta| + |\gamma\delta| > 1$. Hence for $p = 1$ we get:

$$\begin{aligned} (|\alpha| + |\beta|)^2 + (|\gamma| + |\delta|)^2 &= |\alpha|^2 + 2|\alpha||\beta| + |\beta|^2 + |\gamma|^2 + 2|\gamma||\delta| + |\delta|^2 \\ &> 2 + 2 + 2 \cdot 1 = 6. \end{aligned} \tag{4.8}$$

From the proof of Theorem 4.2, we know that $\pi_s(A) = \sum_{i,j=1}^{n} \left|A(i,j)\right|$ if $p = 1$. Therefore the minimum is equal to 6 and it can not be attained under the conditions of case (ii). Then for different values of p the minimum, which was calculated by Mathematica, is given by:

- $p = 1$: The minimum is equal to 6 and it is attained in case

    (i) for all $\alpha \in \left[0, \frac{1}{\sqrt{2}}\right]$

    (ii) /

    In the case $p = 1$ the minimum is attained for a nonengative decomposition. In this case it is also possible to exclude the decompositions that belong to case (ii) because of the special property (4.8) above.

- $p = 2$: The minimum is equal to 4 and it is attained in the given cases for all

    (i) $\alpha \in \left[0, \frac{1}{\sqrt{2}}\right]$

    (ii) $\alpha \in [0, 1.2233]$

    If $p = 2$, then the minimum is attained in each of the cases (i) and (ii). Therefore it is possible to consider just the nonnegative decompositions to determine the minimum confirming, what was shown in Theorem 4.2.

- $p = \frac{3}{2}$:

    (i) min = 4.4377 for $\alpha = 0.3667$

    (ii) min = 4.4945 for $\alpha = 0$

    For $p = \frac{3}{2}$ we get a nonnegative decomposition for which the minimum is attained: $\alpha = 0.3667, \beta = 1.3662, \gamma = 1.3659, \delta = 0.3654$.

- $p = 1.35$:

    (i) min = 4.7034 for $\alpha = 0.3667$

    (ii) min = 4.7656 for $\alpha = 0$

    For $p = 1.35$ the minimum is only attained for a single value of α in each case. Moreover there exists a nonnegative decomposition with $\alpha = 0.3667, \beta = 1.3662, \gamma = 1.3659, \delta = 0.3654$ for which the minimum is attained.

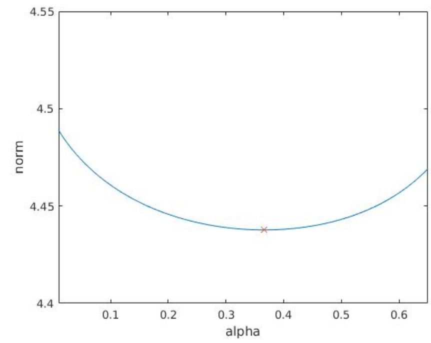


Figure 4.1: $p = \frac{3}{2}$, case (i)

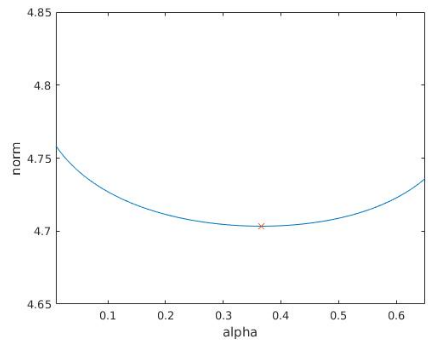


Figure 4.2: $p = 1.35$, case (i)

For each of our choices of p we receive a nonnegative decomposition of A which leads the minimal sum of squared norms as the graphics show. Furthermore in the special cases $p = 1$ and $p = 2$ the results illustrate the property we proved in Theorem 4.2 that the minimum is attained for every feasible decomposition. The other cases determine one nonnegative decomposition, where the sum of squared norms is minimal.

Example 4.10. Let $A = \begin{pmatrix} 3 & 1 \\ 1 & 1/2 \end{pmatrix}$. Then we know from (4.5) that $|\alpha| \leq \sqrt{3}$ and $|\beta| \leq \sqrt{1/2}$. Note that if $p = 1$, then the minimum of the right-hand side of (4.6) is not attained in case (ii). The argument for that is again:

$$\begin{aligned} \alpha^2 + \gamma^2 &= 3 \\ \beta^2 + \delta^2 &= \tfrac{1}{2} \\ \alpha\beta + \gamma\delta &= 1. \end{aligned}$$

As before the last equation gives via the triangle inequality:

$$|\alpha\beta| + |\gamma\delta| \geq 1.$$

Since we consider case (ii) and we assume again without loss of generality $\alpha > 0$, we have $\alpha\beta < 0$. Thus $|\alpha\beta| + |\gamma\delta| > 1$ and for $p = 1$ we get:

$$\begin{aligned} (|\alpha| + |\beta|)^2 + (|\gamma| + |\delta|)^2 &= |\alpha|^2 + 2|\alpha||\beta| + |\beta|^2 + |\gamma|^2 + 2|\gamma||\delta| + |\delta|^2 \\ &> 3 + \tfrac{1}{2} + 2 \cdot 1 = 5\tfrac{1}{2}. \end{aligned} \tag{4.9}$$

Then for different values of p the minimum, which was calculated by Mathematica, is given as follows:

- $p = 1$: The minimum is equal to 5.5 for all

  (i) $\alpha \in \left[0, \sqrt{2}\right]$

  (ii) /

  The minimum is attained for nonnegative decompositions if $p = 1$. The case (ii) can be ignored again because of the inequality (4.9).

- $p = 2$: The minimum is equal to 3.5 for all

  (i) $\alpha \in \left[0, \sqrt{2}\right]$

  (ii) $\alpha \in [0, 1]$

  Also here for $p = 2$ the minimum is attained for nonnegative decompositions in case (i).

- $p = \frac{3}{2}$:

  (i) min = 3.9542 for $\alpha = 0.1433$

  (ii) min = 3.9602 for $\alpha = 0$

  If $p = \frac{3}{2}$ the minimum is attained in case (i) for the nonnegative decomposition $\alpha = 0.1433, \beta = 0.4546, \gamma = 1.7261, \delta = 0.5416$.

- $p = 1.35$:

  (i) min = 4.2217 for $\alpha = 0.1433$

  (ii) min = 4.2283 for $\alpha = 0$

  If $p = 1.35$, then the minimum is only attained for a single value of α in each case. Moreover the decomposition is nonnegative, since we have $\alpha = 0.1433, \beta = 0.4546, \gamma = 1.7261, \delta = 0.5416$.

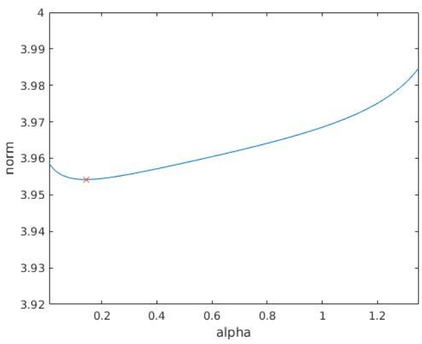


Figure 4.3: $p = \frac{3}{2}$, case (i)

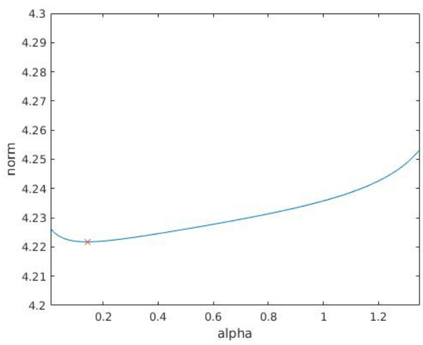


Figure 4.4: $p = 1.35$, case (i)

Also in this second example we obtain a nonnegative decomposition for every case, for which the sum of squared norms is minimal. Moreover the special cases $p = 1$ and $p = 2$ supply the expected result, i.e. that the minimum is attained for every nonnegative decomposition. In the other cases we determined a minimal decomposition as the graphics show.

Example 4.11. Let $A = \begin{pmatrix} 4 & 1 \\ 1 & 1/2 \end{pmatrix}$. Then we have by (4.5) that $|\alpha| \leq 2$ and $|\beta| \leq \sqrt{1/2}$. Note that if $p = 1$, then the minimum of the right-hand side of (4.6) is not attained in case (ii). The argument for that is again:

$$\begin{aligned} \alpha^2 + \gamma^2 &= 4 \\ \beta^2 + \delta^2 &= \tfrac{1}{2} \\ \alpha\beta + \gamma\delta &= 1. \end{aligned}$$

The last equation gives with the help of the triangle inequality:

$$|\alpha\beta| + |\gamma\delta| \geq 1.$$

Because of case (ii) and without loss of generality $\alpha > 0$, we have $\alpha\beta < 0$ as well as $|\alpha\beta| + |\gamma\delta| > 1$. Thus for $p = 1$ we get:

$$\begin{aligned} (|\alpha| + |\beta|)^2 + (|\gamma| + |\delta|)^2 &= |\alpha|^2 + 2|\alpha||\beta| + |\beta|^2 + |\gamma|^2 + 2|\gamma||\delta| + |\delta|^2 \\ &> 4 + \tfrac{1}{2} + 2 \cdot 1 = 6\tfrac{1}{2}. \end{aligned} \tag{4.10}$$

Then for different values of p the minimum, calculated by Mathematica, is given by:

- $p = 1$: The minimum is equal to 6.5 for all

  (i) $\alpha \in \left[0, \sqrt{2}\right]$

  (ii) /

  If $p = 1$, then the minimum is attained for nonnegative decompositions. Also here the case (ii) can be ignored because of (4.10).

- $p = 2$: The minimum is equal to 4.5 for all

  (i) $\alpha \in \left[0, \sqrt{2}\right]$

  (ii) $\alpha \in \left[0, \sqrt{2}\right]$

As before the minimum is attained for $p = 2$ and nonnegative decompositions.

- $p = \frac{3}{2}$:
  (i) min = 4.9231 for $\alpha = 0.1233$
  (ii) min = 4.9302 for $\alpha = 0$

  Again the minimum is attained for $p = \frac{3}{2}$ and a nonnegative decomposition with $\alpha = 0.1233, \beta = 0.5299, \gamma = 1.9962, \delta = 0.4682$.

- $p = 1.35$:
  (i) min = 5.1868 for $\alpha = 0.1233$
  (ii) min = 5.1949 for $\alpha = 0$

  If $p = 1.35$, then the minimum is again only attained for a single value of α in each case. Moreover the decomposition is nonnegative with $\alpha = 0.1233, \beta = 0.5299, \gamma = 1.9962, \delta = 0.4682$.

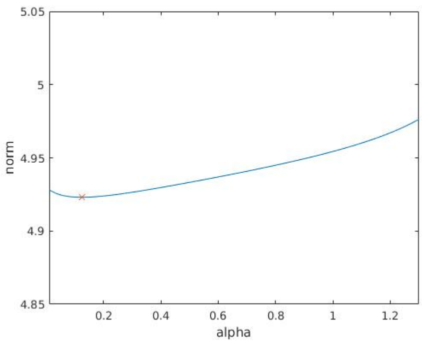


Figure 4.5: $p = \frac{3}{2}$, case (i)

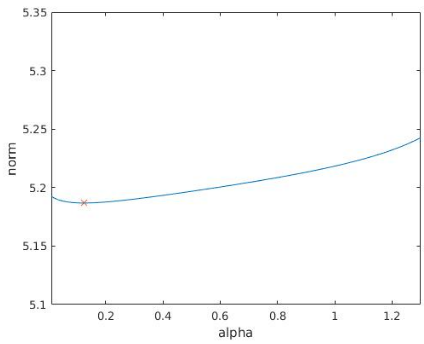


Figure 4.6: $p = 1.35$, case (i)

Also in this last example we were able to replicate the results from the examples before. Again for $p = 1$ and $p = 2$ the minimum was attained for every nonnegative decomposition and in the other cases we gave a nonnegative decomposition for which the sum of squared norms is minimal.

Conclusion

In this section we considered several matrix decompositions. Our goal was to get more information about the case $p \in (1, 2)$ and to justify our conjecture

(4.2). Here we discussed which of the decompositions would be suitable and restricted them to two different cases. Furthermore we calculated the derivative of the sum of squared norms of this special decomposition and equate it to zero to determine the minimum. Unfortunately Mathematica was not able to solve this problem analytically. For some numerical results we discussed three different examples and for every example the two different cases (i) and (ii). For each case we calculated the minimum of the sum of p-norms numerically.
For each example and each case of p we found a nonnegative decomposition for which the minimum is attained. In particular for $p = \frac{3}{2}$ and $p = 1.35$ there was a unique α for which the corresponding decomposition is nonnegative.
These examples strengthened our conjecture

$$\pi_{s,p}(z) \stackrel{?}{=} \inf\left\{\sum_{\nu=1}^{N} \|x_\nu\|_p^2 : z = \sum_{\nu=1}^{N} x_\nu \otimes x_\nu,\ x_\nu \geq 0, N \in \mathbb{N}\right\},$$

i.e. that the $\pi_{s,p}$-norm can be calculated by considering only the nonnegative decompositions if the matrix is completely positive for $p \in (1, 2)$.

4.3 Interpretation for duality

In this section we will discuss the meaning of the symmetric projective norm for our goal to generalize the topic of copositivity and completely positivity in infinite dimension. In the cases $p = 1$ and $p = 2$ it is much easier to calculate the projective norm, since the sum of squared norms is constant for all nonnegative decompositions and the minimum is attained.

For the sake of simplicity we will discuss the needed spaces only for vectors.

First we will consider the case $p = 1$, which leads to $\mathbb{R}^n$ equipped with the 1-norm $\|\cdot\|_1$. This space is equivalent to the space of Borel measures $M(\{1, \dots, n\})$ equipped with the total variation norm $\|\cdot\|_{tv}$. Note that two pairs (V, ϕ) and (W, ψ) are called *equivalent* if there is an isomorphism $f\colon V \to W$ such that $\psi = f \circ \phi$, cf. [29, VI. Chapter 3]. We use the symbol $\cong$ to denote the equivalence of two products; thus:

$$(\mathbb{R}^n, \|\cdot\|_1) \cong (M(\{1, \dots, n\}), \|\cdot\|_{tv}),$$

where the total variation norm is defined as follows:

Definition 4.12. [48, Section 5.1] Let $\mathcal{A}$ be a σ-algebra of subsets of a set Ω. A *partition* of $\Omega \in \mathcal{A}$ is any finite set $\{A_1, \dots, A_n\}$ of pairwise disjoint subsets of $\mathcal{A}$ satisfying $\bigcup_{i=1}^n A_i = \Omega$. If $\mu \colon \mathcal{A} \to X$ is a vector measure, i.e. a countable additive function with values in a Banach space X, then the *total variation norm* of μ is given by

$$\left\|\mu\right\|_{\mathrm{tv}} = |\mu|(\Omega) = \sup \left\{ \sum_{i=1}^{n} |\mu(A_i)| \colon \ \{A_1, \dots, A_n\} \text{ is a partition of } \Omega \right\}.$$

The dual space of $(\mathbb{R}^n, \|\cdot\|_1)$ is $(\mathbb{R}^n, \|\cdot\|_\infty)$. This space is equivalent to $(C(\{1, \dots, n\}), \|\cdot\|_\infty)$, i.e.

$$(\mathbb{R}^n, \|\cdot\|_\infty) \cong (C(\{1, \dots, n\}), \|\cdot\|_\infty).$$

The second case $p = 2$ leads to the space $\mathbb{R}^n$ equipped with the 2-norm $\|\cdot\|_2$. This space is equivalent the Hilbert space $L^2(m)$, where m denotes the counting measure, equipped with the same norm $\|\cdot\|_2$, i.e.

$$(\mathbb{R}^n, \|\cdot\|_2) \cong \left(L^2(m), \|\cdot\|_2\right).$$

A very fundamental property in this context is that this Hilbert space is selfdual, which is why we can consider the same space on the primal and on the dual side.

On that basis we will discuss in the next chapter a generalization of copositivity and completely positivity that results from the first case $p = 1$. After that we will discuss in Chapter 6 the generalizations of both cones that follows from second case $p = 2$.

Chapter 5

Copositivity for continuous kernels

In Chapters 4.1 and 4.3 we discussed that there are two practical cases ($p = 1$ and $p = 2$), in which the symmetric projective norm can be calculated comparatively easily. This leads us to the conclusion that in these two cases also a generalization of the concept of copositivity would be possible to be formulated in an appropriate way. In this chapter we will have a look at the first case $p = 1$ and the resulting generalization which deals with the copositive cone of continuous Hilbert-Schmidt kernels and the completely positive cone of symmetric, signed Radon measures. Furthermore we will give a representation of the completely positive cone by its extreme rays and provide a copositive formulation of the stable set problem. The theory which is discussed in this chapter is based on [24]. The formulated statements and their proofs can be found there or in the references given therein.

5.1 Theoretical results

Let V be a compact metric space with probability measure ω, which has to be strictly positive on open sets. Such a measure ω always exists for the following reason: If V is compact and metrizable, then V is also separable, see Proposition A.4. So V has a countable dense subset (see Definition A.2) and that is why there exists a sequence $(x_n)_{n\in\mathbb{N}}$ which lies dense in V. Furthermore we choose a sequence $a_n > 0$ with $\sum_{n\in\mathbb{N}} a_n = 1$ and also define $\omega(A) := \sum_{n\in\mathbb{N}} a_n \mathcal{X}_A(x_n)$ for any set $A \subseteq V$. Here $\mathcal{X}_T(x)$ denotes the indicator function $\mathcal{X}_T\colon V \supset T \to \{0,1\}$, which is equal to 1 if $x \in T$ and otherwise equal to 0. Thus ω is a probability measure which is strictly positive on open sets.

A natural generalization of finite $n\times n-$matrices are real-valued continuous Hilbert-Schmidt kernels on a compact Hausdorff space V. The set of continuous Hilbert-Schmidt kernels is defined as follows:

$$C(V \times V)_{\mathrm{sym}} = \{K\colon V \times V \to \mathbb{R}\colon K \text{ is symmetric and continuous}\}.$$

Here symmetry means that K is symmetric concerning its arguments, i.e. for all $x, y \in V$ we have $K(x, y) = K(y, x)$.

Analogously to the finite case we give the following definition of copositivity:

Definition 5.1. We name a kernel $K \in C(V \times V)_{\mathrm{sym}}$ *copositive* if

$$\int_V \int_V K(x,y) f(x) f(y)\, d\omega(x) d\omega(y) \geq 0 \quad \text{for all } f \in C(V)_{\geq 0},$$

where $C(V)_{\geq 0}$ denotes the convex cone of all nonnegative continuous functions on V. The convex *cone of copositive kernels* is denoted by $\mathcal{COP}_V$.

Furthermore we have to think about a generalization of duality in infinite dimension because we also want to find a representation of the dual cone of the cone of copositive kernels. The dual space of $C(V \times V)_{\mathrm{sym}}$ equipped with the supremum norm $\|\cdot\|_\infty$ consists of all continuous linear functionals equipped with the $1-$norm $\|\cdot\|_1$. By the Riesz' representation theorem A.5 the space of all continuous linear functionals equipped with the 1-norm $\|\cdot\|_1$ can be identified with the space of symmetric, signed Radon measures $M(V \times V)_{\mathrm{sym}}$ equipped with the total variation norm $\|\cdot\|_{\mathrm{tv}}$, see [27, 1.1 Definition] and Definition 4.12.

Therefore let K be a continuous kernel and let μ be a symmetric, signed Radon measure on $V \times V$. Then the duality is given by the pairing

$$\langle K, \mu\rangle = \int_{V\times V} K(x,y)\, d\mu(x,y). \tag{5.1}$$

Moreover we have to equip the spaces $C(V \times V)_{\mathrm{sym}}$ and $M(V \times V)_{\mathrm{sym}}$ with a topology, e.g. to define closure properties. Therefore we use the weakest topologies which are compatible with the pairing (5.1), which are the weak topology on $C(V \times V)_{\mathrm{sym}}$ and the weak* topology on $M(V \times V)_{\mathrm{sym}}$.

Moreover with (5.1) we can give the following definition:

Definition 5.2. The dual cone of the cone of copositive kernels is given by

$$\mathcal{CP}_V = \left\{ \mu \in M(V \times V)_{\mathrm{sym}} \colon \langle K, \mu \rangle \geq 0 \text{ for all } K \in \mathcal{COP}_V \right\},$$

which we call the *cone of completely positive measures*.

In the finite case a subset of the copositive cone is the cone of positive semidefinite matrices. In infinite dimension we have the following definition:

Definition 5.3. A kernel $K \in C(V \times V)_{\mathrm{sym}}$ is called *positive (semi-)definite* if

$$\int_V \int_V K(x, y) f(x) f(y)\, d\omega(x) d\omega(y) \geq 0 \quad \text{for all } f \in C(V),$$

where $C(V)$ denotes the space of continuous functions. We denote the set of positive semidefinite kernels by $\mathcal{S}_V^+$.

In [6, Lemma 1] Bochner proved that a kernel $K \in C(V \times V)_{\mathrm{sym}}$ is positive semidefinite if and only if for any choice of $x_1, \ldots, x_N$ of finitely many points in V, the matrix $(K(x_i, x_j))_{i,j=1}^N$ is positive semidefinite. This is a generalization of the finite characterization, cf. [5, Theorem 1.10 (c)].

In [24, Lemma 2.1] a similar property is shown for copositive kernels:

Lemma 5.4. *Let V be a compact space with probability measure ω which is strictly positive on open sets. A kernel $K \in C(V \times V)_{\mathrm{sym}}$ is copositive if and only if for any choice of finitely many points $x_1, \ldots, x_N \in V$, the matrix $(K(x_i, x_j))_{i,j=1}^N$ is copositive.*

Remark 5.5. An important aspect, which follows from Lemma 5.4, is that the set $\mathcal{COP}_V$ is independent of the choice of ω.

An alternative characterization of copositive kernels is the following, see [24, Lemma 2.2]:

Lemma 5.6. *A kernel $K \in C(V \times V)_{\mathrm{sym}}$ is copositive if and only if for any choice of finitely many points $x_1, \ldots, x_N$ in V, the sum $\sum_{i=1}^N \sum_{j=1}^N K(x_i, x_j)$ is nonnegative.*

Next we have a look at a representation of the cone of completely positive measures by its extreme rays, for which the proof can be found in [24, Proposition 2.3].

Proposition 5.7. *The cone of completely positive measures is equal to*

$$\mathcal{CP}_V = \text{cl cone}\left\{\sum_{i=1}^{N} a_i\delta_{x_i} \otimes a_i\delta_{x_i} : N \in \mathbb{N}, x_i \in V, a_i \geq 0\right\},$$

where the closure is taken with respect to the weak topology $\mathbb{R}^n \hat{\otimes}^s_{\pi_s} \mathbb{R}^n$ and by δ_x we denote the Dirac measure of x.*

Next we will show that the set of extreme rays of the completely positive cone consists of all product measures of the form $\mu \otimes \mu$, where μ is a nonnegative measure on V. For this purpose we have to formulate some helping properties first. The next two statements and their proofs can be found in [24, Lemma 2.5 and Proposition 2.6].

Lemma 5.8. *Let $\mathcal{B}$ be a compact set in a topological space such that $0 \notin \mathcal{B}$. Then the set $\mathcal{K}$ defined by the union $\mathcal{K} = \bigcup_{\lambda \geq 0} \lambda\mathcal{B}$ is closed.*

Proposition 5.9. *The set*

$$\mathcal{B} = \text{cl conv}\left\{\sum_{i=1}^{N} a_i\delta_{x_i} \otimes a_i\delta_{x_i} : N \in \mathbb{N}, x_i \in V, a_i \geq 0, \sum_{i=1}^{N} a_i = 1\right\}$$

is weak compact and the equality*

$$\mathcal{CP}_V = \bigcup_{\lambda \geq 0} \lambda\mathcal{B}$$

holds. Hence the extreme rays of $\mathcal{CP}_V$ are exactly the rays generated by the extreme points of $\mathcal{B}$.

Furthermore we need two results from Choquet theory [45, Proposition 1.5 and Chapter 3, Theorem]:

Theorem 5.10 (Milman's converse of the Krein-Milman theorem)**.** *Suppose that X is a compact convex subset of a locally convex space. Suppose further that $Z \subseteq X$, and that $X = \text{cl conv } Z$. Then the extreme points of X are contained in the closure of Z, i.e.* $\text{Ext}(X) \subseteq \text{cl}\, Z$.

Theorem 5.11 (Choquet)**.** *Suppose that X is a metrizable compact convex subset of a locally convex space E, an let $x_0 \in X$. Then there exists a probability measure P on X which represents x_0, i.e.*

$$u(x_0) = \int_X u(x)\, dP(x) \text{ for every continuous linear functional } u \text{ on } E,$$

and is supported by the extreme points of X, i.e. $P(X \setminus \text{Ext}(X)) = 0$.

With these properties we can now prove the following theorem concerning the extreme rays of the cone of completely positive measures, cf. [24, Theorem 2.4].

Theorem 5.12. *A measure generates an extreme ray of the cone $\mathcal{CP}_V$ of completely positive measures if and only if it is a product measure of the form $\mu \otimes \mu$, where $\mu \in M(V)$ is a nonnegative measure on V.*

Proof. First we define the following sets

$$\begin{aligned} \mathcal{M}_1^+(V) &:= \{\mu \in M(V) \colon \mu \geq 0, \mu(V) = 1\} \\ \mathcal{K}_1 &:= \left\{\mu \otimes \mu \colon \mu \in \mathcal{M}_1^+\right\}. \end{aligned}$$

We will prove that

$$\operatorname{Ext}(\operatorname{cl}\operatorname{conv}\mathcal{K}_1) = \mathcal{K}_1,$$

i.e. the extreme rays of the closed convex hull of $\mathcal{K}_1$ are equal to $\mathcal{K}_1$. Since $\mathcal{K}_1$ is weak* compact we get the first inclusion from Theorem 5.10:

$$\operatorname{Ext}(\operatorname{cl}\operatorname{conv}\mathcal{K}_1) \subseteq \mathcal{K}_1.$$

To prove the other inclusion, we assume that $\mu \otimes \mu \in \mathcal{K}_1$ can be represented as $\mu \otimes \mu = \frac{1}{2}(\nu_1 + \nu_2)$ for some $\nu_1, \nu_2 \in \operatorname{cl}\operatorname{conv}\mathcal{K}_1$, i.e. that $\mu \otimes \mu$ can be written as a convex combination of two product measures in $\operatorname{cl}\operatorname{conv}\mathcal{K}_1$. Furthermore $\mathcal{K}_1$ is weak* compact and weak* metrizable. Then by Theorem 5.11 there exist two probability measures P_1, P_2 on $\mathcal{M}_1^+$ such that for all $u \in \left(M(V \times V)_{\text{sym}}\right)'$ the following description holds:

$$u(\nu_i) = \int_{\mathcal{M}_1^+} u(\rho \otimes \rho)\, dP_i(\rho), \quad i = 1, 2.$$

Let $P := \frac{1}{2}(P_1 + P_2)$, then it follows that for all $F \in C(V \times V)_{\text{sym}}$

$$\begin{aligned} (\mu \otimes \mu)(F) &= \frac{1}{2}\left(\int_{\mathcal{M}_1^+} (\rho \otimes \rho)(F)\, dP_1(\rho) + \int_{\mathcal{M}_1^+} (\rho \otimes \rho)(F)\, dP_2(\rho)\right) \\ &= \int_{\mathcal{M}_1^+} (\rho \otimes \rho)(F)\, d\left(\frac{1}{2}(P_1 + P_2)\right)(\rho) \\ &= \int_{\mathcal{M}_1^+} (\rho \otimes \rho)(F)\, dP(\rho). \end{aligned}$$

According to our conditions V is a compact metrizable space and hence separable. Therefore the space $C(V)$ of continuous functions on V is separable too, cf. [49, p.33]. Consequently there exists a countable dense subset H of $C(V)_{\geq 0}$. Furthermore let $f \in H$ and let $\mathbf{1}_V$ be the constant function equal to 1 on V. Then we consider

$$F := \frac{1}{2}(f \otimes \mathbf{1}_V) + \frac{1}{2}(f \otimes \mathbf{1}_V).$$

Therefore we have

$$\begin{aligned} \mu(f) &= (\mu \otimes \mu)(F) \\ &= \int_{\mathcal{M}_1^+} (\rho \otimes \rho)(F)\, dP(\rho) \\ &= \int_{\mathcal{M}_1^+} \rho(f)\, dP(\rho). \end{aligned} \tag{5.2}$$

Analogously we set $F' = f \otimes f$ and get

$$\begin{aligned} \mu(f)^2 &= (\mu \otimes \mu)(F') \\ &= \int_{\mathcal{M}_1^+} (\rho \otimes \rho)(F')\, dP(\rho) \\ &= \int_{\mathcal{M}_1^+} \rho(f)^2\, dP(\rho). \end{aligned} \tag{5.3}$$

If $\mu(f) = 0$, then by (5.2) it follows that $\rho(f) = 0$ P-almost everywhere. Furthermore if $\mu(f) > 0$, we divide (5.2) and (5.3) by $\mu(f)$ and $\mu(f)^2$, respectively, and get the following equations:

$$\int_{\mathcal{M}_1^+} \frac{\rho(f)}{\mu(f)}\, dP(\rho) = 1 = \int_{\mathcal{M}_1^+} \frac{\rho(f)^2}{\mu(f)^2}\, dP(\rho).$$

Therefore there exists a set $N_f \subset \mathcal{M}_1^+$ with $P(N_f) = 0$ such that $\rho(f) = \mu(f)$ for all $\rho \in \mathcal{M}_1^+ \setminus N_f$. Furthermore we set $N = \bigcup_{f \in H} N_f$ and since H is countable we know that $P(N) = 0$ and $\rho(f) = \mu(f)$ for all $\rho \in \mathcal{M}_1^+ \setminus N$ and for all $f \in H$. Since H is dense in $C(V)_{\geq 0}$ it follows that $\rho = \mu$ for all $\rho \in \mathcal{M}_1^+ \setminus N$.

Because $0 \leq P_i(N) \leq 2P(N) = 0$ we get for $i = 1, 2$ and for all $F \in C(V \times V)_{\mathrm{sym}}$

that

$$\begin{aligned}\nu_i(F) &= \int_{\mathcal{M}_1^+\setminus N} (\rho\otimes\rho)(F)\,dP_i(\rho)\\ &= \int_{\mathcal{M}_1^+\setminus N} (\mu\otimes\mu)(F)\,dP_i(\rho)\\ &= (\mu\otimes\mu)(F)\int_{\mathcal{M}_1^+\setminus N} dP_i(\rho)\\ &= (\mu\otimes\mu)(F).\end{aligned}$$

Hence, $\nu_1 = \nu_2 = \mu\otimes\mu$, so $\mu\otimes\mu$ can be represented only by a convex combination of itself, i.e. $\mu\otimes\mu \in \operatorname{Ext}(\operatorname{cl}\operatorname{conv}\mathcal{K}_1)$ and the second inclusion is proved. Consider now $\operatorname{cl}\operatorname{conv}\mathcal{K}_1 = \mathcal{B}$ and with Proposition 5.9 the theorem follows. □

Remark 5.13. Theorem 5.12 also gives us another representation of the cone of completely positive measures by its extreme rays:

$$\mathcal{CP}_V = \operatorname{cl}\left\{\sum_{\nu=1}^{N}\mu_\nu\otimes\mu_\nu\colon \mu_\nu\in M(V), \mu_\nu\geq 0, \nu=1,\dots,N, N\in\mathbb{N}\right\}.$$

5.2 Duality

In this section we will have a look at two special applications of copositive optimization in infinite dimension. The first one is the stable set problem that determines the stability number. In [24, Chapter 3] the authors formulated the stable set problem of an infinite graph as a completely positive program and by duality this leads to a copositive program. For both problems strong duality holds. The second application we will consider in this section is the so-called kissing number problem, which can be illustrated as the stable set problem over a sphere. In [24, Chapter 4] the authors gave a copositive formulation of this problem too.

First we have to define a generalization of a graph in infinite dimension, cf. [24, Definition 1.1]:

Definition 5.14. Consider a graph whose vertex set is a Hausdorff topological space. An *open clique* is an open subset of the vertex set, where every two vertices are adjacent. The graph is called *topological packing graph* if each finite clique is contained in an open clique.

As a generalization of (2.6) in [24, Chapter 3] the authors gave the following copositive formulation of the maximum stable set problem: Let $G = (V, E)$ be a compact topological packing graph with metrizable vertex set. Furthermore let $\bar{E} = \{\{x, y\} : x \neq y, \{x, y\} \notin E\}$ and $D = \{(x, x) \colon x \in V\}$. Then the copositive program which supplies the stability number is the following:

$$\begin{aligned} \alpha(G) = \inf\ & t \\ \text{s.t. } & t \in \mathbb{R}, K \in \mathcal{COP}_V \\ & K(x, x) = t - 1 \qquad \text{for all } x \in V \\ & K(x, y) = -1 \qquad\ \ \text{for all } \{x, y\} \notin E. \end{aligned} \tag{5.4}$$

To receive the dual program the authors followed the arguments in [3, Chapter IV]. The problem (5.4) can be seen as a general conic problem of the form

$$\begin{aligned} & \inf \langle x, c \rangle_1 \\ & \text{s.t. } x \in \mathcal{K},\ Ax = b \end{aligned} \tag{5.5}$$

with the notations as follows:

$$\begin{aligned} & x = (t, K) \in \mathbb{R} \times C(V \times V)_{\text{sym}} \\ & c = (1, 0) \in \mathbb{R} \times M(V \times V)_{\text{sym}} \\ & \langle \cdot, \cdot \rangle_1 : \left(\mathbb{R} \times C(V \times V)_{\text{sym}}\right) \times \left(\mathbb{R} \times M(V \times V)_{\text{sym}}\right) \to \mathbb{R} \\ & \mathcal{K} = \mathbb{R}_{\geq 0} \times \mathcal{COP}_V \\ & A \colon \mathbb{R} \times C(V \times V)_{\text{sym}} \to C(V) \times C(\bar{E}) \\ & \quad A(t, K) = (x \mapsto K(x, x) - t, (x, y) \mapsto K(x, y)) \\ & b = (-1, -1) \in C(V) \times C(\bar{E}). \end{aligned}$$

Indeed, the condition $t \in \mathbb{R}$ can be easily replaced by $t \in \mathbb{R}_{\geq 0}$ in (5.4) since $t \geq 0$ holds automatically because diagonal elements of copositive kernels are nonnegative. The dual problem of (5.5) has the following form:

$$\begin{aligned} \sup\ & \langle b, y \rangle_2 \\ & c - A^* y \in \mathcal{K}^*. \end{aligned}$$

We use the following notation:

$$\begin{aligned} & \langle \cdot, \cdot \rangle_2 : (C(V) \times C(\bar{E})) \times (M(V) \times M(\bar{E})) \to \mathbb{R} \\ & y = (\mu_0, \mu_1) \in M(V) \times M(\bar{E}) \\ & A^* \colon M(V) \times M(\bar{E}) \to \mathbb{R} \times M(V \times V)_{\text{sym}} \\ & \quad A^*(\mu_0, \mu_1) = (-\mu_0(V), \mu_0 + \mu_1) \\ & \mathcal{K}^* = \mathbb{R}_{\geq 0} \times \mathcal{CP}_V. \end{aligned}$$

Note that the map A^* is the adjoint of A because

$$\begin{aligned}\langle A(t,K),(\mu_0,\mu_1)\rangle_2 &= \int_V K(x,x) - t\,d\mu_0(x) + \int_{\bar{E}} K(x,y)\,d\mu_1(x,y)\\ &= -t\mu_0(V) + \int_{V\times V} K(x,y)\,d(\mu_0+\mu_1)(x,y)\\ &= \langle (t,k), A^*(\mu_0,\mu_1)\rangle_1 .\end{aligned}$$

Thus, we receive the dual of (5.4), which is a completely positive program:

$$\begin{aligned} \sup\ & \mu(V\times V)\\ \text{s.t. }\ & \mu \in \mathcal{CP}_V\\ & \mu(D) = 1\\ & \operatorname{supp}\mu \subseteq D\cup \bar{E}, \end{aligned} \tag{5.6}$$

where $\mu = -\mu_0 - \mu_1$ and the support of a measure μ is given by

$$\operatorname{supp}\mu = (V\times V)\setminus O$$

and

$$O = \bigcup_{\substack{W \text{ open in } V\times V\\ \mu(W)=0}} W.$$

Furthermore the authors were able to prove that the optimal value of the completely positive program (5.6) is equal to the stability number too:

Theorem 5.15. *Let $G = (V,E)$ be a compact topological packing graph. Then the optimal value of the completely positive program (5.6) is attained and equals $\alpha(G)$.*

In [24, Theorem 3.2] they were also able to prove that strong duality holds for (5.4) and (5.6):

Theorem 5.16. *There is no duality gap between the primal copositive program (5.4) and the dual completely positive program (5.6). In particular the optimal value of both programs equals $\alpha(G)$.*

5.3 Application

A special application of the maximum stable set problem is the so-called kissing number problem. The kissing number in $\mathbb{R}^n$ is the maximum number τ_n of unit spheres that can be arranged simultaneously around a unit sphere in an n-dimensional space without pairwise overlapping. A extensive overview over the known kissing numbers so far and the development of this topic can be found in [2] and the references given therein. The following facts can be found in [2, Introduction]. Up to now the values of τ_n are only known for $n = 1, 2, 3, 4, 8, 24$. For the other values of n only lower and upper bounds are known. It is not difficult to identify the kissing number for $n = 1$ and $n = 2$, where the centers of the spheres are arranged in a line and in a plane, respectively. It becomes more difficult for larger n. For $n = 3$ the kissing number is $\tau_3 = 12$. Furthermore in this case the points can be moved around obtaining infinitely many new suitable configurations apart from the configurations that arise by orthogonal transformation. In dimension $n = 4$ it is known that $\tau_4 = 24$, but a proof of the uniqueness up to orthogonal transformations of the optimal configuration of points is still an open problem. Moreover the known kissing numbers in higher dimensions are $\tau_8 = 240$ and $\tau_{24} = 196560$. In contrast to dimension 4, in these cases it is known that the optimal configuration of points of the spheres is unique up to orthogonal transformations. Additionally the difference between the lower and upper bounds can be very big, e.g. for $n = 5$ the lower bound is equal to 40 and the upper bound is equal to 44, while the bounds for $n = 20$ are 17400 and 36764. Moreover in [2] an approach by semidefinite programming is used to verify the known kissing numbers ($n = 3, 4, 8, 24$) and to calculate new upper bounds for the kissing number in dimension $n = 5, 6, 7, 9, 10, 13, 14, 15, 16, 17$.

In [24, Chapter 4] the authors gave the following copositive formulation of the kissing number problem:

$$
\begin{aligned}
&\inf t \\
&\text{s.t. } t \in \mathbb{R}, K \in \mathcal{COP}_{S^{n-1}} \\
&\qquad K(x,x) = t - 1 \qquad \text{for all } x \in S^{n-1} \\
&\qquad K(x,y) = -1 \qquad\;\; \text{for all } x, y \in S^{n-1} \text{ with } x^\top y \in [-1, 1/2]\,,
\end{aligned}
\tag{5.7}
$$

where $S^{n-1} = \{x \in \mathbb{R}^n : x^\top x = 1\}$ denotes the n-dimensional unit sphere.

In [24, Chapter 4] they also showed that the copositive program (5.7) can be approximated by a sequence of semi-infinite linear programs, i.e. programs with a finite number of variables and an infinite number of constraints. Their reasoning is the following: The packing graph is invariant under the orthogonal group. Therefore the copositive formulation is invariant under this group too. By convexity the copositive formulation (5.7) can be restricted to copositive kernels which are invariant under the orthogonal group. Thus $K(x, y)$ depends only on the inner product $x^\top y$.

With the Theorem of Stone-Weierstrass it follows that polynomials lie dense in $C([-1, 1])$, cf. [49, Satz I.2.10 and Satz VIII.4.7]. Therefore we can approximate $K(x, y)$ by $\sum_{k=0}^{d} c_k (x^\top y)^k$. Then, by Lemma 5.6 the copositivity condition $K \in \mathcal{COP}_{S^{n-1}}$ translates to

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=0}^{d} c_k (x_i^\top x_j)^k \geq 0 \text{ for all } N \in \mathbb{N} \text{ and } x_1, \ldots, x_N \in S^{n-1}. \tag{5.8}$$

The other constraints of (5.7) change in the same way and by considering that $x^\top x = 1$ for $x \in S^{n-1}$, the following semi-infinite linear program results. Its optimal value converges to the kissing number if the degree d tends to infinity.

$$\begin{aligned}
\inf\ & 1 + \sum_{k=0}^{d} c_k \\
& c_0, \ldots, c_d \in \mathbb{R} \\
& \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=0}^{d} c_k (x_i^\top x_j)^k \geq 0 \text{ for all } N \in \mathbb{N} \text{ and } x_1, \ldots, x_N \in S^{n-1} \\
& \sum_{k=0}^{d} c_k s^k \leq -1 \text{ for all } s \in [-1, 1/2]\,.
\end{aligned}$$

The condition $\sum_{k=0}^{d} c_k s^k = -1$ is relaxed to $\sum_{k=0}^{d} c_k s^k \leq -1$ to make the problem feasible for finite degree d. This relaxation does not affect the optimal value if d goes to infinity.

Further on all the difficulty of the problem lies in constraint (5.8). In contrast to this, the other constraint $\sum_{k=0}^{d} c_k s^k \leq -1$ for all $s \in [-1, 1/2]$ is quite easy. Although there are infinitely many linear conditions on the coefficients c_k, this can be modeled equivalently as a semidefinite constraint using the sum of squares technique for polynomial optimization, cf. [43, 37].

5.4 Approximation hierarchies

Based on the results of [24], Kuryatnikova and Vera discussed in [36] different ways to approximate the copositive cone $\mathcal{COP}_V$. Therein the authors generalized the approaches from Peña et al. [44] and de Klerk and Pasechnik [16] for an inner approximation of the copositive cone from finite to infinite dimension. Note that also in infinite dimension the easiest way to approximate $\mathcal{COP}_V$ is to replace it by $\mathcal{S}_V^+$. Derived from Chapter 2.4 we will present here the generalizations of $\mathcal{C}_n^r$ and $\mathcal{Q}_n^r$. To that purpose we repeat the definitions:

$$\mathcal{C}_n^r = \left\{ A \in \mathcal{S}_n \colon P_A(x) \cdot \left(\sum_{i=1}^n x(i)^2 \right)^r \text{ has nonnegative coefficients} \right\},$$

where $P_A(x) = \sum_{i=1}^n \sum_{j=1}^n A(i,j)x(i)^2 x(j)^2$ for a given symmetric matrix A. Furthermore we had

$$\mathcal{E}_n^r = \left\{ \sum_{\substack{\beta \in \mathbb{N}^n, \\ |\beta| = r}} x^\beta x^\top (S_\beta + N_\beta) x \colon S_\beta \in \mathcal{S}_n^+, N_\beta \in \mathcal{N}_n \right\}$$

$$\mathcal{Q}_n^r = \left\{ A \in \mathcal{S}_n^+ \colon x^\top A x \left(\sum_{i=1}^n x(i)^2 \right)^r \in \mathcal{E}_n^r \right\}.$$

To approximate the kissing number, Kuryatnikova and Vera [36] extended the definitions of $\mathcal{C}_n^r$ and $\mathcal{Q}_n^r$. Since they considered a special kind of tensors which are used in numerical applications we will only present a short summary of their generalizations and results. Of course we will adjust our notation to their theory. For $F \in C(V^{d+2})$ and $v = (v_1, \ldots, v_d) \in V^d$ let $F^v = F(:, :, v_1, \ldots, v_d)$ be the 2-slice of F obtained by fixing all but the first two variables. Moreover a function $F \in C(V^{d+2})$ is called 2-*psd* if F^v is positive definite for all $v \in V^d$.

We set P_d as the group of permutations on d elements. Then for any $d \in \mathbb{N}$ we define $\theta \colon C(V^d) \to C(V^d)$ as the symmetrization (Reynolds) operator:

$$\theta(F) := \frac{1}{d!} \sum_{\eta \in P_d} F(v^\eta) \quad \forall v \in V^d,$$

where $v^\eta = (v_{\eta(1)}, \ldots, v_{\eta(d)})$ for a permutation $\eta \in P_d$. If $d \in \mathbb{N}$, then for any $r \in \mathbb{N}$, $v \in V^d$ the lifting operator $\cdot^{\oplus r} \colon C(V^d) \to C(V^{d+r})$ is defined as

$$F(v) \mapsto F^{\oplus r}(v, u_1, \ldots, u_r) \text{ for all } u_1, \ldots, u_r \in V.$$

Then the authors introduced the following sets:

$$\begin{aligned}
\mathcal{C}_V^r &:= \left\{K \in \mathcal{C}(V \times V)_{\text{sym}} \colon \theta(K^{\oplus r}) = \theta(T),\ T \in \mathcal{C}(V^{r+2})_{\geq 0}\right\}, \\
\mathcal{Q}_V^r &:= \left\{K \in \mathcal{C}(V \times V)_{\text{sym}} \colon \theta(K^{\oplus r}) = \theta(S) + \theta(T), \right. \\
&\qquad \left. S, T \in \mathcal{C}(V^{r+2}),\ T \in \mathcal{C}(V^{r+2})_{\geq 0},\ S \text{ is 2-psd}\right\},
\end{aligned}$$

where $\mathcal{C}(V^d)_{\geq 0} := \left\{F \in \mathcal{C}(V^d) \colon F(v) \geq 0,\ \forall v \in V^d\right\}$ denotes the set of non-negative continuous functions on V^d. In particular if $V = \{1, \dots, n\}$, then $\mathcal{C}_V^r = \mathcal{C}_n^r$ and $\mathcal{Q}_V^r = \mathcal{Q}_n^r$ for any r. Furthermore the authors were able to prove for $V \subseteq \mathbb{R}^n$ compact that

$$\begin{aligned}
\mathcal{C}_V^0 \subseteq \mathcal{C}_V^1 \subseteq \ldots \subseteq \mathcal{COP}_V, \\
\mathcal{Q}_V^0 \subseteq \mathcal{Q}_V^1 \subseteq \ldots \subseteq \mathcal{COP}_V.
\end{aligned}$$

Define a kernel K as *strictly copositive* if there exists an $\varepsilon > 0$ such that for any choice of $v_1, \dots, v_n \in V$ we have

$$\min_{x \in \Delta} \sum_{i=1}^{n} \sum_{j=1}^{n} K(v_i, v_j) \geq \varepsilon,$$

where $\Delta = \{x \in \mathbb{R}^n \colon e^\top x = 1, x \geq 0\}$ denotes the standard simplex in $\mathbb{R}^n$. The set of strictly copositive kernels is denoted by $\mathcal{COP}_V^+$.

For this generalization of the interior of the copositive cone the authors were also able to prove that

$$\mathcal{COP}_V^+ \subseteq \bigcup_{r \in \mathbb{N}_0} \mathcal{C}_V^r \subseteq \bigcup_{r \in \mathbb{N}_0} \mathcal{Q}_V^r.$$

Apart from these convergence results they have also shown that the bounds obtained by replacing $\mathcal{COP}_V$ by $\mathcal{C}_V^r$ or $\mathcal{Q}_V^r$ converge to the stability number $\alpha(G)$ of a compact topological packing graph $G = (V, E)$. With this techniques and $V = S^{n-1}$ they were able verify the previously known kissing numbers and give new lower and upper bounds for the unknown kissing numbers.

For the sake of completeness we have to mention here that De Laat and Vallentin presented in [17] the hierarchy las_r^* which is based on Lasserre's hierarchy for finite graphs [38]. This hierarchy can be seen as a generalization of $\mathcal{K}_n^r$. A description of these concepts in this thesis would go to far and we refer to [17, Section 3].

In this chapter we discussed copositivity for Hilbert-Schmidt kernels and completely positivity for symmetric, signed Radon measures. Besides the definition, we gave some interesting results in infinite dimension and in particular the proof of extreme rays of the cone of completely positive measures as an essential aspect. Furthermore we presented as an applications the generalization of the maximum stable set problem and the kissing number problem. All of these results were obtained from [24]. Last we cited the approximation hierarchies by Kuryatnikova and Vera [36]. In the next chapter we will discuss what happens in the special case $p = 2$ and consider the copositive as well as the completely positive cone in the selfdual Hilbert space $L^2(\mu)$.

Chapter 6

Copositivity in Hilbert spaces

In this chapter we will discuss a generalization of the copositive and the completely positive cone as well as their characteristics for the case $p = 2$ which leads to the space $L^2(\mu)$, where μ is a σ-finite measure. We will formulate and prove important properties of these cones and compare them with the theory in finite dimension and the results from the case $p = 1$, respectively. Furthermore we will give impressions of the difficulties that result in this kind of generalization.

6.1 The copositive cone

First we will give a definition of the copositive cone concerning the space $L^2(\mu)$. For this purpose, we will repeat the definition of the copositive cone concerning $C(V)$ from Definition 5.1 and show the connection to the new cone. Let E be either the space $C(V)$ of all continuous real valued functions on a metrizable compact space V or $E = L^p(\mu)$, with $1 \leq p < +\infty$, for some $\sigma-$finite measure μ on some measurable space $(V, \mathcal{A})$, where $\mathcal{A}$ is assumed to be countably generated. Note that $E = C(V)$ leads to the theory which we discussed in Chapter 5 and $E = L^2(\mu)$ leads to a new case, which we will consider in this chapter. We remark that it is possible to give the following definitions and to prove some of the results in the much more general context of Riesz-spaces, but for our purpose the above spaces are more than sufficient.

We will introduce in this chapter the copositive cone concerning $L^2(\mu)$, thus we denote from now on $\mathcal{COP}_V =: \mathcal{COP}_{C(V)}$ to get a more noticeable distinction between both cones. In [24, Inequality (2)] the cone $\mathcal{COP}_{C(V)}$ of copositive kernels on a metrizable compact space V is defined as the cone of all symmetric continuous functionals F on $V \times V$ such that

$$\int_V \int_V F(x,y)f(x)f(y)\,d\omega(x)d\omega(y) \geq 0$$

for all $f \in C(V)$ and $f \geq 0$. There ω is a regular measure on V such that $\omega(O) > 0$ for all $\emptyset \neq O \subset V$ is open. We remark that $\mathcal{COP}_{C(V)}$ is independent of such a measure ω. It is clear that $\mathcal{COP}_{C(V)}$ equals to the cone of all symmetric continuous functions F on $V \times V$ such that

$$\int_V \int_V F(x,y)\,d\mu(x)d\mu(y) \geq 0$$

for all regular Radon measure $\mu \geq 0$ on V, cf. (5.1), Remark 5.13 and [24]. Expressed in the language of the symmetric tensor product, $\mathcal{COP}_{C(V)}$ is the cone of all $F \in C(V)\hat{\otimes}^s_{\varepsilon_s} C(V)$ such that

$$\langle F, \mu \otimes \mu \rangle \geq 0$$

for all $\mu \in C(V)' = M(V)$ and $\mu \geq 0$. This is what we now take as a general definition:

Definition 6.1. Let $E = L^p(\mu)$ for $1 \leq p < +\infty$ or $E = C(V)$. Then we set

$$\mathcal{COP}_E := \left\{ z \in E\hat{\otimes}^s_{\varepsilon_s} E \colon \langle z, x' \otimes x' \rangle \geq 0 \text{ for all } x' \in E', x' \geq 0 \right\}.$$

The copositive cone $\mathcal{COP}_{C(V)}$ regarding the continuous functions over V and its properties were discussed in Chapter 5. In this chapter we will now have a closer look at $\mathcal{COP}_{L^2(\mu)}$:

$$\mathcal{COP}_{L^2(\mu)} = \left\{ z \in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu) \colon \langle z, x' \otimes x' \rangle \geq 0 \text{ for all } x' \in L^2(\mu)', x' \geq 0 \right\}.$$

If $z \in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)$, then there is a unique compact self-adjoint operator $T = T_z$ such that

$$\langle Tf, f \rangle_{L^2(\mu)} = \langle z, f \otimes f \rangle_{L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)}\,.$$

With this identification the cone $\mathcal{COP}_{L^2(\mu)}$ can be considered as the cone of all self-adjoint compact operators T on $L^2(\mu)$ such that $\langle Tf, f \rangle_{L^2(\mu)} \geq 0$ for all $f \in L^2(\mu)$ and $f \geq 0$.

In contrast to the characteristics in finite dimension in Proposition 2.3, we can prove the following theorem:

Theorem 6.2. *The copositive cone $\mathcal{COP}_{L^2(\mu)}$ is a closed convex and pointed cone. Furthermore it has no interior points if* $\dim(L^2(\mu)) = +\infty$.

Proof. The copositive cone in $L^2(\mu)\hat{\otimes}^s_{\varepsilon_s}L^2(\mu)$ is the dual cone of the set $\operatorname{span}\{f\otimes f\colon f\in L^2(\mu), f\geq 0\}$ and hence convex and closed. To prove that $\mathcal{COP}_{L^2(\mu)}$ is pointed we assume that there is a $z\in\mathcal{COP}_{L^2(\mu)}$ such that also $-z\in\mathcal{COP}_{L^2(\mu)}$. Then $\langle z, x'\otimes x'\rangle = 0$ for all $x'\in L^2(\mu)'$. Thus $\varepsilon_s(z) = 0$ and hence $z = 0$. Last we prove that $\mathcal{COP}_{L^2(\mu)}$ has no interior point. For this, let $(V,\mathcal{A},\mu)$ be a σ-finite measure space with $\dim L^2(\mu) = +\infty$. We assume that z is an interior point of $\mathcal{COP}_{L^2(\mu)}$. Then there exists an $\varepsilon > 0$ such that $z - w\in\mathcal{COP}_{L^2(\mu)}$ for all $w\in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s}L^2(\mu)$ with $\varepsilon_s(w)\leq\varepsilon$. Additionally we denote $z = \sum_{\nu=0}^{\infty}\lambda_\nu f_\nu\otimes f_\nu$ with an orthonormal basis $(f_\nu)_{\nu\in\mathbb{N}_0}$ and $\lambda_\nu\in\mathbb{R}$ with $\lambda_\nu\to 0$ as $\nu\to\infty$. Then there exists an $n\in\mathbb{N}$ with $|\lambda_\nu| < \varepsilon/2$ for all $\nu\geq n$. Furthermore for $k\in\mathbb{N}$ let $w_k\in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s}L^2(\mu)$ be given by $w_k = \sum_{\nu=n}^{\infty}\lambda_\nu f_\nu\otimes f_\nu + \frac{\varepsilon}{2}\sum_{\nu=n}^{k} f_\nu\otimes f_\nu$. Thus $z - w_k\in\mathcal{COP}_{L^2(\mu)}$ and for $f\in L^2(\mu)$ it is valid that

$$\begin{aligned}\langle z - w_k, f\otimes f\rangle &= \sum_{\nu=0}^{n-1}\lambda_\nu\langle f, f_\nu\rangle^2 - \frac{\varepsilon}{2}\sum_{\nu=n}^{k}\langle f, f_\nu\rangle^2\\ &\leq \max_{0\leq k\leq n-1}|\lambda_k|\sum_{\nu=0}^{n-1}\langle f, f_\nu\rangle^2 - \frac{\varepsilon}{2}\sum_{\nu=n}^{k}\langle f, f_\nu\rangle^2 .\end{aligned}$$

Denote $\max_{1\leq k\leq n-1}|\lambda_k| =: s$. To construct $k\in\mathbb{N}$ and $f\in L^2(\mu)$ with $\langle z - w, f\otimes f\rangle < 0$ it is sufficient to find an $f\geq 0$ such that

$$s\sum_{\nu=0}^{n-1}\langle f, f_\nu\rangle^2 - \frac{\varepsilon}{2}\sum_{\nu=n}^{\infty}\langle f, f_\nu\rangle^2 < 0.$$

We assume that such an f does not exist. Then for all $f\in L^2(\mu)$ with $f\geq 0$ we have

$$\begin{aligned}\|f\|_{L^2(\mu)} &= \left(\sum_{\nu=0}^{\infty}\langle f, f_\nu\rangle^2\right)^{\frac{1}{2}}\\ &\leq \left(1 + \frac{2s}{\varepsilon}\right)^{\frac{1}{2}}\left(\sum_{\nu=0}^{n-1}\langle f, f_\nu\rangle^2\right)^{\frac{1}{2}} .\end{aligned}$$

We pick $X_k\in\mathcal{A}$ pairwise disjoint with $\mu(X_k) > 0$ and $\bigcup_k X_k = V$. For the existence see [4, Chapter 7, Ex. 1.3]. Set $f_k := \frac{1}{\mu(X_k)^{1/2}}\mathcal{X}_{X_k}\geq 0$, where $\mathcal{X}_A$

denotes again the indicator function. Then $\|f_k\|_{L^2(\mu)} = 1$ holds and for all $g \in L^2(\mu)$ by the Cauchy-Schwarz inequality we have

$$|\langle f_k, g\rangle_{L^2(X_k)}| \leq \|f_k\|_{L^2(X_k)} \|g\|_{L^2(X_k)}.$$

Since $\|g\|^2_{L^2(\mu)} = \sum_{k\in\mathbb{N}} \|g\|^2_{L^2(X_k)}$ it follows that $\|g\|_{L^2(X_k)} \to 0$ for $k \to \infty$. Therefore

$$\langle f_k, g\rangle_{L^2(\mu)} = \langle f_k, g\rangle_{L^2(X_k)} \to 0 \text{ for } k \to \infty.$$

Indeed, the equality holds by the definition of f_k. So in particular

$$\langle f_k, f_\nu\rangle \to 0 \text{ for } k \to \infty \text{ and } 0 \leq \nu \leq n.$$

Thus we have

$$\begin{aligned} 1 &= \|f_k\|_{L^2(\mu)} \\ &\leq \left(1 + \frac{2s}{\varepsilon}\right)^{\frac{1}{2}} \left(\sum_{\nu=0}^{n-1} \langle f_k, f_\nu\rangle^2\right)^{\frac{1}{2}} \\ &\to 0 \text{ for } k \to \infty, \end{aligned}$$

which is a contradiction. Therefore there exists an $f \in L^2(\mu)$ with $f \geq 0$ such that $\langle z - w_k, f \otimes f\rangle < 0$. But this is a contradiction to $z - w_k \in \mathcal{COP}_{L^2(\mu)}$. Thus $z - w \notin \mathcal{COP}_{L^2(\mu)}$ for all $w \in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)$ with $\varepsilon_s(w) \leq \varepsilon$. Consequently z can not be an interior point of $\mathcal{COP}_{L^2(\mu)}$ and the proof is complete. □

Remark 6.3. In contrast to $\mathcal{COP}_{L^2(\mu)}$ the cone $\mathcal{COP}_{C(V)}$ has a nonempty interior, for example $\mathbf{1}_{V\times V}$ is an interior point of $\mathcal{COP}_{C(V)}$.

Analogous to Lemma 5.4 we can formulate and prove the following property:

Lemma 6.4. *Let $z \in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)$. Then $z \in \mathcal{COP}_{L^2(\mu)}$ if and only if for all decompositions $\mathcal{E} = \{E_1, \ldots, E_n\}$ of V with $E_i \cap E_j = \emptyset$ $(i \neq j)$ and $0 < \mu(E_i) < +\infty$, $1 \leq j \leq n$ holds that $(\langle z, \mathcal{X}_{E_\nu} \otimes \mathcal{X}_{E_\iota}\rangle)_{1\leq\nu,\iota\leq n}$ is copositive.*

Proof. Let $z \in \mathcal{COP}_{L^2(\mu)}$ with $z = \sum_{k=0}^{\infty} \varrho_k h_k \otimes h_k$ and let $h_0, h_h, \ldots$ be an orthonormal system in $L^2(\mu)$. Then we have:

$$\begin{aligned}
\langle z, \mathcal{X}_{E_\nu} \otimes \mathcal{X}_{E_\iota} \rangle &= \sum_{k=0}^{\infty} \varrho_k \langle h_k, \mathcal{X}_{E_\nu} \rangle \langle h_k, \mathcal{X}_{E_\iota} \rangle \\
&= \sum_{k=0}^{\infty} \varrho_k \int_{E_\nu} h_k \, d\mu \int_{E_\iota} h_k \, d\mu \\
&= \sum_{\nu=1}^{n} \sum_{\iota=1}^{n} \langle z, \alpha_\nu \mathcal{X}_{E_\nu} \otimes \alpha_\iota \mathcal{X}_{E_\iota} \rangle \\
&= \left\langle z, \sum_{\nu=1}^{n} \sum_{\iota=1}^{n} \alpha_\nu \alpha_\iota \mathcal{X}_{E_\nu} \otimes \mathcal{X}_{E_\iota} \right\rangle \\
&= \left\langle z, \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \otimes \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \right\rangle \\
&\geq 0, \quad \text{for all } \alpha_\nu \geq 0.
\end{aligned}$$

Thus $(\langle z, \mathcal{X}_{E_\nu} \otimes \mathcal{X}_{E_\iota} \rangle)_{1 \leq \nu, \iota \leq n}$ is copositive.
Conversely, let $f \in L^2(\mu)$ and $f \geq 0$. For $\varepsilon > 0$ choose $E_1, \ldots, E_n$ pairwise disjoint with $0 < \mu(E_j) < +\infty$ for all j and $\alpha_1, \ldots, \alpha_n \geq 0$ such that $\left\| f - \sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right\|_2 < \varepsilon$. Then:

$$\begin{aligned}
&\pi \left(f \otimes f - \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \otimes \sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \right) \\
\leq &\pi \left(f \otimes \left(f - \sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) + f \otimes \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) - \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \otimes \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \right) \\
\leq &\|f\| \, \varepsilon + \varepsilon \left(\|f\| + \varepsilon \right).
\end{aligned}$$

Therefore for $\varepsilon > 0$ and $f \in L^2(\mu)$ with $f \geq 0$ we can choose $E_1, \ldots, E_n$ measurable and $\alpha_1, \ldots, \alpha_n \geq 0$ such that

$$\pi_s \left(f \otimes f - \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \otimes \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \right) < \varepsilon.$$

Recall that in $L^2(\mu)$ holds: $\pi(z) = \pi_s(z)$ for all z. Hence:

$$\begin{aligned}
&\langle z, f \otimes f \rangle \\
= &\left\langle z, f \otimes f - \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \otimes \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \right\rangle + \left\langle z, \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \otimes \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu} \right) \right\rangle \\
\geq &-\varepsilon.
\end{aligned}$$

Consequently we have

$$\left\langle z, \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu}\right) \otimes \left(\sum_{\nu=1}^{n} \alpha_\nu \mathcal{X}_{E_\nu}\right)\right\rangle = \sum_{\nu=1}^{n} \sum_{\iota=1}^{n} \alpha_\nu \alpha_\iota \left\langle z, \mathcal{X}_{E_\nu} \otimes \mathcal{X}_{E\iota}\right\rangle \geq 0,$$

and hence $z \in \mathcal{COP}_{L^2(\mu)}$. □

By duality we have the following definition of the dual cone of $\mathcal{COP}_E$:

Definition 6.5. The *completely positive cone* $\mathcal{CP}_{E'}$ for E', where $E' = C(V)'$ or $E' = L^2(\mu)'$ is given as

$$\mathcal{CP}_{E'} := \left\{z \in \left(E \hat{\otimes}^s_{\varepsilon_s} E\right)' : \langle z, w\rangle \geq 0 \ \forall w \in \mathcal{COP}_E\right\}.$$

With this definition the completely positive cone regarding $C(V)$ and $L^p(\mu)$ with $1 \leq p < +\infty$ is given as follows. Since

$$\left(C(V) \hat{\otimes}^s_{\varepsilon_s} C(V)\right)' = \{\mu \in C(V \times V)' \colon \mu \text{ symmetric}\},$$

we get for the completely positive cone regarding $C(V)$ that

$$\mathcal{CP}_{C(V)'} = \left\{\mu \in C(V \times V)' \colon \mu \text{ symmetric and } \int_V K \, d\mu \geq 0 \ \forall K \in \mathcal{COP}_{C(V)}\right\}.$$

If we consider $E = L^p(\mu)$ with $1 \leq p < +\infty$ we have

$$\left(L^p(\mu) \hat{\otimes}^s_{\varepsilon_s} L^p(\mu)\right)' = L^q(\mu) \hat{\otimes}^s_{\pi_s} L^q(\mu)$$

for $\frac{1}{p} + \frac{1}{q} = 1$, cf. [19, Section 6.2]. With this property the completely positive cone regarding $L^p(\mu)$ is

$$\mathcal{CP}_{L^p(\mu)} = \left\{z \in L^q(\mu) \hat{\otimes}^s_{\pi_s} L^q(\mu) \colon \langle z, w\rangle \geq 0 \text{ for all } w \in \mathcal{COP}_{L^p(\mu)}\right\}.$$

With these characteristics we will have a closer look at the completely positive cone concerning $L^2(\mu)$ and its special properties in the next section.

6.2 The completely positive cone and its extreme rays

In this section we will have a closer look at the completely positive cone and especially at its extreme rays. Recall here some important aspects of functional analysis which can be found in the Appendix. They will be useful for some results we need to prove the representation of the completely positive cone by its extreme rays, which is the goal of the current section.

We define the following set

$$\mathcal{B} := \left\{ f \in L^2(\mu) \colon f \geq 0, \left\| f \right\|_{L^2(\mu)} \leq 1 \right\}.$$

By Proposition A.3 the space $L^2(\mu)$ is separable since by assumption (Introduction of Section 6.1) μ is σ-finite and $\mathcal{A}$ is countably generated. Thus $\mathcal{B}$ equipped with the $\sigma(L^2(\mu), L^2(\mu)') = \sigma(L^2(\mu), L^2(\mu))$-topology (= weak topology) is compact and metrizable, and hence by Proposition A.4 a separable topological space. Therefore due to duality it is also possible to consider the weak*- topology $\sigma(L^2(\mu)\hat{\otimes}_{\pi_s} L^2(\mu), L^2(\mu)\hat{\otimes}_{\varepsilon_s} L^2(\mu))$.

In contrast to the properties in the finite dimensional case in Proposition 2.3 we can prove the following theorem:

Theorem 6.6. *$\mathcal{CP}_{L^2(\mu)}$ is a closed convex pointed cone in $L^2(\mu)\hat{\otimes}_{\pi_s} L^2(\mu)$ and it has no interior points if* $\dim(L^2(\mu)) = +\infty$.

Proof. Since the completely positive cone $\mathcal{CP}_{L^2(\mu)}$ is the dual cone of the copositive cone $\mathcal{COP}_{L^2(\mu)}$, it is closed and convex. To prove that $\mathcal{CP}_{L^2(\mu)}$ is pointed we assume that $z \in \mathcal{CP}_{L^2(\mu)}$ and that also $-z \in \mathcal{CP}_{L^2(\mu)}$. Then $\langle w, z \rangle = 0$ for all $w \in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)$. Thus $\pi_s(z) = 0$ and hence $z = 0$. To prove that $\mathcal{CP}_{L^2(\mu)}$ has no interior point we show that the generalization of $\mathcal{S}^+_n$ in $L^2(\mu)$

$$\mathcal{S}^+_{L^2(\mu)} := \left\{ z \in L^2(\mu)\hat{\otimes}^s_{\pi_s} L^2(\mu) \colon \langle z, f \otimes f \rangle \geq 0 \;\forall f \in L^2(\mu) \right\}$$

has no interior point. Since $\mathcal{CP}_{L^2(\mu)} \subseteq \mathcal{S}^+_{L^2(\mu)}$ it follows directly that also $\mathcal{CP}_{L^2(\mu)}$ has no interior point. Therefore let $z \in \mathcal{S}^+_{L^2(\mu)}$ be an interior point. Then we have $z = \sum_{\nu=0}^{\infty} \lambda_\nu f_\nu \otimes f_\nu$, where $f_0, f_1, \ldots$ is an orthonormal system in $L^2(\mu)$, and $0 \leq \lambda_\nu \downarrow 0$ with $\sum_{\nu=0}^{\infty} \lambda_\nu < \infty$. Therefore there exists an $\varepsilon > 0$

with $z - \varepsilon g \otimes g \in \mathcal{S}^+_{L^2(\mu)}$ for all $g \in L^2(\mu)$ with $\left\|g\right\|_{L^2(\mu)} = 1$. Let $\lambda_{\nu_0} < \varepsilon$ and $g = f_{\nu_0}$, then

$$\sum_{\nu=0}^{\infty} \lambda_\nu f_\nu \otimes f_\nu - \varepsilon f_{\nu_0} \otimes f_{\nu_0} = (\lambda_{\nu_0} - \varepsilon)\, f_{\nu_0} \otimes f_{\nu_0} + \sum_{\substack{\nu=0\\ \nu\neq\nu_0}}^{\infty} \lambda_\nu f_\nu \otimes f_\nu .$$

With $\lambda_{\nu_0} - \varepsilon < 0$ it follows that $z - \varepsilon g \otimes g \notin \mathcal{S}^+_{L^2(\mu)}$ since it has a negative eigenvalue. Therefore $\mathcal{S}^+_{L^2(\mu)}$ has no interior points and hence neither has $\mathcal{CP}_{L^2(\mu)}$. □

Remark 6.7. Just like $\mathcal{CP}_{L^2(\mu)}$ the cone $\mathcal{CP}_{C(V)}$ has no interior points. Let $x_0, x_1 \in V$ with $x_1 \neq x_0$ and consider $\delta_{x_0} \otimes \delta_{x_0} \in \mathcal{CP}_{C(V)}$, where δ_x denotes the Dirac measure of x. Then $\delta_{x_0} \otimes \delta_{x_0} - \varepsilon\,(\delta_{x_1} \otimes \delta_{x_1}) \notin \mathcal{CP}_{C(V)}$ for all $\varepsilon > 0$.

Remark 6.8. A preferable result in this context would be a generalization of Lemma 2.4 in infinite dimension. Unfortunately this is not possible: Consider

$$\mathcal{S}^+_{(L^2(\mu),\sigma)} := \left\{ z \in L^2(\mu) \hat{\otimes}^s_\sigma L^2(\mu) \colon \ \langle z, f \otimes f \rangle \geq 0 \ \forall f \in L^2(\mu) \right\}.$$

This cone is pointed and has no interior point. Moreover $\mathcal{S}^+_{(L^2(\mu),\sigma)}$ is selfdual, thus its dual cone also has no interior points.

Remark 6.9. Another interesting question in this context is whether a $z \in \mathcal{CP}_{L^2(\mu)}$ can be approximated by a sequence $z_n \in \mathcal{CP}_{L^2(\mu)}$ with $z_n \in \operatorname{span}(\mathcal{X}_{E_\nu} \otimes \mathcal{X}_{E_\iota})$ such that $z - z_n \in \mathcal{CP}_{L^2(\mu)}$, i.e. $z_n = \sum_{\iota=1}^{\infty} h_\iota \otimes h_\iota$ with $h_\iota \geq 0$ and further $\sum_{\iota=1}^{\infty} \|h_\iota\|^2_{L^2(\mu)} < \pi_s(z) + \varepsilon$. However, we were not able to prove the existence of such a sequence so far.

Next we will repeat the definition of the trace for our purpose. It can be found in [19, Section 2.5]:

Definition 6.10. Let E be a vector space. For a linear map $T \colon E \to E$ with $T = \sum_{l=1}^{m} x_l \otimes x_l$ and representing matrix $(\langle e_l, Te_k \rangle)_{l_k}$, where $e_1, \ldots, e_n$ is a basis of E with coefficient functionals e_k^*, cf. [39, p. 340, Definition], the *trace* is

$$\begin{aligned} \operatorname{tr}(T) &= \operatorname{tr}\left(\sum_{l=1}^{m} x_l \otimes x_l \right) = \sum_{l=1}^{m} \langle x_l, x_l \rangle = \sum_{l=1}^{m} \left\langle \sum_{k=1}^{n} \langle x_l, e_k \rangle\, e_k^*, x_l \right\rangle \\ &= \sum_{k=1}^{n} \left\langle e_k^*, \sum_{l=1}^{m} \langle x_l, e_k \rangle\, x_l \right\rangle = \sum_{k=1}^{n} \left\langle e_k^*, Te_k \right\rangle . \end{aligned}$$

Hence the trace of T is the sum of the diagonal elements of the representing matrix of the mapping T. Because of the circumstance that we consider here the case $p = 2$, where the symmetric projective norm of a completely positive operator is constant for all decompositions by Theorem 4.2, we do not need the representing matrix to calculate the trace.

Analogously to Theorem 4.2 we prove the following theorem:

Theorem 6.11. *Let $z \in L^2(\mu)\hat{\otimes}^s_{\pi_s} L^2(\mu)$ with $z = \sum_{\nu=1}^N f_\nu \otimes f_\nu$ and $f_\nu \geq 0$, then*

$$\pi_s(z) = \operatorname{tr}(z) = \sum_{\nu=1}^{N} \left\|f_\nu\right\|^2_{L^2(\mu)}. \tag{6.1}$$

Proof. The idea of the proof of (6.1) is to show both inequalities and thus the equality holds. The trace is $\operatorname{tr}\colon E \otimes E \to \mathbb{R}$ with $\sum_\nu x_\nu \otimes x_\nu \mapsto \sum x'_\nu(x_\nu)$. Let $z = \sum_{\nu=1}^n f_\nu \otimes f_\nu$, then:

$$\begin{aligned}
\operatorname{tr}(z) &= \operatorname{tr}\left(\sum_{\nu=1}^{n} f_\nu \otimes f_\nu\right) \\
&= \sum_{\nu=1}^{n} \langle f_\nu, f_\nu \rangle \\
&= \sum_{\nu=1}^{n} \int_V f_\nu f_\nu \, d\mu \\
&= \sum_{\nu=1}^{n} \left\|f_\nu\right\|^2_{L^2(\mu)} \\
&\geq \pi_s(z).
\end{aligned}$$

Conversely let $z = \sum_{\nu=1}^N g_\nu \otimes h_\nu$, then:

$$\begin{aligned}
\operatorname{tr}(z) &= \operatorname{tr}\left(\sum_{\nu=1}^{N} g_\nu \otimes h_\nu\right) \\
&= \sum_{\nu=1}^{N} \langle g_\nu, h_\nu \rangle \\
&\leq \sum_{\nu=1}^{N} \left\|g_\nu\right\|_{L^2(\mu)} \|h_\nu\|_{L^2(\mu)}.
\end{aligned}$$

The last inequality holds by the Cauchy-Schwarz inequality. Using the infimum supplies the definition of the π-norm on the right-hand side of the inequality:

$$\operatorname{tr}(z) \leq \pi(z) \leq \pi_s(z)$$

The last inequality holds by [28, Section 2.3]. Combining both inequalities leads to (6.1):

$$\pi_s(z) = \operatorname{tr}(z) = \sum_{\nu=1}^{N} \left\| f_\nu \right\|_{L^2(\mu)}^2 .$$

□

Remark 6.12. Note that the restriction $f \geq 0$ is not necessary for the proof. So this holds for any positive semidefinite z, and not only for completely positive z.

Next for $z \in \mathcal{CP}_{L^2(\mu)}$ with $\pi_s(z) = 1$ we want to construct a probability space $(V, \mathscr{S}, P)$ such that $\omega \mapsto f_\omega \otimes f_\omega$ is measurable with respect to the Borel σ-algebra and restricted as well as $z = \int_V f_\omega \otimes f_\omega \, dP(\omega)$. Note here that the measurability poses no problem, since every continuous map $f \colon A \to B$ on a metric (or topological) space A in the metric (or topological) space B is Borel-measurable, cf. [27, III, Corollary 1.4].

But first we will prove the following lemma:

Lemma 6.13. *The convex set* $\operatorname{conv}\{f \otimes f \colon f \in L^2(\mu), f \geq 0\}$ *is* π_s*-dense in* $\mathcal{CP}_{L^2(\mu)}$.

Proof. We use: If $M \subseteq L^2(\mu)$ finite and $\varepsilon > 0$, then there exists a measurable decomposition $\mathcal{E} = \{E_1, \dots, E_n\}$ of V such that

$$\left\| f - \sum_{\nu=1}^{n} \frac{1}{\mu(E_\nu)} \int_{E_\nu} f \, d\mu \mathcal{X}_{E_\nu} \right\|_{L^2(\mu)} < \varepsilon \quad \text{for all } f \in M.$$

We set $S_{\mathcal{E}}\colon L^2(\mu) \to L^2(\mu)$ with $f \mapsto \sum_{\nu=1}^{n} \left(\frac{1}{\mu(E_\nu)} \int_{E_\nu} f\, d\mu\right) \mathcal{X}_{E_\nu}$. Then:

$$\begin{aligned}
\left\|S_{\mathcal{E}}(f)\right\|_{L^2(\mu)}^2 &= \int_V \left(\sum_{\nu=1}^{n} \left(\frac{1}{\mu(E_\nu)} \int_{E_\nu} f\, d\mu\right) \mathcal{X}_{E_\nu}\right)^2 d\mu \\
&= \sum_{\nu=1}^{n} \sum_{\iota=1}^{n} \frac{1}{\mu(E_\nu)} \frac{1}{\mu(E_\iota)} \left(\int_{E_\nu} f\, d\mu\right)\left(\int_{E_\iota} f\, d\mu\right) \mu(E_\nu \cap E_\iota) \\
&= \sum_{\nu=1}^{n} \frac{1}{\mu(E_\nu)^2} \left(\int_{E_\nu} f\, d\mu\right)^2 \mu(E_\nu) \\
&= \sum_{\nu=1}^{n} \frac{1}{\mu(E_\nu)} \left(\int_{E_\nu} f\, d\mu\right)^2 \\
&\le \sum_{\nu=1}^{n} \frac{1}{\mu(E_\nu)} \left(\int_{E_\nu} |f|\, d\mu\right)^2 \\
&\le \sum_{\nu=1}^{n} \frac{1}{\mu(E_\nu)} \sqrt{\mu(E_\nu)}^2 \left(\int_{E_\nu} f^2\, d\mu\right)^{\frac{1}{2}\cdot 2} \\
&= \sum_{\nu=1}^{n} \int_{E_\nu} f^2\, d\mu \\
&= \int_V f^2\, d\mu \\
&= \left\|f\right\|_{L^2(\mu)}^2 .
\end{aligned}$$

Furthermore we have $S_{\mathcal{E}}(f) \geq 0$ if $f \geq 0$ and $S_{\mathcal{E}}$ is self-adjoint. For a measurable decomposition $\mathcal{E}$ of V we set $R\colon L^2(\mu)\hat{\otimes}_{\pi_s}^{s} L^2(\mu) \to L^2(\mu)\hat{\otimes}_{\pi_s}^{s} L^2(\mu)$ continuous and linear with $\sum_{\nu=0}^{\infty} \lambda_\nu g_\nu \otimes g_\nu \mapsto \sum_{\nu=0}^{\infty} \lambda_\nu S_{\mathcal{E}}(g_\nu) \otimes S_{\mathcal{E}}(g_\nu)$, where $\lambda_\nu \downarrow 0$ and $g_0, g_1, \ldots$ an orthonormal system of $L^2(\mu)$. This map R fulfills $\pi_s(R(z)) \leq \pi_s(z)$ and further

$$\operatorname{im}(R) \subset \operatorname{span}\left\{\frac{1}{2}\left(\mathcal{X}_{E_\nu} \otimes \mathcal{X}_{E_\iota} + \mathcal{X}_{E_\iota} \otimes \mathcal{X}_{E_\nu}\right) : 1 \leq \nu, \iota \leq n\right\} \subset L^2(\mu) \otimes L^2(\mu).$$

Furthermore we have that $R(\mathcal{CP}_{L^2(\mu)}) \subset \mathcal{CP}_{L^2(\mu)}$. In the following we need to show for any $z \in \mathcal{CP}_{L^2(\mu)}$, i.e. $\langle z, w\rangle \geq 0$ for all $w \in \mathcal{COP}_{L^2(\mu)}$ that $R(z)$ is also in $\mathcal{CP}_{L^2(\mu)}$, i.e. $\langle R(z), w\rangle \geq 0$ for all $w \in \mathcal{COP}_{L^2(\mu)}$. For that purpose let

$z = \sum_{\nu=0}^{\infty} \lambda_\nu g_\nu \otimes g_\nu$ and $w = \sum_{\iota=0}^{\infty} \varrho_\iota h_\iota \otimes h_\iota$. Then we have:

$$\begin{aligned}
\langle R(z), w\rangle &= \sum_{\nu=0}^{\infty}\sum_{\iota=0}^{\infty} \lambda_\nu \varrho_\iota \left\langle S_{\mathcal{E}}(g_\nu) \otimes S_{\mathcal{E}}(g_\nu), h_\iota \otimes h_\iota \right\rangle \\
&= \sum_{\nu=0}^{\infty}\sum_{\iota=0}^{\infty} \lambda_\nu \varrho_\iota \left\langle S_{\mathcal{E}}(g_\nu), h_\iota \right\rangle^2 \\
&= \sum_{\nu=0}^{\infty}\sum_{\iota=0}^{\infty} \lambda_\nu \varrho_\iota \left\langle g_\nu, S_{\mathcal{E}}(h_\iota) \right\rangle^2 \\
&= \sum_{\nu=0}^{\infty}\sum_{\iota=0}^{\infty} \lambda_\nu \varrho_\iota \left\langle g_\nu \otimes g_\nu, S_{\mathcal{E}}(h_\iota) \otimes S_{\mathcal{E}}(h_\iota) \right\rangle \\
&= \sum_{\iota=0}^{\infty} \varrho_\iota \left\langle z, S_{\mathcal{E}}(h_\iota) \otimes S_{\mathcal{E}}(h_\iota) \right\rangle \\
&= \left\langle z, \sum_{\iota=0}^{\infty} \varrho_\iota S_{\mathcal{E}}(h_\iota) \otimes S_{\mathcal{E}}(h_\iota) \right\rangle.
\end{aligned}$$

It remains to show that $\sum_{\iota=0}^{\infty} \varrho_\iota S_{\mathcal{E}}(h_\iota) \otimes S_{\mathcal{E}}(h_\iota) \in \mathcal{COP}_{L^2(\mu)}$. Indeed since $f \geq 0$ we have:

$$\left\langle \sum_{\iota=0}^{\infty} \varrho_\iota S_{\mathcal{E}}(h_\iota) \otimes S_{\mathcal{E}}(h_\iota), f \otimes f \right\rangle = \left\langle \sum_{\iota=0}^{\infty} \varrho_\iota h_\iota \otimes h_\iota, S_{\mathcal{E}}(f) \otimes S_{\mathcal{E}}(f) \right\rangle \geq 0.$$

Therefore for $z \in \mathcal{CP}_{L^2(\mu)}$ we have shown that $R(z) \in \mathcal{CP}_{L^2(\mu)}$ and further that $R(z) = \sum_{k=1}^{K} \delta_k f_k \otimes f_k$ with $f_k \geq 0$. In the last step we need to show that for a $z \in \mathcal{CP}_{L^2(\mu)}$ with $z = \sum_{\nu=0}^{\infty} g_\nu \otimes g_\nu$ and for all $\varepsilon > 0$ there exists a decomposition $\mathcal{E}$ such that $\pi_s(z - R_{\mathcal{E}}(z)) < \varepsilon$. Furthermore we choose n_0 with $\sum_{\nu=n_0+1}^{\infty} |\lambda_\nu| < \varepsilon/4$ and select $\mathcal{E}$ such that

$$\pi_s \left(S_{\mathcal{E}}(g_\nu) \otimes S_{\mathcal{E}}(g_\nu) - g_\nu \otimes g_\nu \right) < \frac{\varepsilon}{2} \frac{1}{\sum_{\nu=0}^{n_0} |\lambda_\nu| + 1}.$$

Then:

$$\begin{aligned}
\pi_s \left(z - R_{\mathcal{E}}(z) \right) \leq& \sum_{\nu=0}^{n_0} |\lambda_\nu| \pi_s \left(S_{\mathcal{E}}(g_\nu) \otimes S_{\mathcal{E}}(g_\nu) - g_\nu \otimes g_\nu \right) + \sum_{\nu=n_0+1}^{\infty} |\lambda_\nu| \left\| S_{\mathcal{E}}(g_\nu) \right\|_{L^2(\mu)}^2 \\
&+ \sum_{\nu=n_0+1}^{\infty} |\lambda_\nu| \left\| g_\nu \right\|_{L^2(\mu)}^2 \\
&< \frac{\varepsilon}{2} + 2 \cdot \frac{\varepsilon}{4} = \varepsilon.
\end{aligned}$$

Therefore $\operatorname{conv}\{f \otimes f \colon f \in L^2(\mu), f \geq 0\}$ lies π_s-dense in $\mathcal{CP}_{L^2(\mu)}$ and the proof is complete. $\square$

We will show now:

Theorem 6.14. *Let $(X, \mathscr{A}, \mu)$ be a σ-finite measure space such that $\mathscr{A}$ is countably generated. Then for all $z \in \mathcal{CP}_{L^2(\mu)}$ with $\pi_s(z) = 1$ there exists a probability measure P on $\mathcal{B}$, equipped with the Borel σ-algebra in such a way that*

$$z = \int_{\mathcal{B}} f \otimes f \, dP(f).$$

Furthermore for every representation of that kind we have

$$1 = \pi_s(z) = \int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \, dP(f),$$

which is equivalent to

$$P\left(\left\{f \in \mathcal{B} \colon \|f\|_{L^2(\mu)} < 1\right\}\right) = 0.$$

Proof. Let $f_\nu^{(n)} \in L^2(\mu)$ with $f_\nu^{(n)} \geq 0$ for $n \in \mathbb{N}$ and $\nu = 1, \dots, n$. Furthermore let

$$\pi_s\left(z - \sum_{\nu=1}^{n} f_\nu^{(n)} \otimes f_\nu^{(n)}\right) \to 0 \text{ for } n \to \infty. \tag{6.2}$$

Moreover without loss of generality we have

$$\sum_{\nu=1}^{n} \left\|f_\nu^{(n)}\right\|^2_{L^2(\mu)} = \pi_s\left(\sum_{\nu=1}^{n} f_\nu^{(n)} \otimes f_\nu^{(n)}\right) = 1 \tag{6.3}$$

simply by scaling. Write

$$\sum_{\nu=1}^{n} f_\nu^{(n)} \otimes f_\nu^{(n)} = \int_{\mathcal{B}} f \otimes f \, dP_n(f), \tag{6.4}$$

where $P_n = \sum_{\nu=1}^{n} \left\|f_\nu^{(n)}\right\|^2_{L^2(\mu)} \delta_{\frac{f_\nu^{(n)}}{\left\|f_\nu^{(n)}\right\|_{L^2(\mu)}}}$. Thus every P_n is a probability measure on $\mathcal{B}$.

Since $C(\mathcal{B})$ is separable (cf. [49, p. 33]), there exists a convergent subsequence, which we denote by $(P_n)_{n\in\mathbb{N}}$. Then $(P_n)_{n\in\mathbb{N}}$ converges to a probability measure P on $\mathcal{B}$ with respect to the weak*-topology $\sigma(M_e(\mathcal{B}), C(\mathcal{B}))$ by the compactness of probability measures.

For this we have to show that $\int_{\mathcal{B}} f \otimes f \, dP(f)$ exists. Thus we have to prove that the map $f \mapsto f \otimes f$ is continuous. Let $w \in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)$, then we have to show that for $n \to \infty$ holds:

$$\langle f_n \otimes f_n, w\rangle \to \langle f_0 \otimes f_0, w\rangle .$$

For this we choose a $z \in L^2(\mu)\otimes L^2(\mu)$ such that $\varepsilon_s(w-z) < \varepsilon$ and we consider

$$\langle f_n \otimes f_n - f_0 \otimes f_0, w\rangle = \langle f_n \otimes f_n - f_0 \otimes f_0, w - z\rangle + \langle f_n \otimes f_n - f_0 \otimes f_0, z\rangle .$$

Thus we have $|\langle f_n \otimes f_n - f_0 \otimes f_0, w - z\rangle| \le 2\varepsilon$. Let $z = \sum_{\nu=0}^{\infty} h_\nu \otimes h_\nu$. We know that for all $h \in L^2(\mu)$ it is valid that $\langle f_n, h\rangle \to \langle f_0, h\rangle$. With $\langle f_n, h\rangle^2 \to \langle f_0, h\rangle^2$ it follows that

$$\langle f_n \otimes f_n, h \otimes h\rangle \to \langle f_0 \otimes f_0, h \otimes h\rangle .$$

Hence

$$\langle f_n \otimes f_n, z\rangle \to \langle f_0 \otimes f_0, z\rangle$$

and therefore

$$\langle f_n \otimes f_n - f_0 \otimes f_0, z\rangle \to 0 \ \text{ as } n \to \infty.$$

Thus we have for n large enough

$$|\langle f_n \otimes f_n - f_0 \otimes f_0, w - z\rangle| \le 2\varepsilon.$$

This proves that $f \mapsto f \otimes f$ is continuous and that the integral $\int_{\mathcal{B}} f \otimes f \, dP(f)$ exists.

Then for all $h \in L^2(\mu)$ it follows:

$$\begin{aligned}
\langle z, h \otimes h\rangle &= \lim_{n\to\infty} \sum_{\nu=1}^{n} \langle f_\nu^{(n)}, h\rangle^2 \\
&= \lim_{n\to\infty} \int_{\mathcal{B}} \langle f, h\rangle^2 \, dP_n(f) \\
&= \int_{\mathcal{B}} \langle f, h\rangle^2 \, dP(f) \\
&= \langle \int_{\mathcal{B}} f \otimes f \, dP(f), h \otimes h\rangle.
\end{aligned}$$

With this we get $z = \int_{\mathcal{B}} f \otimes f \, dP(f)$. The first equality uses property (6.2) and the second equality uses (6.4), i.e. that P_n converges on continuous functions. The third equality holds by the weak* convergence of P_n to P and the last equality is valid by the Theorem of Hille [23, II Theorem 6].

Last we show the second property of the theorem:

$$\begin{aligned}
\int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \, dP(f) &= \lim_{n\to\infty} \int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \, dP_n(f) \\
&= \lim_{n\to\infty} \int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \sum_{\nu=1}^{n} \|f_\nu\|^2_{L^2(\mu)} \, d\delta_{f_\nu^{(n)}}(f) \\
&= \lim_{n\to\infty} \int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \, \pi_s\left(\sum_{\nu=1}^{n} f_\nu^{(n)} \otimes f_\nu^{(n)}\right) d\delta_{f_\nu^{(n)}}(f) \\
&= \lim_{n\to\infty} \int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \, d\delta_{f_\nu^{(n)}}(f) \\
&= \lim_{n\to\infty} \sum_{\nu=1}^{n} \left\|f_\nu^{(n)}\right\|^2_{L^2(\mu)} \\
&= \lim_{n\to\infty} \pi_s\left(\sum_{\nu=1}^{n} f_\nu^{(n)} \otimes f_\nu^{(n)}\right) \\
&= \pi_s(z) \\
&= 1.
\end{aligned}$$

The first equality holds by the weak convergence of P_n to P and the second by the definition of P_n. The third, the fourth and the sixth equality are fulfilled by (6.3). The seventh is satisfied by (6.2) and the last by assumption. The property $P\left(\left\{f \in \mathcal{B} \colon \|f\|_{L^2(\mu)} < 1\right\}\right) = 0$ follows directly with the equation $\int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \, dP(f) = 1$. □

In the next theorem we will prove a representation of the extreme rays of the completely positive cone.

Theorem 6.15. *The set* $\operatorname{Ext}\left(\mathcal{CP}_{L^2(\mu)}\right)$ *of extreme rays of the completely positive cone* $\mathcal{CP}_{L^2(\mu)}$ *is given by*

$$\operatorname{Ext}\left(\mathcal{CP}_{L^2(\mu)}\right) = \left\{[0,\infty)\; f \otimes f \colon f \in L^2(\mu), f \geq 0\right\}.$$

Proof. Let $A = [0,\infty)\, z_0$ be an extreme ray of $\mathcal{CP}_{L^2(\mu)}$ with $z_0 \neq 0$ and without loss of generality $\pi_s(z_0) = 1$. Then by Theorem 6.14 we have

$$z_0 = \int_{\mathcal{B}} f \otimes f \, dP(f)$$

for some probability measure P. For every measurable subset $S \subset \mathcal{B}$ we have

$$z_0 = \int_S f \otimes f \, dP(f) + \int_{\mathcal{B}\setminus S} f \otimes f \, dP(f)$$

and since z_0 is an extreme direction there exists some $\lambda_S \geq 0$ such that

$$\lambda_S z_0 = \int_S f \otimes f \, dP(f) \tag{6.5}$$

for all measurable $S \subset \mathcal{B}$. Then $f \mapsto f \otimes f$ is P-almost surely constant. The argument for this is the following: By equation (6.5) we know that $\int_S f \otimes f \, dP(f) \in \operatorname{span}(z_0)$. We consider $\{z_0\}^\circ = \left\{w \in L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu) \colon w(z_0) = 0\right\}$, where $\{z_0\}^\circ$ denotes the annihilator of $\{z_0\}$ in $L^2(\mu)\hat{\otimes}^s_{\pi_s} L^2(\mu)$, cf. [39, p. 48/49 Definition and Remark]. Then let $A \subset \{z_0\}^\circ$ be countable and dense (this is possible since $L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)$ is a separable Banach space). Then for all $w \in A$ it is valid that

$$\int_S w(f \otimes f) \, dP(f) = 0 \quad \text{for all measurable } S \subset \mathcal{B}.$$

Then

$$f \mapsto w(f \otimes f) \equiv 0 \;\; P\text{–a.e.}$$

Since A is countable there exists a measurable set $N \subseteq \mathcal{B}$ with $P(N) = 0$ such that $w(f \otimes f) = 0$ for all $w \in A$ and for all $f \in \mathcal{B} \setminus N$. Hence for all $f \in \mathcal{B} \setminus N$ it holds $f \otimes f \in A^\circ = \operatorname{span}(z_0)$. Therefore we have $z_0 = f_0 \otimes f_0$ with $f_0 \in L^2(\mu)$ and $f_0 \geq 0$.

Conversely let $f_0 \in L^2(\mu)$ with $f_0 \geq 0$ and $\|f_0\|_{L^2(\mu)} = 1$. Furthermore let $f_0 \otimes f_0 = z_1 + z_2$ with $z_1, z_2 \in \mathcal{CP}_{L^2(\mu)}$. Then for $j = 1, 2$ we have:

$$\frac{1}{\pi_s(z_j)} z_j = \int_{\mathcal{B}} f \otimes f \, dP_j(f), \text{ with } P_j\left(\left\{f \colon \|f\|_{L^2(\mu)} < 1\right\}\right) = 0. \tag{6.6}$$

And therefore:

$$f_0 \otimes f_0 = \int_{\mathcal{B}} f \otimes f \, d(\pi_s(z_1)P_1 + \pi_s(z_2)P_2)(f). \tag{6.7}$$

Set $P(f) := [\pi_s(z_1)P_1 + \pi_s(z_2)P_2](f)$. Consequently we have:

$$\begin{aligned} 1 &= (\pi_s(f_0 \otimes f_0))^2 \\ &= \int_{\mathcal{B}} \|f\|^2_{L^2(\mu)} \, dP(f) \\ &= \int_{\mathcal{B}} 1 \, d(\pi_s(z_1)P_1 + \pi_s(z_2)P_2)(f) \\ &= \pi_s(z_1) + \pi_s(z_2). \end{aligned}$$

Thus P is a probability measure. Furthermore it follows:

$$\begin{aligned} 1 &= \|f_0\|^2_{L^2(\mu)} \\ &= \langle f_0 \otimes f_0, f_0 \otimes f_0 \rangle \\ &= \int_{\mathcal{B}} \langle f, f_0 \rangle^2 \, dP(f). \end{aligned}$$

The first equation holds by assumption. The second is true because of $\|f_0\|^2_{L^2(\mu)} = \langle f_0, f_0 \rangle^2 = \langle f_0 \otimes f_0, f_0 \otimes f_0 \rangle$ and the third follows with equality (6.7):

$$\begin{aligned} \langle f_0 \otimes f_0, f_0 \otimes f_0 \rangle &= \left\langle \int_{\mathcal{B}} f \otimes f \, dP(f), f_0 \otimes f_0 \right\rangle \\ &= \int_{\mathcal{B}} \langle f \otimes f, f_0 \otimes f_0 \rangle \, dP(f) \\ &= \int_{\mathcal{B}} \langle f, f_0 \rangle^2 \, dP(f). \end{aligned}$$

Consequently we get $\langle f, f_0 \rangle^2 = 1$ and $\|f_0\|_{L^2(\mu)} = \|f\|_{L^2(\mu)} = 1$ for P-almost all f. Thus $f = \pm f_0$ is valid for P-almost all f and with $f \in \mathcal{B}$ it follows that $f \geq 0$. This is why we have $f = f_0$ P-almost everywhere as well as P_1- and P_2-almost everywhere too. With $f = f_0$ P-almost everywhere, $f_0 \otimes f_0 = z_1 + z_2$ and (6.6) we have:

$$z_1 = \pi_s(z_1) \cdot z_0 \quad \text{and} \quad z_2 = \pi_s(z_2) \cdot z_0.$$

Therefore z_1 and z_2 lie in the same extreme ray and thus $f_0 \otimes f_0$ is an extreme ray of $\mathcal{CP}_{L^2(\mu)}$. □

Remark 6.16. With the result from Theorem 6.15 we get a new representation of the completely positive cone $\mathcal{CP}_{L^2(\mu)}$ by its extreme rays:

$$\mathcal{CP}_{L^2(\mu)} = \operatorname{cl}\operatorname{conv}\operatorname{cone}\left\{f \otimes f \colon f \in L^2(\mu), f \geq 0, \left\|f\right\|_{L^2(\mu)} \leq 1\right\},$$

where the closure is taken with respect to the weak*- topology $\sigma(L^2(\mu)\hat{\otimes}_{\pi_s} L^2(\mu), L^2(\mu)\hat{\otimes}_{\varepsilon_s} L^2(\mu))$.

Remark 6.17. The proof of Theorem 6.15 also gives a characterization of the extreme rays of the positive semidefinite cone $\mathcal{S}^+_{L^2(\mu)}$:

$$\operatorname{Ext}\left(\mathcal{S}^+_{L^2(\mu)}\right) = \left\{[0, \infty)\, f \otimes f \colon f \in L^2(\mu)\right\}.$$

6.3 Duality and application?

With these theoretical basics the next step would be to formulate the primal and dual optimization problems over the copositive as well as the completely positive cone similarly to Section 5.2. Unfortunately a direct reformulation is not possible. The main difficulty with an exact reformulation is that the trace is not defined in $L^2(\mu)\hat{\otimes}^s_{\varepsilon_s} L^2(\mu)$.
An interesting application of these optimization problems would be the kissing number problem, i.e. the maximum stable set problem over a circle. But also for this problem a direct copositive formulation is not possible since we need measures that describe the feasible set correctly. In a direct way this is very difficult to be realized in $L^2(\mu)$, perhaps another way via approximations would work.

Although we considered in this chapter another generalization concerning $L^2(\mu)$ instead of the set of continuous functions the results have a similar structure like the outcome of Chapter 5. Of course both chapters differ in essential characteristics, in particular the existence of usual tools, e.g. the trace, was difficult to realize. But nevertheless we obtained a generalization of the principle of copositivity as well as completely positivity that fits suitably into the previous results of copositivity in finite and even more in infinite dimension.

6.4 Current literature

The topic of copositive optimization in infinite dimension is a topical subject in cone optimization hence not solely we have been doing research in this topic. Therefore we will set [18] in contrast to our work.

The paper [18] by DeCorte, de Oliveira Filho and Vallentin, which was published during our work, also treats a generalization of copositive and completely positive optimization. In this paper the copositive and the completely positive cones are also discussed concerning the space $L^2(\mu)$. Some of their definitions and results may look similar to our work although the research was done independently. But in contrast to our theory, the authors considered the tensor product concerning the Hilbert-Schmidt norm. In our notation this is the space $L^2(\mu)\hat{\otimes}^s_{\sigma_s} L^2(\mu) = L^2(\mu \otimes \mu)$, which is the simplest way to generalize the theory. But in this space there is no trace available. We discussed this topic in a more general way since we can define the theory for all kernels, not only for continuous ones as it is the case in [18]. Furthermore our generalization provides a trace that we can define uniquely and that gives us the possibility to work via properties of the diagonal or instruments like the polarization formula. In particular the trace is an essential component in the formulation of the optimization problems. Our theory in this chapter corresponds to the choice of a maximal sized copositive and a minimal sized completely positive cone, which we control via the tensor product concerning the projective or the injective norm, respectively. In the paper of DeCorte et al. both cones were defined regarding the tensor product equipped with the Hilbert-Schmidt norm, which can be seen compared to our definitions as an in-between concerning their size.

Chapter 7

Conclusions and outlook

7.1 Summary

The goal of this thesis was to generalize the topic of copositivity and completely positivity suitably to an infinite dimensional space. In order to do so we introduced the topic of copositivity in finite dimension in Chapter 2. After a short motivation and preliminaries we presented some basic properties, the definition of the copositive and the completely positive cone as well as their special properties, i.e. both are pointed closed convex cones with nonempty interior. Furthermore we gave a representation of the respective interiors and considered their extreme rays. In particular the related optimization problems and duality statements were formulated. As an application we considered binary quadratic problems and problems from combinatorial optimization which can be formulated as copositive programs. The advantage of the reformulation is that the complexity of the new problem lies entirely in the cone constraint. Last we cite some approximation hierarchies of the copositive cone.
In order to generalize this theory in infinite dimension we introduced the background from functional analysis in Chapter 3 especially for Banach spaces. In this chapter we presented the tensor product of two vector spaces and three different tensor norms (projective, injective, Hilbert-Schmidt norm), which can be defined over normed and especially Hilbert spaces. For these norms we formulated special properties and explained the duality between the projective and the injective norm. As a special topic we presented the algebraic theory of the symmetric tensor product, which forms the basis for the generalization of copositivity and completely positivity later. Also here we gave three different symmetric tensor norms: the

symmetric projective, the symmetric injective and the symmetric Hilbert-Schmidt norm. A crucial point is that these norms are in general not simply the restrictions of the general norms to the symmetric spaces. They are calculated only for the symmetric decompositions. Also for these norms we presented some special properties and discussed the duality between the symmetric projective and the symmetric injective norm. We illustrated this chapter with some appropriate examples.
In Chapter 4 we discussed the symmetric projective norm in more detail. We were able to prove that in two spacial cases ($p = 1$ and $p = 2$) the symmetric projective norm can be determined by considering only the nonnegative decompositions and in particular that one nonnegative decomposition is enough for its calculation. Especially for the case $p \in (1, 2)$ we compute many different examples. For a special decomposition of a positive semidefinite matrix we tried to verify our conjecture that also in this case the symmetric projective norm can be determined via the nonnegative decompositions. These examples strengthen our conjecture but unfortunately we were not able to prove this equality.
The symmetric projective norm leads not only to nice numerical results but also gave us hints for our goal to generalize the topic of copositivity and completely positivity suitably to infinite dimension. It is intuitive that if the norm is quite easy to calculate, then we would also get a duality theory in which calculations are possible.
The first case $p = 1$ leads to the setting of copositivity in Hilbert-Schmidt kernels, which was first introduced in [24]. Analogous to the finite case we gave in Chapter 5 the definitions of the cone of copositive kernels and the cone of completely positive measures. Moreover we proved special properties, in particular the representation of the completely positive cone by its extreme rays. It is pleasant to see that the representations and properties look similar to the finite case. Nonetheless the proofs here are more complicated. As an application we considered some problems from combinatorial optimization. The kissing number problem, which is a special formulation of the stable set problem, was of particular interest. Regarding this application we cited approximation hierarchies as a generalization of the finite ones.
In Chapter 6 we discussed the respective generalizations of the copositive and the completely positive cone for the second special case of the projective norm, i.e. $p = 2$. The advantage here in contrast to Chapter 5 is that this case leads to the selfdual Hilbert space $L^2(\mu)$, i.e. both cones, the copositive and the completely positive, are subsets of the same space. Analogous to the previous case, we defined the copositive and the completely positive cone in $L^2(\mu)$. As a pleasant result we proved that both cones are closed

convex and pointed. In contrast to the finite case they have no interior points. Another useful result was the characterization of the set of extreme rays of the completely positive cone. The proof to this representation was quite extensive and required some auxiliary statements which we proved too. In this context we mentioned the difficulties relating to reformulate the optimization problems. Last we referenced to the very topical literature.

7.2 Further research

As presented in this thesis copositive optimization in infinite dimension is a new topic in optimization. Therefore we had to discuss the basic properties and generalizations of statements from finite dimension first. But now there is a more specific framework to refine this topic. In particular for Hilbert spaces the theory can be further developed.
There are many possibilities to expand this theory. Starting with finding more representations, characteristics and properties of the copositive and the completely positive cone, respectively. Like in Sections 2.4 and 5.4 an approximation of $\mathcal{COP}_{L^2(\mu)}$ would be a desirable extension of this topic. A next step could be the formulation of the primal and the dual optimization problems such that the basic properties in functional analysis hold true. This theory around the primal-dual pair would offer many possibilities for further research too. For example finding constraint qualifications or necessary optimality conditions would be essential points in optimization. A fundamental result in this context would be the proof of strong duality between the primal and the dual program.
Furthermore a transfer of the known solution techniques from finite dimension like the simplex partition via inner and outer approximations or the interior point method would be a new field in this context. Moreover an application of this new theory to the topic of quadratic programs, especially standard quadratic programs, or to optimization problems with binary constraints would also be worth considering. Apart from the kissing number problem there are many more problems from combinatorial optimization that can be formulated as copositive program. It would be interesting to transfer those into our context.

Appendix

Here we will state some definitions and properties concerning separability, which are relevant in Chapters 5 and 6. The use of these aspects from functional analysis is explained in these chapters too.

Definition A.1. [15, p. 102] Let $\mathscr{A}$ be a σ-algebra on the set X. Then $\mathscr{A}$ is *countably generated* if there is a countable subfamiliy $\mathscr{C}$ of $\mathscr{A}$ such that $\mathscr{A} = \sigma(\mathscr{C})$.

Definition A.2. [15, p. 86] A metric space is called *separable* if it has a countable dense subset.

Proposition A.3. *[15, Proposition 3.4.5] Let $(X, \mathscr{A}, \mu)$ be a measure space and $1 \leq p < +\infty$. If μ is σ-finite and $\mathscr{A}$ is countably generated, then $L^p(X, \mathscr{A}, \mu)$ is separable.*

Proposition A.4. *[35, §4, Section 5 (2)] Let (X, d) be a compact metrizable space. Then (X, d) is separable.*

Definition .1. [47, 3.16 Definition] A complex function f on a locally compact Hausdorff space X is said to *vanish at infinity* if for every $\varepsilon > 0$ there exists a compact set $S \subset X$ such that $|f(x)| < \varepsilon$ for all x not in S. We denote the class of all continuous functions on X that vanish at infinity by $C_0(X)$.

Theorem A.5. *[47, 6.19 Theorem] If X is a locally compact Hausdorff space, then every bounded linear functional Φ on $C_0(X)$ is represented by a unique regular complex Borel measure μ, in the sense that*

$$\Phi f = \int_X f \, d\mu$$

for every $f \in C_0(X)$. Moreover, the norm of Φ is the total variation norm of μ:

$$\|\Phi\| = \|\mu\|_{\mathrm{tv}}(X).$$

Bibliography

[1] K.M. Anstreicher, S. Burer, P.J.C. Dickinson, *An algorithm for computing the cp-factorization of a completely positive matrix*, Working Paper.

[2] C. Bachoc, F. Vallentin, *New upper bounds for kissing number from semidefinite programming*, Journal of the American Mathematical Society 21-3, 909–924, 2008.

[3] A. Barvinok, *A course in convexity*, Graduate Studies in Mathematics 54, Oxford University Press, 2002.

[4] Y.M. Berezansky, Z.G. Sheftel, G.F. Us, *Functional Analysis Vol. I*, Operator Theory: Advances and Applications Vol. 85, Birkhäuser, 1996.

[5] A. Berman, N. Shaked-Monderer, *Completely positive matrices*, World Scientific Publishing Co., 2003.

[6] S. Bochner, *Hilbert distances and positive definite functions*, Annals of Mathematics Vol. 42, No. 3, 647-656, 1941.

[7] I.M. Bomze, *Copositive optimization - recent developments and applications*, European Journal of Operational Research 216(3), 509-520, 2012.

[8] I.M. Bomze, M. Budinich, P.M. Pardalos, M. Pelillo, *The Maximum Clique Problem*, Handbook of Combinatorial Optimization, Kluwer Academic Publishers, 1-74, 1999.

[9] I.M. Bomze, M. Dür, E. de Klerk, C. Roos, A.J. Quist, T. Terlaky, *On copositive programming and standard quadratic optimization problems*, Journal of Global Optimization 18(4), 301-320, 2000.

[10] S. Bundfuss, *Copositive matrices, copositive programming and applications*, Ph.D. Thesis, Technical University Darmstadt, 2009.

[11] S. Bundfuss, M. Dür, *An adaptive linear approximation algorithm for copositive programs*, SIAM Journal of Optimization 20(1), 30-53, 2009.

[12] S. Bundfuss, M. Dür, *Algorithmic copositivity detection by simplicial partitions*, Linear Algebra and its Applications 428(7), 1511-1523, 2008.

[13] S. Burer, *Copositive Programming*, In: M.F. Anjos, J.B. Lasserre (Eds.), Handbook of Semidefinite, Conic and Polynomial Optimization, International Series in Operations Research + Management Science 166, Springer Science+Business Media, 2012.

[14] S. Burer, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Mathematical Programming 120 Ser. A, 479-495, 2009.

[15] D.L. Cohn, *Measure Theory*, Birkhäuser Basel, 2nd Edition, 2013.

[16] E. de Klerk, D.V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM Journal on Optimization 12(4), 875-892, 2002.

[17] D. de Laat, F. Vallentin, *A semidefinite programming hierarchy for packing problems in discrete geometry*, Mathematical Programming 151(2), 529-553, 2015.

[18] E. DeCorte, F.M. de Oliveira Filho, F. Vallentin, *Complete positivity and distance avoiding sets*, arXiv:1804.09099v2, 2018.

[19] A. Defant, K. Floret, *Tensor Norms and Operator Ideals*, North-Holland and Mathematics Studies, Volume 176, 1993.

[20] P.J.C. Dickinson, *The copositive cone, the completely positive cone and their generalisations*, Ph.D. Thesis, University of Groningen, 2013.

[21] P.J.C. Dickinson, *A new Certificate for Copositivity*, Preprint, `http://www.optimization-online.org/DB_HTML/2017/07/6146.html`, 2018.

[22] P.J.C. Dickinson, L. Gijben *On the Computational Complexity of Membership Problems for the Completely Positive Cone and its Dual*, Computational Optimization and Applications 57(2), 403–415, 2014.

[23] J. Diestel, J.J. Uhl, Jr., *Vector Measures*, Mathematical Surveys and Monographs, Volume 15, American Mathematical Society, 1977.

[24] C. Dobre, M. Dür, L. Frerick, F. Vallentin, *A copositive formulation for the stability number of infinite graphs*, Mathematical Programming 160 Ser. A, 65-83, 2016.

[25] M. Dür, *Copositive Programming - a Survey*, In: M. Diehl, F. Glineur, E. Jarlebring, W. Michiels (Eds.), Recent Advances in Optimization and its Applications in Engineering, Springer Berlin Heidelberg, 3-20, 2010.

[26] M. Dür, P. Groetzner, *A factorization method for completely positive matrices*, Preprint, `http://www.optimization-online.org/DB_HTML/2018/03/6511.html`, 2018.

[27] J. Elstrodt *Maß- und Integrationstheorie*, Springer, 6. Edition, 2010.

[28] K. Floret, *Natural norms on symmetric tensor products of normed spaces*, Note di Mathematica 17, 153–188, 1997.

[29] S.A. Gaal, *Linear Analysis and Representation Theory*, Springer Science+Business Media, 1973.

[30] W. Hackbusch, *Tensor spaces and numerical tensor calculus*, Springer, 2012.

[31] M. Hall, Jr., M. Newman, *Copositive and completely positive quadratic forms*, Mathematical Proceedings of the Cambridge Philosophical Society 59, 329-339, 1963.

[32] C. Helmberg, *Semidefinite Programming for Combinatorial Optimization*, Habilitation, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 2000.

[33] A.J. Hoffman, F. Pereira, *On copositive matrices with* $-1, 0, 1$ *entries*, Journal of Combinatorial Theory 14(3), 302-309, 1973.

[34] R.V. Kadison, J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras, Bd. 1: Elementary Theory*, Academic Press, 1983.

[35] G. Köthe, *Topological Vector Spaces I*, Springer, 1983.

[36] O. Kuryatnikova, J.C. Vera Lizcano, *Approximating the cone of copositive kernels to estimate the stability number of infinite graphs*, Electronic Notes in Discrete Mathematics 62, 303-308, 2017.

[37] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal of Optimization 11(3), 769-817, 2001.

[38] J.B. Lasserre, *An explicit equivalent positive semidefinite program for nonlinear 0-1 programs*, SIAM Journal of Optimization 12(3), 756-769, 2002.

[39] R. Meise, D. Vogt, *Introduction to Functional Analysis*, Oxford Graduate Texts in Mathematics 2, 1997.

[40] T.S. Motzkin, E.G. Straus, *Maxima for graphs and a new proof of a theorem of Turan*, Carnadian Journal of Mathematics 17, 533-540, 1965.

[41] K.G. Murty, S.N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Mathematical Programming 39, 117-129, 1987.

[42] P. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, Ph.D. Thesis, California Institute of Technology, Pasadena, 2000.

[43] P. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Mathematical Programming Ser. B 96, 293-320, 2003.

[44] J. Peña, J.C. Vera, L.F. Zuluaga, *Computing the stability number of a graph via linear and semidefinite programming*, SIAM Journal of Optimization 18(1), 87-105, 2007.

[45] R.R. Phelps, *Lectures on Choquet's Theorem*, Springer, 2001.

[46] W. Rudin, *Functional Analysis*, McGraw-Hill Book Company, 1973.

[47] W. Rudin, *Real and Complex Analysis*, McGraw-Hill Book Company, 3. Edition, 1987.

[48] R.A. Ryan, *Introduction to Tensor Products of Banach Spaces*, Springer, 2002.

[49] D. Werner, *Funktionalanalysis*, Springer, 5. Edition, 2005.