



Delft University of Technology

Document Version

Final published version

Citation (APA)

Bekker, A. J. F. (2026). *Optimization hierarchies for extremal geometry through complete positivity*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:cc5f3ff8-1adb-4742-89a3-2cd2f526b5b8>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

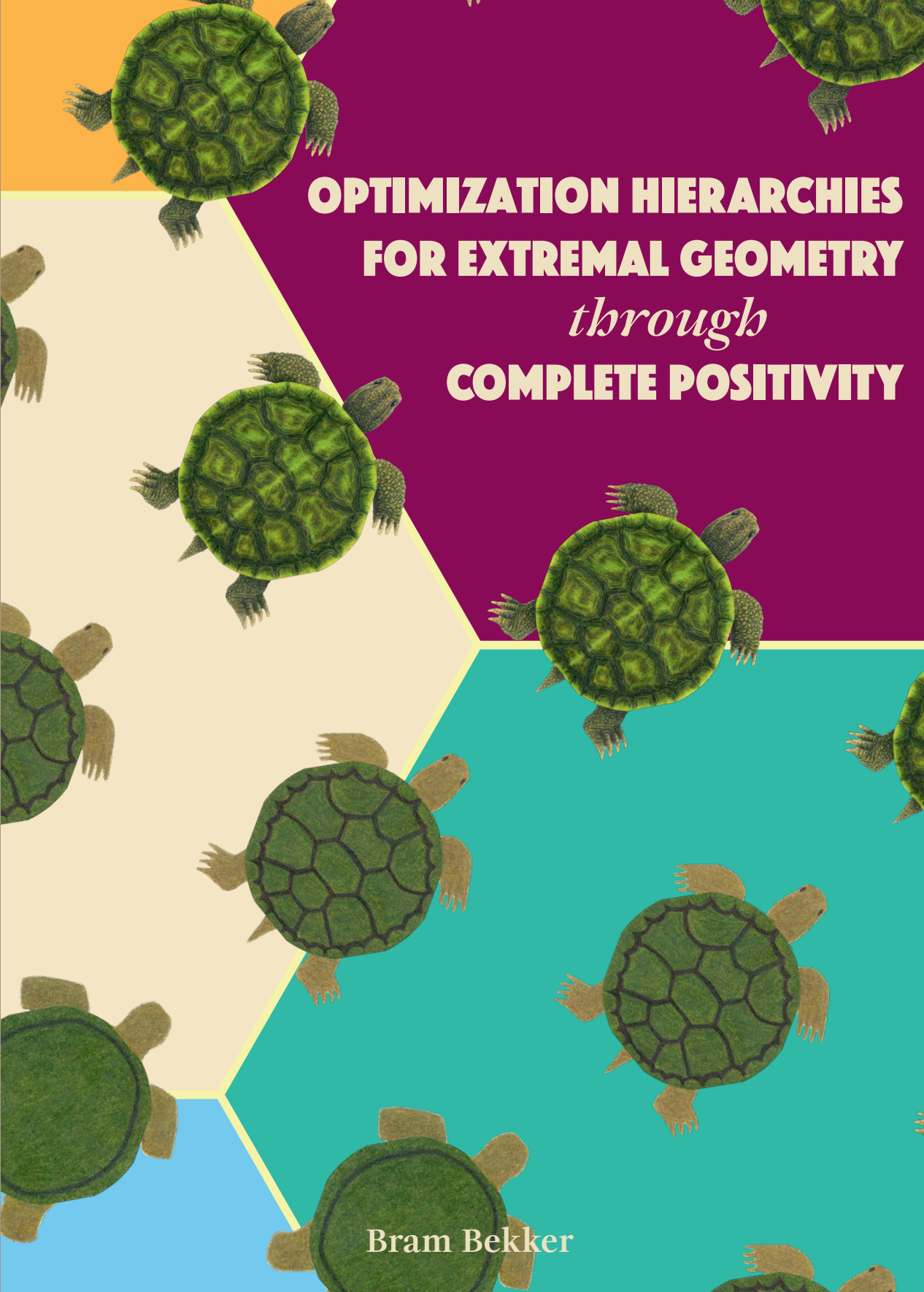
Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

The background of the cover is a repeating pattern of green turtles with brown limbs, swimming in a clockwise direction. The turtles are positioned on a grid of large hexagons. The hexagons are colored in a sequence: orange, purple, yellow, teal, and light blue, creating a vibrant, geometric backdrop.

OPTIMIZATION HIERARCHIES FOR EXTREMAL GEOMETRY *through* **COMPLETE POSITIVITY**

Bram Bekker

Optimization hierarchies for extremal geometry through complete positivity

Optimization hierarchies for extremal geometry through complete positivity

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof. dr. ir. H. Bijl;
Chair of the Board for Doctorates
to be defended publicly on
Thursday, 22 October 2026 at 15:00

by

Abraham Johannes Franciscus BEKKER

This dissertation has been approved by the promotors and the external advisor.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Dr. F.M. de Oliveira Filho	Delft University of Technology, promotor
Prof. dr. D.C. Gijswijt	Delft University of Technology, promotor
Dr. P. Moustrou	University of Toulouse – Jean Jaurès, France, external advisor

Independent members:

Prof. dr. ir. M.C. Veraar	Delft University of Technology
Prof. dr. F. Vallentin	University of Cologne, Germany
Prof. dr. E. de Klerk	Tilburg University
Prof. dr. A. Wiegele	University of Klagenfurt, Austria
Prof. dr. J.M.A.M. van Neerven	Delft University of Technology, <i>reserve member</i>

The research in this thesis is funded by the grant OCENW.KLEIN.024 of the Dutch Research Council (NWO).



Copyright © 2026 by A.J.F. Bekker

Artwork by Roos Ramakers

Printed by Ridderprint

ISBN 978-94-6563-021-2

An electronic version of this dissertation is available at
<http://repository.tudelft.nl/>.

Contents

Summary	ix
Samenvatting	xi
Chapter 1. Introduction	1
1. Some notation	4
2. The independence number of a finite graph	5
3. The copositive cone in finite dimensions	9
Part 1. The completely positive cone	13
Chapter 2. The completely positive cone of a measure space	15
1. Notation and preliminaries	16
2. Projective approximation and martingales	18
3. A Pólya-type approximation on finite measure spaces	23
4. An outer approximation on σ -finite measure spaces	27
Chapter 3. Complete positivity and symmetry	31
1. Harmonic analysis	31
2. Averaging	34
3. A group-invariant outer approximation	38
Part 2. Finite measure spaces	45
Chapter 4. Completely positive formulations	47
1. Upper bounds on the measurable independence number	49
2. Density and thick sets	52
3. The thick setting	56
4. The homogeneous setting	60
Chapter 5. Convergence through completely positive programming	65
1. Spaces of subsets	65

2. Three hierarchies and how they compare	67
3. Convergence of the completely positive hierarchy	70
4. Discussion and future directions	74
Chapter 6. Application: Witsenhausen's problem	77
1. Invariant positive-semidefinite kernels	79
2. The failure of the block moment hierarchy	80
3. The fix: another hierarchy	81
4. Implementation and verification of the bound	85
Acknowledgments	91
Part 3. Packings in compact spaces	93
Chapter 7. Copositive programming for compact packing graphs	95
1. Spaces of measures	97
2. Compact packing graphs	97
3. The moment and block moment hierarchies	99
4. The copositive hierarchy	101
5. Comparing the hierarchies	105
6. What about packing hypergraphs?	109
Chapter 8. Application: obtuse almost-equiangular sets	111
1. Preliminaries	115
2. The block moment hierarchy for almost-equiangular sets	115
3. Upper bounds from the three-point bound	121
4. Realizability of anti-triangle-free graphs	124
5. Maximum obtuse almost-equiangular sets	134
6. Classification in dimensions 2 and 3	142
Acknowledgments	144
Part 4. Euclidean space	145
Chapter 9. Distance-avoiding sets	147
1. Lattices, tori, and densities	148
2. Convergence of the completely positive hierarchy	151
3. Some discussion of the implementation	154
Chapter 10. Sphere packing	155
1. Sphere-packing densities	156
2. Some more harmonic analysis	157
3. Upper bounds on the maximal sphere-packing density	161

4. Exactness and convergence of the copositive hierarchy	164
5. Comparison to other known bounds	166
Conclusion and discussion	169
Discussion of the theory	169
Discussion of the computations	172
Appendix A. Preliminaries from analysis	175
1. Locally convex spaces, duality, and cones	175
2. p -integrable functions	177
3. Radon measures	178
4. Spaces of continuous linear operators	178
5. Invariant measures	179
Bibliography	181
Index	189
Dankwoord / Acknowledgements	195
List of publications	197
Curriculum vitae	199

Summary

Completely positive functions are an extension of completely positive matrices. They are known to characterize maximal spherical codes and maximum-density distance-avoiding subsets of $\mathbb{R}^n$ and certain compact metric spaces. This thesis expands this framework to related classes of problems in finite measure spaces and to the sphere-packing problem. For the latter, this is sharpened to show that the optimal sphere-packing density can be approximated using Schwartz functions.

Converging hierarchies of semidefinite programming bounds on the size of optimal spherical codes are known, based on approximations of completely positive functions and the Lovász theta number of a graph. This thesis extends these hierarchies to distance-avoiding sets and similar problems and to the sphere-packing problem, and proves their convergence to the maximum density. For distance-avoiding sets, additional hierarchies, such as the moment hierarchy, are introduced and shown to be stronger than the completely positive hierarchy, hence they also converge. These bounds are investigated for Witsenhausen's problem, which asks for the maximum fraction α_n of the n -dimensional unit sphere that is coverable by a set avoiding orthogonal pairs, obtaining the best known upper bounds on α_n in low dimensions.

The comparison of hierarchies for distance-avoiding sets moreover inspires a proof that the k -point bound for compact packing problems due to De Laat, Machado, Oliveira, and Vallentin is stronger than a converging completely positive hierarchy by Kuryatnikova and Vera. This proves convergence of the k -point bound. A related three-point bound is introduced for the t -almost-equiangular-set problem: finding the maximum size $\alpha(n, t)$ of a subset of the n -dimensional unit sphere in which every triple contains a pair with inner product $t \in [-1, 1)$. An analytic solution to this bound yields an enumeration of optimal constructions for $n = 2$ and 3 when $t \geq 0$.

Samenvatting

Volledig positieve functies zijn een uitbreiding van volledig positieve matrices. Het is bekend dat ze maximale sferische codes en afstandsvermijdende deelverzamelingen van $\mathbb{R}^n$ en bepaalde compacte metrische ruimten met maximale dichtheid karakteriseren. Dit proefschrift breidt dit raamwerk uit naar verwante klassen van problemen in eindige maatruimten en naar het bolstapelingsprobleem. Voor het laatste wordt dit verscherpt om aan te tonen dat de optimale bolstapelingsdichtheid benaderd kan worden met behulp van Schwartz-functies.

Convergerende hiërarchieën van bovengrenzen op de grootte van optimale sferische codes via semidefinietprogrammeren zijn bekend, gebaseerd op benaderingen van volledig positieve functies en het Lovász-thetagetal van een graaf. Dit proefschrift breidt deze hiërarchieën uit naar afstandsvermijdende verzamelingen en vergelijkbare problemen en naar het bolstapelingsprobleem, en bewijst hun convergentie naar de maximale dichtheid. Voor afstandsvermijdende verzamelingen worden aanvullende hiërarchieën, zoals de momentenhiërarchie, geïntroduceerd en aangetoond sterker te zijn dan de volledig positieve hiërarchie, en dus convergeren zij ook. Deze grenzen worden onderzocht voor het probleem van Witsenhausen, dat vraagt naar de maximale fractie α_n van de n -dimensionale eenheidssfeer die overdekt kan worden door een verzameling die orthogonale paren vermijdt. Dit leidt tot de best bekende bovengrenzen op α_n in lage dimensies.

De vergelijking van hiërarchieën voor afstandsvermijdende verzamelingen inspireert bovendien een bewijs dat de k -punts grens voor compacte stapelingsproblemen van De Laat, Machado, Oliveira en Vallentin sterker is dan een convergerende volledig positieve hiërarchie van Kuryatnikova en Vera. Dit bewijst de convergentie van de k -punts grens. Een verwante driepunts grens wordt geïntroduceerd voor het t -bijna-gelijkhoekige-verzamelingprobleem: het vinden van de maximale grootte $\alpha(n, t)$ van een deelverzameling van de n -dimensionale eenheidssfeer waarin elk drietal een paar bevat met inwendig

product $t \in [-1, 1)$. Een analytische oplossing van deze grens maakt een opsomming van de optimale constructies voor $n = 2$ en 3 wanneer $t \geq 0$ mogelijk.

CHAPTER 1

Introduction

The interest in completely positive programming originates from the ubiquity of problems of the form

$$\min_{x \in \Delta_S} x^\top Q x. \quad (1)$$

Here, S is a finite set, $Q \in \mathbb{R}^{S \times S}$ is a real symmetric matrix, and Δ_S is the standard simplex in $\mathbb{R}^S$: the set of all nonnegative vectors whose coefficients sum to 1. By writing $x^\top Q x = \text{Tr}(Q x x^\top)$, where Tr denotes the matrix trace, and noting that $x \in \Delta_S$ if and only if $x \geq 0$ and $\sum_{i,j} x_i x_j = 1$, it is apparent that (1) is equivalent to an optimization problem over the intersection of the affine space $\{A \in \mathbb{R}^{S \times S} : \sum_{i,j} A_{i,j} = 1\}$ and the extreme rays of the completely positive cone, which is the convex cone generated by matrices of the form $x x^\top$, with x a nonnegative vector. The study of problems of this kind was initiated in the works by Bomze [17] and Bomze, Dür, De Klerk, Roos, Quist, and Terlaky [18].

Calculating the independence number $\alpha(G)$ of a finite graph G is an example of an NP-hard problem [53] that can be expressed in the form (1). The independence number is the maximum cardinality a set of vertices can have without containing edges. Denoting the vertex set of G by V , the adjacency matrix of G by A , and the identity matrix by I , Motzkin and Straus [76] showed that

$$\frac{1}{\alpha(G)} = \min_{x \in \Delta_V} x^\top (A + I) x. \quad (2)$$

This is not terribly interesting in itself: as computing the independence number is NP-hard, it just means that optimizing over the completely positive cone is difficult. However, in his thesis, Parrilo [84] introduced systems of linear matrix inequalities that approximate the copositive cone—the conic dual of the completely positive cone—in finitely many steps. De Klerk and Pasechnik [54, 55], Bomze and De Klerk [19], and Peña, Vera, and Zuluaga [85] investigated this method closer, and used it to define linear and semidefinite

optimization hierarchies that approximate the independence number of a finite graph, also in finitely many steps.

This thesis presents extensions of these ideas to four classes of questions from extremal geometry that are modelled as a kind of independence-number problem on an infinite hypergraph. The following are typical examples of a problem in each class. The dimension n is fixed.

- I What is the maximum fraction of the $(n-1)$ -dimensional unit sphere that a set can cover without containing pairs of orthogonal vectors?
- II How many unit balls can touch a central unit ball, if their interiors do not overlap?
- III What is the maximum fraction of n -dimensional Euclidean space that a set can cover without containing pairs at distance 1?
- IV What is the maximum fraction of n -dimensional Euclidean space that a set of $(n-1)$ -dimensional unit balls can cover, if their interiors do not overlap?

Each of these has its own characteristics, but all can be formulated as an independence number of a graph $G = (V, E)$, where the correct notion of “size” of an independent set is not necessarily its cardinality.

For example, for Problem I, take $V = S^{n-1}$ the unit sphere, and

$$E = \{ (x, y) \in (S^{n-1})^2 : x^\top y = 0 \}.$$

In these terms, Problem I asks for the largest fraction of the sphere that is covered by an independent set of G . Here, “fraction” means the total surface measure of such a set, since such a set in general is not finite. Indeed, a spherical cap of nonzero angular radius less than $\pi/4$ is a valid construction which contains uncountably many points. Hence, we take the uniform probability measure μ on S^{n-1} , with which the question can be formulated as: find

$$\sup \{ \mu(I) : I \subseteq S^{n-1} \text{ measurable and independent} \}.$$

This problem is called *Witsenhausen’s problem*, and was first posed in 1974 by Witsenhausen [102]. It is conjectured that the optimal value is given by twice the measure of a spherical cap of angular radius $\pi/4$ [52, Conjecture 2.8], which is $(1/\sqrt{2} + o(1))^n$. This has only been confirmed for $n = 2$. Linear optimization bounds were introduced by Bachoc, Nebe, Oliveira, and Vallentin [5] and Oliveira [80]. Parts 1 and 2 of this thesis give a theory of completely positive programming for this class of problems, which results in the best bounds known on Witsenhausen’s problem in low dimensions. These results were partially published in [12].

Problem II is an example of a problem where the objective value is the cardinality of a set; it is called the *kissing-number problem*. Here, again, we

take $V = S^{n-1}$, but now $E = \{ (x, y) \in (S^{n-1})^2 : x^\top y \in (1/2, 1) \}$. The kissing number asks for the largest size $|I|$, where I is an independent set of this graph. To see that this is the correct choice for the edge set, radially project a valid configuration of balls onto the central ball. Each ball touching the central ball corresponds to a spherical cap of angular radius $\pi/3$ under this projection, explaining the maximally allowed inner product $\cos(\pi/3) = 1/2$.

Delsarte, Goethals, and Seidel [38] introduced linear programming bounds to the kissing-number problem. They did this for a more general class of problems, where the inner product $1/2$ is replaced by any inner product. After many improvements on this bound and adaptations to other problems, Bachoc and Vallentin [6] gave the first semidefinite programming bounds on the kissing number. More details on the history are presented in Chapter 7.

A theory of copositive programming for problems of this type was initiated by Kuryatnikova and Vera [57, 56]. In Part 3 we complement their work by comparing it to other well-known optimization methods. We also explore a problem on a 3-uniform hypergraph that is similar to Problem II. These results are based on the preprints [13] and [4].

Problems III and IV take place on a graph with vertex set $V = \mathbb{R}^n$. These are examples of problems where even the Lebesgue measure of a typical set satisfying the requirements is infinite; instead, the objective value is a limit over local densities, which we make more precise in Part 4. For now, we denote it by $\bar{\delta}(I)$ without definition.

Problem III is known as the *1-avoiding-set problem*. The edge set associated to it is $E = \{ (x, y) \in (\mathbb{R}^n)^2 : \|x - y\| = 1 \}$. The problem asks for the supremum $m_1(\mathbb{R}^n)$ of $\bar{\delta}(I)$ ranging over all measurable independent sets of the graph $(\mathbb{R}^n, E)$. Recently, Ambrus, Csiszárík, Matolcsi, Varga, and Zsámboi [2], showed that $m_1(\mathbb{R}^2) \leq 1/4$, settling a conjecture by Erdős [42]. Their method was based on convex programming bounds first introduced by Oliveira and Vallentin [82]. Chapter 9 contains more background on the problem.

Problem IV is called the *sphere packing problem*. It can be formulated as an independent set problem on the graph $(\mathbb{R}^n, E)$, with $E = \{ (x, y) \in (\mathbb{R}^n)^2 : \|x - y\| \in (0, 2) \}$. Again, the objective is to maximize $\bar{\delta}(S)$ over some sets S . However, the sets are not directly independent sets of the graph, but rather we are looking for

$$\sup \left\{ \bar{\delta} \left(\bigcup_{x \in I} B_1(x) \right) : I \text{ measurable and independent} \right\},$$

where $B_1(x)$ is the unit ball centered at x . In other words, at every point of an independent set we attach a unit ball centered at that point; it is the density of such a set we are interested in.

The sphere packing problem is perhaps the most famous of these four problems. The most notable results in this topic are the linear programming bound by Cohn and Elkies [27] and the proof that the Cohn-Elkies bound is exact for $n = 8$ by Viazovska [100], and for $n = 24$ by and Cohn, Kumar, Radchenko, and Viazovska [28], including calculations of the optimal values. The introduction of Chapter 10 goes deeper into this history.

Chapter 3 investigates the completely positive cone for problems like Problem III and Problem IV on spaces similar to $\mathbb{R}^n$. In Part 4 all the work in this thesis comes together to describe completely positive and copositive programming approaches to Problem III and IV. These results are new and unpublished.

I have included extensive preliminaries. To keep the main text to the point, I have moved the most elementary of these to the appendix. Notions and theorems that cannot be found in the preliminary section of the relevant chapter can probably be found there.

1. Some notation

The natural numbers start at 0. For integer $n \geq 1$, $[n] = \{1, \dots, n\}$. Given an index set I , we will often abbreviate $\{x_i : i \in I\}$ as $\{x_i\}_{i \in I}$; when I is ordered, and we intend a collection of x_i s to be ordered accordingly, we use the common notation $(x_i)_{i \in I}$. If S is a set of real-valued vectors, and $r \in \mathbb{R}$, then $S_{\geq r}$ is the set of all vectors in S with all coordinates at least r . In particular, for every $r \in \mathbb{R}$, $\mathbb{N}_{\geq r}$ is the set of natural numbers greater than or equal to r .

Let $k \in \mathbb{N}_{\geq 1}$, denote by $\mathfrak{S}_k$ the permutation group on k letters. Fix S a finite set. The space $\text{Sym}(S, k)$ is the vector space of *symmetric k -tensors* on S , that is, functions $S^k \rightarrow \mathbb{R}$ such that for all elements $s_1, \dots, s_k \in S$ and permutations $\pi \in \mathfrak{S}_k$: $T(s_{\pi 1}, \dots, s_{\pi k}) = T(s_1, \dots, s_k)$. The space $\text{Sym}(S, 2)$ is nothing more than the space of symmetric matrices indexed by S , and we will often denote it by $\text{Sym}(S)$. If $n \in \mathbb{N}$, then $\text{Sym}(n, k) = \text{Sym}([n], k)$, and $\text{Sym}(n) = \text{Sym}([n])$. We will use both the function notation $T(s_1, \dots, s_k)$ and the index notation $T_{s_1, \dots, s_k}$.

Denote the transpose of a vector $x \in \mathbb{R}^S$ by $x^\top$. We always consider the inner product of x and $y \in \mathbb{R}^S$ to be $x^\top y$, hence, unless indicated otherwise, the norm $\|x\| = \sqrt{x^\top x}$. The inner product on $\text{Sym}(S, k)$ is the usual Euclidean inner product, and is denoted $\langle \cdot, \cdot \rangle$. Indicate the cone of positive semidefinite matrices by

$$\text{Sym}(S)_{\geq 0} = \{ M \in \text{Sym}(S) : \langle M, xx^\top \rangle \geq 0 \text{ for all } x \in \mathbb{R}^S \}.$$

If $M \in \text{Sym}(S)_{\geq 0}$, we write $M \succeq 0$.

Let $k \in \mathbb{N}_{\geq 1}$. A choice of k elements $x_1, \dots, x_k \in \mathbb{R}^S$ defines a k -tensor

$$(x_1 \otimes \dots \otimes x_k)_{i_1, \dots, i_k} = (x_1)_{i_1} \dots (x_k)_{i_k}.$$

The vector $\mathbf{1} \in \mathbb{R}^S$ indicates the all-one vector— $\mathbf{1}_s = 1$ for all $s \in S$ —and $J = \mathbf{1}\mathbf{1}^\top$ is the all-one matrix. For a set $I \subseteq S$, the vector $\mathbf{1}_I$ is the indicator function of I .

Given a set S and $r \in \mathbb{N}$, the set $\text{Sub}_r(S)$ is the family of all subsets of S of cardinality at most r , including the empty set. Likewise, $\text{Sub}_{=r}(S)$ is the family of subsets of cardinality exactly r . Again, if $n \in \mathbb{N}$, then $\text{Sub}_r(n) = \text{Sub}_r([n])$ and $\text{Sub}_{=r}(n) = \text{Sub}_{=r}([n])$. We will call a set of cardinality r an r -set.

When we say that $H = (V, E)$ is a hypergraph, we mean that V is its vertex set and E is its edge set, which is a collection of subsets of V . For edges $\{v_1, \dots, v_k\} \in E$, we often omit the brackets, and write $v_1 \dots v_k \in E$. For an integer $k \geq 2$, a hypergraph is k -uniform if $E \subseteq \text{Sub}_{=k}(V)$. In this case, we will often interpret E as a symmetric subset of V^k by identifying it with the set of all k -tuples $(v_1, \dots, v_k)$ such that $\{v_1, \dots, v_k\} \in E$. The *automorphism group* of H is the set of bijections $\sigma : V \rightarrow V$ such that $\sigma(v_1) \dots \sigma(v_k) \in E$ if and only if $v_1 \dots v_k \in E$, and is denoted $\text{Aut}(H)$. We denote the action of an element $\sigma \in \text{Aut}(H)$ without brackets: $\sigma v = \sigma(v)$ for all $v \in V$. A hypergraph is called *vertex transitive* if the action of $\text{Aut}(V)$ is transitive on V .

2. The independence number of a finite graph

Let $k \geq 2$ be an integer and $H = (V, E)$ be a k -uniform hypergraph with vertex set V and edge set E . An *independent set* of H is a subset $S \subseteq V$ such that no k -subset of S is an edge. The *independence number* $\alpha(H)$ of H is

$$\alpha(H) = \sup\{|I| : I \subseteq V \text{ independent}\}$$

and is either attained or infinite. In this thesis we will discuss upper bounds on several extensions of the independence number to hypergraphs with a possibly infinite vertex set. The remainder of this chapter forms an overview of the finite-graph setting.

Fix a finite set V and a graph $G = (V, E)$. A common starting point for optimization methods for the independence number is the *Lovász theta number* of G , introduced by Lovász [70]. It is the semidefinite program

$$\begin{aligned} \vartheta(G) = \sup \quad & \sum_{v,w \in V} X_{vw} \\ & \sum_{v \in V} X_{vv} = 1, \\ & X_{vw} = 0 \quad \text{for all } vw \in E, \\ & X \in \text{Sym}(V)_{\succeq 0}. \end{aligned} \tag{3}$$

We use the same name and symbol for an optimization problem and its optimal value; so, “Lovász theta number” refers to both the program displayed in (3) and the number it produces.

The Lovász theta number is an upper bound on the independence number of G , that is $\vartheta(G) \geq \alpha(G)$. Indeed, and take $I \subseteq V$ an independent set. Then, the matrix $\mathbf{1}_I \mathbf{1}_I^\top / |I|$ is a feasible solution of $\vartheta(G)$ with objective $|I|$.

The Lovász theta number is appealing from a computational perspective, as the ellipsoid method offers a proof that it can be solved in polynomial time to any fixed precision [49]. In practice, interior point methods offer a polynomial-time implementation which is preferred. Moreover, the theta number is amenable to many well-known techniques, like restriction of its feasible region and sparsity arguments, to improve the bound and reduce the size of the program. Symmetries of the graph and of the constraints and objective can also often be exploited to the same end.

We will say that a bound on $\alpha(G)$ is *sharp* or *exact* if it is equal to $\alpha(G)$. Rarely is $\vartheta(G)$ a sharp bound on $\alpha(G)$. However, there are many ways to add or modify constraints, which has produced many of the best upper bounds on $\alpha(G)$.

One of the ways in which the theta number can be strengthened, is by replacing the positive-semidefinite cone by the *completely positive cone*

$$\text{CP}(V) = \left\{ \sum_{i=1}^m x_i x_i^\top : m \in \mathbb{N} \text{ and } x_i \in \mathbb{R}_{\geq 0}^V \text{ for } i \in [m] \right\}.$$

Let for any convex cone $\mathcal{C} \subseteq \text{Sym}(V)$,

$$\begin{aligned} \vartheta(G, \mathcal{C}) = \sup \quad & \sum_{v,w \in V} X_{vw} \\ & \sum_{v \in V} X_{vv} = 1, \\ & X_{vw} = 0 \quad \text{for all } vw \in E, \\ & X \in \mathcal{C}, \end{aligned}$$

so that $\vartheta(G, \text{Sym}(V)_{\succeq 0}) = \vartheta(G)$. Then, $\vartheta(G, \text{CP}(V))$ is a sharp bound on $\alpha(G)$ which is closely related to the formulation (2) of $1/\alpha(G)$. That it is an upper bound on $\alpha(G)$, follows from the same argument as for the theta number. It is moreover exact, since the optimal value is attained at an extreme point, and the extreme points of the feasible region are of the form $xx^\top$ with $x \geq 0$, $\|x\| = 1$, and the support of x is an independent set. Take an optimal solution $xx^\top$, with $x \geq 0$ a vector with support I . It has objective value $(\mathbf{1}_I^\top x)^2 \leq \|\mathbf{1}_I\|^2 \|x\|^2 = |I|$ by the Cauchy-Schwarz inequality, hence the conclusion follows. We call $\vartheta(G, \text{CP}(V))$ a *completely positive formulation* of $\alpha(G)$ for finite graphs G . In Chapter 4 we will see a proof of a much more general statement.

Instead of optimizing over $\text{CP}(V)$, upper bounds on the independence number are often obtained by optimizing over its conic dual

$$\text{COP}(V) = \{ M \in \text{Sym}(V) : \langle M, xx^\top \rangle \geq 0 \text{ for all } x \in \mathbb{R}_{\geq 0}^V \},$$

which is called the *copositive cone*. The accompanying optimization problem is $\vartheta^*(G, \text{COP}(V))$, where for any convex cone $\mathcal{C} \subseteq \text{Sym}(V)$,

$$\begin{aligned} \vartheta^*(G, \mathcal{C}) &= \inf t \\ F_{vv} &= t - 1 \quad \text{for all } v \in V, \\ F_{vw} &= -1 \quad \text{for all } v \neq w \text{ and } vw \notin E, \\ F &\in \mathcal{C}. \end{aligned}$$

If $\mathcal{C}$ is a closed convex cone, the program $\vartheta^*(G, \mathcal{C})$ is the dual of $\vartheta(G, \mathcal{C}^*)$, and under mild conditions, strong duality holds; that is $\vartheta(G, \mathcal{C}) = \vartheta^*(G, \mathcal{C}^*)$. In particular, we have $\vartheta(G, \text{CP}(V)) = \vartheta^*(G, \text{COP}(V))$, so that the latter also is an exact bound on $\alpha(G)$. We call $\vartheta^*(G, \text{COP}(V))$ the *copositive formulation* of $\alpha(G)$ for finite graphs. Of course, that these programs return exactly the independence number must mean that they are hard to compute. In Section 3 of this introduction we go further into this.

Other well-studied ways to improve the Lovász theta number stem from the moment hierarchy, which was developed for general 0-1 programming by Lasserre [65, 66] and described in more detail for the independence number by Laurent [67]. For $r \in \mathbb{N}$, let M_r be the operator

$$M_r : \mathbb{R}^{\text{Sub}_{2r}(V)} \rightarrow \text{Sym}(\text{Sub}_r(V)), \quad (M_r \nu)_{S,T} = \nu_{S \cup T}.$$

The *moment hierarchy*—also called the *Lasserre hierarchy*—is the sequence of programs

$$\begin{aligned} M_r(G) = \sup \quad & \sum_{v \in V} \nu_{\{v\}} \\ & \nu_{\emptyset} = 1, \\ & \nu_S = 0 \quad \text{for all } S \text{ not independent,} \\ & \nu \in \mathbb{R}_{\geq 0}^{\text{Sub}_{2r}(V)}, \quad M_r \nu \in \text{Sym}(\text{Sub}_r(V))_{\geq 0}. \end{aligned}$$

For all r , $M_r(G)$ is an upper bound on $\alpha(G)$, and the sequence is decreasing: for an integer $r \geq 1$ and an independent set I , define the vector $\chi_I \in \mathbb{R}_{\geq 0}^{\text{Sub}_{2r}(V)}$, which is 1 on sets $S \in \text{Sub}_r(I)$ and 0 otherwise. This defines a feasible solution to $M_r(G)$. Moreover, extending a feasible solution ν of $M_r(G)$ by zeros gives a feasible solution of $M_s(G)$ for all $s \geq r$. In fact, Laurent [67] showed that

$$M_1(G) \geq M_2(G) \geq \cdots \geq M_{\alpha(G)}(G) = \alpha(G).$$

We say that the hierarchy *converges* to $\alpha(G)$, since the sequence of numbers $(M_1(G), M_2(G), \dots)$ converges to $\alpha(G)$.

The program $M_1(G)$ is known as the *theta-prime number* and is equivalent to the theta number (3), but with the additional constraint that the matrix X is nonnegative. One way to think of the higher levels $M_r(G)$ is that they strengthen the theta number by including correlations between more than two points. We could thus call $M_r(G)$ a $(2r)$ -point bound.

The moment hierarchy is interesting because each level is a semidefinite program that can be solved in polynomial time, but on the other hand it converges to $\alpha(G)$ in finitely many steps. However, the time it takes to compute $M_r(G)$ goes up quickly with r for many graphs, even with efficient algorithms. Thus, already for small r the programs might not be tractable. For some graphs in this thesis, though, we can define a moment hierarchy for which low levels are tractable [61].

That the moment hierarchy might be difficult to compute explains the interest in weaker versions of these programs, which might give worse bounds, but are easier to compute. Schrijver [95] described a three-point bound for a certain combinatorial problem, which Bachoc and Vallentin [6] then used as inspiration for a three-point bound for the kissing-number problem, which was a breakthrough result. Musin [78] extended it to an r -point bound for the kissing-number problem for all $r \geq 2$. Gvozdenović, Laurent, and Vallentin introduced an r -point bound for the independence number of finite graphs [50], which De Laat, Machado, Oliveira, and Vallentin [60] extended to an r -point bound for topological packing graphs, a class that includes all finite graphs.

In this thesis, we investigate the bound by De Laat, Machado, Oliveira, and Vallentin [60], defined as follows. For $r \in \mathbb{N}_{\geq 2}$ and $Q \in \text{Sub}_{r-2}(V)$, define

$$M_Q : \mathbb{R}^{\text{Sub}_r(V)} \rightarrow \text{Sym}(\text{Sub}_1(V)), \quad (M_Q \nu)_{S,T} = \nu_{Q \cup S \cup T}.$$

The r th level of the *block moment hierarchy* is

$$\begin{aligned} \text{blockM}_r(G) = \sup \sum_{v \in V} \nu_{\{v\}} \\ \nu_{\emptyset} = 1, \\ \nu_S = 0 & \quad \text{for all } S \text{ not independent,} \\ M_Q \nu \in \text{Sym}(\text{Sub}_1(V))_{\succeq 0} & \quad \text{for all } Q \in \text{Sub}_{r-2}(V), \\ \nu \in \mathbb{R}_{\succeq 0}^{\text{Sub}_r(V)}. \end{aligned}$$

In the literature, it is called “the k -point bound”, but we will discuss many hierarchies that qualify for this name.

We may think of the matrix $M_Q \nu$ as a principal submatrix of a matrix of the form $M_r \mu$, where $(M_Q \nu)_{S,T} = (M_r \mu)_{Q \cup S, Q \cup T}$. Hence, the operator M_Q selects blocks indexed by sets containing Q . This shows that the block moment hierarchy is indeed weaker than the moment hierarchy, in the sense that there is a number $n \in \mathbb{N}$ such that for all i : $M_{i+n}(G) \leq \text{blockM}_i(G)$. We again have

$$\text{blockM}_1(G) \geq \text{blockM}_2(G) \geq \cdots \geq \alpha(G).$$

We prove that $\alpha(G) = \text{blockM}_{\alpha(G)^2}(G)$ for all finite graphs G in Chapter 5 and 7.

3. The copositive cone in finite dimensions

Let $G = (V, E)$ be a finite graph. In the previous section, we saw the copositive formulation $\vartheta^*(G, \text{COP}(V))$ for the independence number of G . Since computing $\vartheta^*(G, \text{COP}(V))$ is equivalent to computing $\alpha(G)$, it cannot be easier. Parrilo [84], and subsequently De Klerk and Pasechnik [54, 55], and Peña, Vera, and Zuluaga [85], introduced hierarchies of optimization problems based on the copositive formulation of $\alpha(G)$. These problems are computationally much less expensive than $\vartheta^*(G, \text{COP}(V))$, and converge in finitely many steps to $\alpha(G)$: the weakest of these hierarchies can be shown to converge to $\alpha(G)$ in at most $\alpha(G)^2$ steps, after rounding down to an integer [54, Theorem 4.1].

This section describes the background necessary to understand the weakest of these hierarchies better, which we will use to develop more general theory in Chapters 2 and 3. In particular, we define cones $C_r(V)$ for finite V that approximate $\text{COP}(V)$ in the sense that $\text{COP}(V)$ is the closure of $\bigcup_{r \in \mathbb{N}} C_r(V)$. This defines the hierarchy $\vartheta^*(G, C_r(V))$, which converges to $\vartheta^*(G, \text{COP}(V))$ —equal to $\alpha(G)$ —in finitely many steps. The theory will be presented more generally, and these concepts are defined for k -uniform hypergraphs, with $k \in \mathbb{N}_{\geq 2}$.

Convergence of this hierarchy is not proved here, as this thesis contains two generalizations of this statement, one in Chapter 5 and another in Chapter 7.

The copositive hierarchy departs from Pólya's theorem, a positivstellensatz for homogeneous polynomials. Let $p \in \mathbb{R}[x]$ with $x = (x_1, \dots, x_n)$ be a homogeneous polynomial in n variables. If there is an $r \in \mathbb{N}$ such that all coefficients of

$$(\mathbf{1}^\top x)^r p(x) = (x_1 + \dots + x_n)^r p(x), \quad (4)$$

are nonnegative, then $p(x) \geq 0$ for all $x \geq 0$. Pólya's theorem offers a partial converse: if $p(x) > 0$ for all nonzero $x \geq 0$, then there exists an $r \in \mathbb{N}$ such that all coefficients of (4) are nonnegative [89].

Define for a finite set V and an integer $k \geq 2$ the *completely positive cone*

$$\text{CP}(V, k) = \left\{ \sum_{i=1}^m x_i^{\otimes k} : m \in \mathbb{N} \text{ and } x_i \in \mathbb{R}_{\geq 0}^V \text{ for all } i \in [m] \right\}$$

and the *copositive cone*

$$\text{COP}(V, k) = \{ T \in \text{Sym}(V, k) : \langle T, x^{\otimes k} \rangle \geq 0 \text{ for all } x \in \mathbb{R}_{\geq 0}^V \},$$

which is the conic dual of $\text{CP}(V, k)$.

To see the connection between Pólya's theorem and the copositive cone, let V be a finite set with $n = |V|$, let $k \geq 2$ be an integer, and let T be in $\text{Sym}(V, k)$. The function $p : x \mapsto \langle T, x^{\otimes k} \rangle$ on $\mathbb{R}^V$ defines a homogeneous polynomial of degree k in n variables over $\mathbb{R}$. Conversely, for all integers $n \geq 0$ and $k \geq 2$, a homogeneous polynomial of degree k in n variables over $\mathbb{R}$ is given in this way by a unique element of $\text{Sym}(n, k)$.

Let $r \geq 0$ be an integer, and define for all $v_1, \dots, v_{k+r} \in V$,

$$\text{Av}_{\mathfrak{S}_{k+r}} T(v_1, \dots, v_{k+r}) = \frac{1}{(k+r)!} \sum_{\pi \in \mathfrak{S}_{k+r}} T(v_{\pi 1}, \dots, v_{\pi(k+r)}).$$

Rewrite (4) as

$$(\mathbf{1}^\top x)^r p(x) = \langle T \otimes \mathbf{1}^{\otimes r}, x^{\otimes(k+r)} \rangle = \langle \text{Av}_{\mathfrak{S}_{k+r}}(T \otimes \mathbf{1}^{\otimes r}), x^{\otimes(k+r)} \rangle, \quad (5)$$

so $\text{Av}_{\mathfrak{S}_{k+r}}(T \otimes \mathbf{1}^{\otimes r})$ is the unique symmetric $(k+r)$ -tensor that holds the coefficients of $(\mathbf{1}^\top x)^r p(x)$. With this, Pólya's theorem says that if T is in the algebraic interior of $\text{COP}(V, k)$, there exists an $r \in \mathbb{N}$ such that

$$\text{Av}_{\mathfrak{S}_{k+r}}(T \otimes \mathbf{1}^{\otimes r}) \geq 0.$$

For $r \in \mathbb{N}$ and V a finite set, define the *Pólya-type cone* by

$$C_r(V, k) = \{ T \in \text{Sym}(V, k) : \text{Av}_{\mathfrak{S}_{k+r}}(T \otimes \mathbf{1}^{\otimes r}) \geq 0 \},$$

which is a closed cone. This leads to the following conic formulation of Pólya's theorem, where algint denotes the algebraic interior.

THEOREM 1.1 (Pólya's theorem [89]). *For a finite set V and integer and $k \geq 2$,*

$$\text{algint COP}(V, k) \subseteq \bigcup_{r \in \mathbb{N}} C_r(V, k).$$

On the other hand, if $p(x) = \langle T, x^{\otimes k} \rangle$ and there is an $r \in \mathbb{N}$ such that $(1^\top x)^r p(x) \geq 0$, then, by Equation (5), $\langle T, x^{\otimes k} \rangle \geq 0$ for all $x \geq 0$ such that $x \neq 0$. This proves $C_r(V, k) \subseteq \text{COP}(V, k)$ for all r . If $r' \geq r$, then $(1^\top x)^{r'} p(x) \geq 0$ for all $x \geq 0$, so $C_r(V, k) \subseteq C_{r'}(V, k)$. Denoting the dual cone of $C_r(V, k)$ by $C_r(V, k)^*$ and, by Theorem A.2, taking the conic dual results in the following corollary of Pólya's theorem. We call a sequence of cones satisfying the conclusion of the Theorem an *outer approximation* of $\text{CP}(V, k)$.

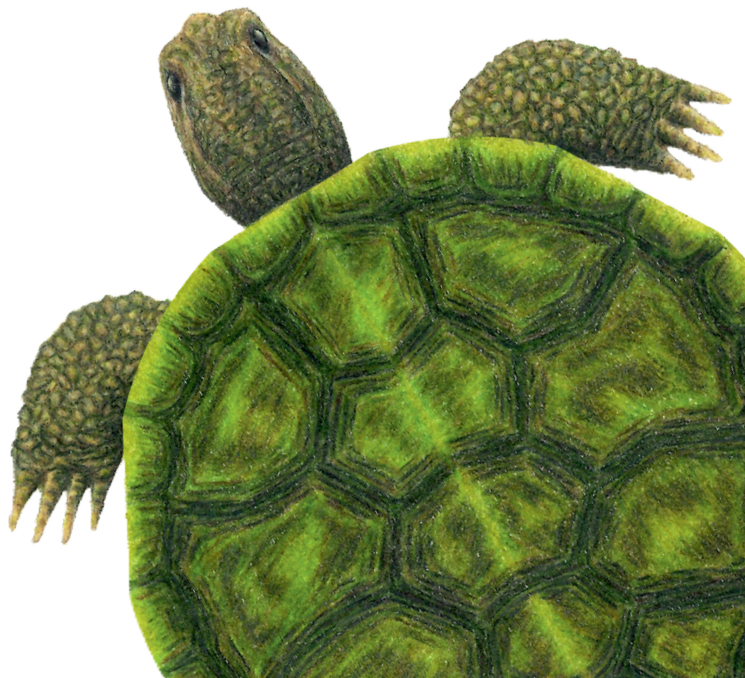
THEOREM 1.2. *For a finite set V and an integer $k \geq 2$,*

$$C_0(V, k)^* \supseteq C_1(V, k)^* \supseteq \cdots \supseteq \text{CP}(V, k) \quad \text{and} \quad \text{CP}(V, k) = \bigcap_r C_r(V, k)^*.$$

It should be said that the cones $C_r(V, k)$ give a weak approximation of the copositive cone. For example, for a finite graph $G = (V, E)$, the program $\vartheta^*(G, C_r(V))$ is infeasible if $r < \alpha(G) - 1$ [54, Theorem 4.2]. It is therefore unsurprising that since the conception of the topic, tighter hierarchies were studied. However, this thesis is mostly concerned with whether certain hierarchies of optimization problems converge, not with the rate of convergence. From this perspective it is interesting that such a weak hierarchy suffices.

Part 1

The completely positive cone



CHAPTER 2

The completely positive cone of a measure space

Problems I, III, and IV from the introduction have in common that they do not ask to optimize a cardinality, but a kind of density. This density is defined by a measure, which leads us to study spaces of p -integrable functions.

Adams [1] and DeCorte, Oliveira, and Vallentin [36] defined a completely positive cone of square-integrable functions on a finite measure space. An outer approximation of this cone by Pólya-type cones similar to Theorem 1.2 was introduced by Bekker, Kuryatnikova, Oliveira, and Vera [12].

In this chapter, we reproduce the results from [12], particularly Theorem 4.1 and Theorem 5.3 therein, by explicitly showing how the finite-dimensional Pólya-type cones lift to the square-integrable setting. The approach is new and comes with tools that are more generally applicable. As opposed to the approach in [12], it does not lean on the action of a locally compact group. Rather, group actions are considered later, in Chapter 3.

Especially the projective approximation lemma, Lemma 2.1, is new. It is an attempt at capturing and generalizing some folklore and intuition about infinite-dimensional analogues of finite-dimensional hierarchies.

The application of the martingale convergence theorem is put to the forefront, which is another difference with the original exposition. This replaces the continuity assumption and an application of a technical theorem by Powers and Reznick in the proof of [12, Theorem A.1]. Whereas [12] restricts itself to continuous and invariant kernels on compact spaces, the treatment here regards square-integrable kernels on finite countably generated measure spaces, but the conclusions are weaker. This approach lends itself to extension to σ -finite countably generated measure spaces, although the hierarchy there is not a direct generalization of the Pólya-type cones.

1. Notation and preliminaries

See the appendix for more background on locally convex analysis and p -integrable functions. Our main reference is Simon's book [97]. All topological vector spaces are Hausdorff, and all functions and measures are real-valued.

Convexity and duality. Section A.1 gives more details about duality and convexity.

Let X be a vector space. The *algebraic dual* X' of X is the space of all linear functionals of X . If X is a topological vector space, its *continuous dual* X^* is the subspace of X' consisting of continuous linear functionals. Let $A : X \rightarrow Y$ is a continuous linear map of topological vector spaces. If its continuous linear adjoint exists, denote it by A^* .

For us, all cones are convex. The convex hull of a subset S of a vector space X is denoted $\text{conv } S$. The closed convex hull of S is $\overline{\text{conv}} S = \text{cl conv } S$. The conic hull is $\text{cone } S$, and the closed conic hull of S is $\overline{\text{cone}} S = \text{cl cone } S$. If (X, Y) is a dual pair of vector spaces, denote the dual cone of a set S by S^* .

Measures and integrable functions. See Section A.2 for more background on measures and integrable functions.

Given a set V and a subset $\mathcal{P}$ of the power set of V , let $\sigma(\mathcal{P})$ be the σ -algebra generated by $\mathcal{P}$. Let $(V, \mathcal{A}, \mu)$ be a measure space with σ -algebra $\mathcal{A}$ and measure μ . One way to interpret μ is as a function from the set of measurable functions to $\mathbb{R} \cup \{\infty\}$ defined by integration over V , which leads to the notation $\mu(f) = \int_V f(x) d\mu(x)$. Denote the usual p -norms by $\|\cdot\|_p$ for $1 \leq p \leq \infty$.

The measure μ^k is the product measure on V^k with the naturally induced σ -algebra $\mathcal{A}^k$; we will sometimes write it as $\mu^{\otimes k}$. Unless specified otherwise, we always understand the product V^k to be equipped with the product measure μ^k , and we denote the naturally induced σ -algebra on V^k by $\mathcal{A}^k$. A square-integrable function f on V^k is called a k -*tensor*. It is called *symmetric* if it is invariant under permutation of its coordinates. We denote the space of μ -equivalence classes of symmetric k -tensors by $L_{\text{sym}}^2(V, k)$, and understand $L_{\text{sym}}^2(V) = L_{\text{sym}}^2(V, 2)$. If $f_1, \dots, f_k \in L^2(V)$, then $f_1 \otimes \dots \otimes f_k$ denotes the element of $L^2(V^k)$ given by

$$f_1 \otimes \dots \otimes f_k(v_1, \dots, v_k) = f_1(v_1) \dots f_k(v_k)$$

for all $v_i \in V$.

The notation $\langle f, g \rangle$ refers to the integral of the pointwise product $\mu(fg)$. If $f \in L^p(V)$ and $g \in L^q(V)$ with $p = 1$ and $q = \infty$, or if $1 < p < \infty$

and $1/p + 1/q = 1$, then the product fg is in $L^1(V)$ and $\langle f, g \rangle \in \mathbb{R}$. In this case the map $\langle \cdot, \cdot \rangle$ defines a duality

$$\langle \cdot, \cdot \rangle : L^p(V) \times L^q(V) \rightarrow \mathbb{R}$$

under the respective norm topologies, hence the notation.

For the following, see the appendix and [96] for more details on Hilbert-Schmidt and trace-class operators.

We isometrically identify the subspace of Hilbert-Schmidt operators with the space of kernel $L^2(V^2)$. Recall that a kernel is called *positive semidefinite* if for all $f \in L^2(V)$ we have $\langle K, f \otimes f \rangle \geq 0$. We denote the cone of positive-semidefinite kernels by $\text{PSD}(V)$.

Let $K \in L^2_{\text{sym}}(V)$. If K is of trace class and Φ an orthonormal basis, the quantity

$$\sum_{\phi \in \Phi} \langle K, \phi \otimes \phi \rangle \quad (6)$$

is independent of Φ and finite. It is called the *trace* of K , denoted $\text{Tr}K$. A positive-semidefinite kernel K is of trace class if and only if $\text{Tr}(K) < \infty$; in this case, $\text{Tr}(K) = \|K\|_{\mathcal{B}^1}$, where the latter denotes the trace norm of K as defined in A.4. If V is a compact Hausdorff topological space and K is positive semidefinite and continuous, then $\text{Tr}(K) = \int_V K(x, x) d\mu(x)$.

Continuous functions and Radon measures. Let V be a locally compact topological space. The space $C(V)$ is the space of continuous functions on V . The space $C_c(V)$ is the space of continuous functions with compact support; if V is compact, then $C_c(V) = C(V)$. The space $C_0(V)$ is the space of continuous functions on V that vanish outside compact sets: if $f \in C_0(V)$, then for every $\epsilon > 0$ there is a compact set $K \subseteq V$ such that $|f(x)| \leq \epsilon$ for all $x \in V \setminus K$. We equip $C_c(V)$ and $C_0(V)$ with the supremum norm. Then, $C_0(V)$ is the norm closure of $C_c(V)$ in $L^\infty(V)$.

Similar to spaces of square-integrable functions, we let $C_{\text{sym}}(V, k)$ be the subspace of $C_0(V^k)$ of functions that are invariant under permutation of their arguments. We will also write $C_{\text{sym}}(V) = C_{\text{sym}}(V, 2)$.

We denote the space of Radon measures by $M(V)$. By identifying it with the dual of $C_0(V)$, we equip it with the operator norm. Let $\mu \in M(V)$. The *support* of μ is

$$\text{supp } \mu = \bigcap_{\substack{U \subseteq V \text{ open} \\ \mu(U)=0}} (V \setminus U).$$

We say that a Radon measure has *full support* if $\mu(U) > 0$ for all open sets U , i.e. if $\text{supp } \mu = V$.

2. Projective approximation and martingales

At the core of this chapter lie two ideas. The first is that under mild conditions an outer approximation of a cone on a collection of simpler spaces lifts to a larger locally convex vector space. In this chapter, this means that we will take, for fixed $k \geq 2$, the outer approximation of $\text{CP}([n], k)$ for all n given in Theorem 1.1, and use it to give an outer approximation of the completely positive cone in $L_{\text{sym}}^2(V, k)$, which we define later for suitable measure spaces $(V, \mathcal{A}, \mu)$. The abstract framework for lifting these approximations is given by the projective approximation lemma, Lemma 2.1. The second idea that this chapter is built on, is that the theory of martingales gives rise to such a lifting. It offers the operators that figure in the projective approximation lemma applied to $L^2(V, k)$.

If X is a topological vector space and $\mathcal{C} \subset X$ is a cone, a continuous linear map $\sigma : X \rightarrow X$ is called a linear automorphism of $\mathcal{C}$ if it is an automorphism of X and $\sigma(\mathcal{C}) \subseteq \mathcal{C}$. If X is part of a dual pair and equipped with a dual topology, and if σ is an automorphism of $\mathcal{C}$, then σ^* is an automorphism of $\mathcal{C}^*$. For details on dual pairs and related topics, see Section 1 of this chapter and the appendix.

LEMMA 2.1 (Projective approximation lemma). *Let (X, Y) be a dual pair of vector spaces, I be a directed set, and $\{(X_i, Y_i)\}_{i \in I}$ be a collection of dual pairs of vector spaces. Let for all i the maps $A_i : X \rightarrow X_i$ and $B_i : Y \rightarrow Y_i$ be linear and continuous under the weak topologies, $\mathcal{C} \subseteq X$ be a closed convex cone and $\{\mathcal{C}_r\}_{r \in R}$ be a family of closed convex cones in X . If $B_i^* A_i$ converges weakly to a linear automorphism of $\mathcal{C}$, and*

- (i) $B_i^* A_i \mathcal{C}_r \subseteq \mathcal{C}_r$ for all i and r ,
- (ii) $B_i^* A_i \mathcal{C} \subseteq \mathcal{C}$ for all i ,
- (iii) $A_i \mathcal{C} = \bigcap_r A_i \mathcal{C}_r$ for all i ,

then $\mathcal{C} = \bigcap_r \mathcal{C}_r$.

Before we move to the proof of this rather technical lemma, let us first look at two examples that show how it can be applied.

Example 2.2. Let V be finite set and $k \geq 2$ be an integer, let $X = \text{Sym}(V, k)$ and let $\mathcal{C} = \text{CP}(V, k)$. Theorem 1.2 says that $\text{CP}(V, k) = \bigcap_r C_r(V, k)^*$. By applying a linear automorphism σ of $\text{CP}(V, k)$ to both sides, we find that $\text{CP}(V, k) = \bigcap_r \sigma(C_r(V, k)^*)$. This conclusion also follows from the projective approximation theorem with $I = \{0\}$ and $A_0 = \sigma$, $B_0 = I$. If σ does not stabilize the $C_r(V, k)^*$, the $\sigma(C_r(V, k)^*)$ form an outer approximation that is distinct from the original.

Example 2.3. Let $X = l^2(\mathbb{N})$ and $\mathcal{C} = l^2(\mathbb{N})_{\geq 0}$, the closed convex cone of nonnegative sequences. Let for $r \in \mathbb{N}$ the set $\mathcal{C}_r$ be the closed convex cone of sequences such that $x_i \geq 0$ for all $i \leq r$. It is of course more than clear that $\mathcal{C} = \bigcap_r \mathcal{C}_r$. We will use the projective approximation lemma to reach the same conclusion.

The space $l^2(\mathbb{N})$ is a Hilbert space with inner product $\langle x, y \rangle = \sum_{i \in \mathbb{N}} x_i y_i$. For every $i \in \mathbb{N}$, let $p_i : l^2(\mathbb{N}) \rightarrow \mathbb{R}^i$ be the projection onto the first i coordinates. Then, $p_i^* p_i$ is the map that sets all coordinates with index larger than i to zero. We see that $p_i^* p_i \mathcal{C} \subseteq \mathcal{C}$ and $p_i^* p_i \mathcal{C}_r = \mathcal{C}_r$ for all r and i . Moreover, the weak limit of $p_i^* p_i$ is the identity on $l^2(\mathbb{N})$.

For all $i \in \mathbb{N}$, $p_i \mathcal{C} = \mathbb{R}_{\geq 0}^i$ and if $r \geq i$, $p_i \mathcal{C}_r = \mathbb{R}_{\geq 0}^i$. For all $r < i$, $p_i \mathcal{C}_r$ is the cone of vectors in $\mathbb{R}^i$ with the first r coordinates nonnegative. Thus, for all i , $\bigcap_r p_i \mathcal{C}_r = p_i \mathcal{C} = \mathbb{R}_{\geq 0}^i = p_i \mathcal{C}$, so that all conditions of the projective approximation lemma are met, and the conclusion follows.

From these examples one might guess that it is usually not necessary—if not pedantic—to use the lemma. One would be correct. The lemma is also generic: it does not say anything what a “best” way of lifting an outer approximation is. Its main purpose is to offer some additional rigor and direction to this exposition.

Proof of Lemma 2.1. For all $i \in I$

$$A_i \mathcal{C} = \bigcap_{r \in R} A_i \mathcal{C}_r \subseteq \bigcap_{r \in R} B_i^{*-1} \mathcal{C}_r. \quad (7)$$

The first equality is implied by property (iii) and the inclusion holds by property (i). Moreover, by Theorem A.2(i) and (iii) and properties (ii) and (iii) above, for all $i \in I$

$$B_i \mathcal{C}^* \subseteq (B_i^{*-1} \mathcal{C})^* \subseteq (A_i \mathcal{C})^* = \left(\bigcap_{r \in R} A_i \mathcal{C}_r \right)^*. \quad (8)$$

Denote the duality of the pair (X, Y) by $\langle \cdot, \cdot \rangle$ and that of (X_i, Y_i) by $\langle \cdot, \cdot \rangle_i$. Weak convergence of $B_i^* A_i$ to a linear automorphism σ of $\mathcal{C}$ means that for all $x \in X$ and $y \in Y$

$$\langle x, y \rangle = \lim_{i \in I} \langle A_i \sigma^{-1} x, B_i y \rangle_i = \lim_{i \in I} \langle A_i x, B_i \sigma^{*-1} y \rangle_i. \quad (9)$$

If $x \in \mathcal{C}$ and $y \in (\bigcap_r \mathcal{C}_r)^*$, then $A_i \sigma^{-1} x \in \bigcap_r B_i^{*-1} \mathcal{C}_r$ by (7). Moreover, $B_i y \in B_i (\bigcap_r \mathcal{C}_r)^* \subseteq (\bigcap_r B_i^{*-1} \mathcal{C}_r)^*$ according to Theorem A.2(iii). It follows from (9) that $\langle x, y \rangle \geq 0$, and $\mathcal{C} \subseteq \bigcap_r \mathcal{C}_r$.

Furthermore, if $x \in \bigcap_r \mathcal{C}_r$ and $y \in \mathcal{C}^*$, then $A_i x \in A_i \bigcap_r \mathcal{C}_r \subseteq \bigcap_r A_i \mathcal{C}_r$ for all $i \in I$. The adjoint σ^* is an automorphism of $\mathcal{C}^*$, so $B_i \sigma^{*-1} y \in (\bigcap_r A_i \mathcal{C}_r)^*$ by (8), so that $\langle x, y \rangle \geq 0$ by (9). This establishes $\bigcap_r \mathcal{C}_r \subseteq \mathcal{C}$ by Theorem A.2(i) and concludes the proof. $\square$

This chapter not only presents a direct L^2 analogue of the tried-and-true Pólya-type cones on $C(V^2)$ for compact V , as introduced in [12, 56, 57], but also a stricter outer approximation of the completely positive cone that extends to a class of σ -finite measure spaces. Both can be understood from the perspective of the projective approximation lemma.

A special case of the theory of martingales offers the operators A_i and B_i of the projective approximation lemma for L^p spaces. For the purposes of this chapter and the next, it simply gives a specific way of describing a function as a limit of simple functions—finite linear combinations of step functions. Chapter 5 of the book by Edwards and Gaudry [41] contains a more complete presentation of the topic.

For a set V , we say that a partition $\mathcal{P}'$ of V *refines* a partition $\mathcal{P}$, if every element of $\mathcal{P}'$ is contained in an element of $\mathcal{P}$; notation $\mathcal{P}' \preceq \mathcal{P}$. Let $(V, \mathcal{A}, \mu)$ be a finite measure space. Call a sequence of partitions $(\mathcal{P}_n)_{n \in \mathbb{N}}$ of V with all $\mathcal{P}_n \subseteq \mathcal{A}$ a *finite-rank approximation* of $(V, \mathcal{A}, \mu)$ if:

- $\mathcal{P}_0 = \{V\}$;
- $\mathcal{P}_{n+1} \preceq \mathcal{P}_n$ for all $n \in \mathbb{N}$;
- each $\mathcal{P}_n$ is finite;
- $\mathcal{A} = \sigma(\bigcup_n \mathcal{P}_n)$.

Given a partition $\mathcal{P}$ of V , call the subset of $\mathcal{P}$ consisting of those sets that have nonzero measure $\mathcal{P}^+$. Define for $1 \leq p < \infty$ the operator

$$E_{\mathcal{P}} : L^p(V) \rightarrow L^p(V), \quad E_{\mathcal{P}} f = \sum_{P \in \mathcal{P}} \langle f, \mathbf{1}_P \rangle / \mu(P).$$

For a sequence of partitions $(\mathcal{P}_n)_{n \in \mathbb{N}}$, let $E_n = E_{\mathcal{P}_n}$.

THEOREM 2.4 (Martingale convergence theorem). *Let $(V, \mathcal{A}, \mu)$ be a finite measure space equipped with a finite-rank approximation $(\mathcal{P}_n)_{n \in \mathbb{N}}$. If $1 \leq p < \infty$, the sequence $(E_n)_{n \in \mathbb{N}}$ converges under the strong operator topology to the identity.*

If $f \in L^p(V)$ for $1 \leq p < \infty$, the sequence $(E_n f(v))_{n \in \mathbb{N}}$ converges to $f(v)$ for almost every $v \in V$.

The second statement of the theorem is quite strong, so it is not surprising that its proof extends beyond what fits in this thesis. It uses a clever but technical trick, which can be found in the book by its inventor [40, Theorem VII.4.1]. The first statement is easier; it uses that $\bigcup_n L^p(V, \sigma(\mathcal{P}_n), \mu)$ is dense in $L^p(V)$ and that the operators E_n are uniformly bounded [41, Theorem 5.2.6].

The functions $E_n f$ are called the *conditional expectations* of f with respect to $(\sigma(\mathcal{P}_n))_{n \in \mathbb{N}}$. They are usually interpreted as functions on the measure space $(V, \sigma(\mathcal{P}_n), \mu)$, in which case the sequence $(E_n f)_{n \in \mathbb{N}}$ is called the *martingale* associated to f . For us, there is no harm in thinking of $E_n f$ as an element of $L^p(V, \mathcal{A}, \mu)$.

Let us investigate how the martingale convergence theorem interacts with the projective approximation lemma. Let $k \geq 2$ be an integer. For given sets $P_1, \dots, P_k \subseteq V$ and a permutation $\pi \in \mathfrak{S}_k$, let $P_{\pi 1, \dots, \pi k}$ denote the set $P_{\pi 1} \times \dots \times P_{\pi k}$. We will be interested in symmetric square-integrable k -tensors, that is, elements of $L^2_{\text{sym}}(V, k)$.

Let $(V, \mathcal{A}, \mu)$ be a finite measure space with a finite partition $\mathcal{P}$. Note that $\mathcal{P}$ induces a partition on V^k , namely

$$\{P_{1, \dots, k} : P_i \in \mathcal{P} \text{ for all } i\}.$$

When $(\mathcal{P}_n)_{n \in \mathbb{N}}$ is a finite-rank approximation of $(V, \mathcal{A}, \mu)$, this construction induces a finite-rank approximation of $(V^k, \mathcal{A}^k, \mu^k)$. We denote the conditional expectation operators on $L^2_{\text{sym}}(V, k)$ with respect to this induced partition by $E_{\mathcal{P}}$ as well.

Define the bounded operators $A_{\mathcal{P}}, B_{\mathcal{P}} : L^2_{\text{sym}}(V, k) \rightarrow \text{Sym}(\mathcal{P}^+, k)$ by

$$A_{\mathcal{P}} K = (\langle K, \mathbf{1}_{P_{1, \dots, k}} \rangle)_{P_1, \dots, P_k \in \mathcal{P}^+} \text{ and } B_{\mathcal{P}} K = \left(\frac{\langle K, \mathbf{1}_{P_{1, \dots, k}} \rangle}{\mu^k(P_{1, \dots, k})} \right)_{P_1, \dots, P_k \in \mathcal{P}^+},$$

with continuous adjoints

$$A_{\mathcal{P}}^* T = \sum_{P_1, \dots, P_k \in \mathcal{P}^+} T_{P_1, \dots, P_k} \mathbf{1}_{P_{1, \dots, k}} \text{ and } B_{\mathcal{P}}^* T = \sum_{P_1, \dots, P_k \in \mathcal{P}^+} T_{P_1, \dots, P_k} \frac{\mathbf{1}_{P_{1, \dots, k}}}{\mu^k(P_{1, \dots, k})}.$$

If $(\mathcal{P}_n)_n$ is a finite-rank approximation of V , denote $A_n = A_{\mathcal{P}_n}$ and $B_n = B_{\mathcal{P}_n}$. Item (i) of the following lemma together with the martingale convergence theorem imply that for a finite-rank approximation $(\mathcal{P}_n)_{n \in \mathbb{N}}$ the weak limit of $B_n^* A_n$ is the identity, which is the first requirement of the projective approximation lemma.

LEMMA 2.5. *Let $(V, \mathcal{A}, \mu)$ be a finite measure space with a partition $\mathcal{P}$. For all integers $k \geq 2$:*

- (i) $E_{\mathcal{P}} = B_{\mathcal{P}}^* A_{\mathcal{P}}$;
- (ii) $A_{\mathcal{P}} B_{\mathcal{P}}^* = I$, the identity on $\text{Sym}(\mathcal{P}^+, k)$;
- (iii) $A_{\mathcal{P}}(f^{\otimes k}) = (A_{\mathcal{P}} f)^{\otimes k}$ for all $f \in L^2(V)$ and $B_{\mathcal{P}}^*(w^{\otimes k}) = (B_{\mathcal{P}}^* w)^{\otimes k}$ for all $w \in \mathbb{R}^{\mathcal{P}^+}$.

Proof. (i) This follows directly from the definition of the maps.

(ii) This is true because $\{\mathbf{1}_{P_1, \dots, k} / \sqrt{\mu^k(P_1, \dots, k)} : P_1, \dots, P_k \in \mathcal{P}^+\}$ is an orthonormal set in $L^2(V)$.

(iii) Indeed,

$$(A_{\mathcal{P}}(f^{\otimes k}))_{P_1, \dots, P_k} = \langle f^{\otimes k}, \mathbf{1}_{P_1, \dots, k} \rangle = \prod_{i=1}^k \langle f, \mathbf{1}_{P_i} \rangle = \prod_{i=1}^k (A_{\mathcal{P}} f)_{P_i},$$

and

$$B_{\mathcal{P}}^*(w^{\otimes k}) = \sum_{P_1, \dots, P_k \in \mathcal{P}^+} w_{P_1} \cdots w_{P_k} \frac{\mathbf{1}_{P_1, \dots, k}}{\mu^k(P_1, \dots, k)} = \left(\sum_{P \in \mathcal{P}^+} w_P \frac{\mathbf{1}_P}{\mu(P)} \right)^{\otimes k}. \quad \square$$

We end this section with a sufficient condition for a measure space to locally admit finite-rank approximations, and for $L^p(V)$ to be separable.

LEMMA 2.6. *If $(V, \mathcal{A}, \mu)$ is countably generated and σ -finite, then every finite measure subspace admits a finite-rank approximation and $L^p(V)$ is separable for all $1 \leq p < \infty$.*

Proof. Let $A \in \mathcal{A}$ have finite measure, and denote the power set of A by $\mathcal{P}(A)$. Let $\{U_0, U_1, \dots\}$ be a countable generator of $\mathcal{A}$. Define $\mathcal{P}_0 = \{A\}$, and define $\mathcal{P}_n$ recursively by

$$\mathcal{P}_{n+1} = \{P \cap U \cap A : P \in \mathcal{P}_n, U \in \{U_n, V \setminus U_n\}\}.$$

Then, $\sigma(\bigcup_n \mathcal{P}_n) = \mathcal{A} \cap \mathcal{P}(A)$, $\mathcal{P}_{n+1}$ refines $\mathcal{P}_n$ for all $n \in \mathbb{N}$, and each $\mathcal{P}_n$ contains only finitely many sets, so $(\mathcal{P}_n)_{n \in \mathbb{N}}$ is a finite-rank approximation of $(A, \mathcal{A} \cap \mathcal{P}(A), \mu)$.

That for all $1 \leq p < \infty$ the space $L^p(V)$ is separable is Proposition 3.4.5. in [25]. $\square$

3. A Pólya-type approximation on finite measure spaces

For a measure space $(V, \mathcal{A}, \mu)$, the cone of *completely positive* k -tensors is

$$\text{CP}(V, k) = \overline{\text{cone}}\{f^{\otimes k} \in L_{\text{sym}}^2(V, k) : f \in L^2(V)_{\geq 0}\}.$$

A k -tensor is called *copositive* if it is in $\text{CP}(V, k)^*$, that is, when it is in

$$\text{COP}(V, k) = \{T \in L_{\text{sym}}^2(V, k) : \langle T, f^{\otimes k} \rangle \geq 0 \text{ for all } f \in L^2(V)_{\geq 0}\}.$$

We derive a version of Theorem 1.2 for these cones when $(V, \mathcal{A}, \mu)$ is countably generated and finite.

From here on, assume $(V, \mathcal{A}, \mu)$ is countably generated and finite. We follow the efforts of Kuryatnikova and Vera [56, 57] and Bekker, Kuryatnikova, Oliveira and Vera [12]. For $r \in \mathbb{N}$, define $\mathcal{T}_r : L^2(V^k) \rightarrow L^2(V^{k+r})$ as

$$\mathcal{T}_r K(v_1, \dots, v_{k+r}) = \frac{1}{(k+r)!} \sum_{\pi \in \mathfrak{S}_{k+r}} K(v_{\pi(1)}, \dots, v_{\pi(k)}).$$

This is a bounded operator, as $\mathcal{T}_r K = \text{Av}_{\mathfrak{S}_{k+r}}(K \otimes \mathbf{1}^{\otimes r})$, and the maps $\text{Av}_{\mathfrak{S}_{k+r}}$ and $K \mapsto K \otimes \mathbf{1}^{\otimes r}$ are bounded. The *Pólya-type cones* of $L_{\text{sym}}^2(V, k)$ are

$$C_r(V, k) = \mathcal{T}_r^{-1} L_{\text{sym}}^2(V, k+r)_{\geq 0} \quad \text{and} \quad C_r(V, k)^* = \mathcal{T}_r^* L_{\text{sym}}^2(V, k+r)_{\geq 0}.$$

Continuity of $\mathcal{T}_r$ implies that the cones are closed. The adjoint of $\mathcal{T}_r$ is given by, for almost all k -tuples $(v_1, \dots, v_k)$,

$$\mathcal{T}_r^*(F)(v_1, \dots, v_k) = \frac{1}{(k+r)!} \int_{V^r} \sum_{\pi \in \mathfrak{S}_{k+r}} F(\pi(v_1, \dots, v_k, v)) d\mu^r(v); \quad (10)$$

here $v \in V^r$ while each v_i is in V , and for $\pi \in \mathfrak{S}_k$ and $v \in V^{k+r}$, we define $\pi(v) = (v_{\pi(1)}, \dots, v_{\pi(k+r)})$. If F is a symmetric tensor, the sum over the permutation group and the factor $1/(k+r)!$ disappear.

The claim is of course that if $(\mathcal{P}_n)_{n \in \mathbb{N}}$ is a finite-rank approximation of V with associated operators A_n and B_n as defined in Section 2, the cones $\text{CP}(V, k)$ and $C_r(V, k)^*$ with these operators satisfy conditions (i)-(iii) in Lemma 2.1, the projective approximation lemma. The following theorem is a step in this direction.

THEOREM 2.7. *Let $(V, \mathcal{A}, \mu)$ be a finite measure space with a finite-rank approximation $(\mathcal{P}_n)_{n \in \mathbb{N}}$. For all integers $k \geq 2$ and $r \geq 0$,*

$$C_r(V, k)^* = \overline{\text{cone}} \bigcup_{n \in \mathbb{N}} B_n^* C_r(\mathcal{P}_n^+, k)^*.$$

The proof of this comes down to showing a degree of compatibility between $\mathcal{T}_r$ and B_n . In the process of proving this, we prove Theorem 2.9. This theorem extends a result by Bomze and De Klerk, who described an explicit generating set for the Pólya-type cones $(C_r(V))^*$ on $\text{Sym}(V)$ for finite V [19, Theorem 2.4].

For a finite set S and $r \in \mathbb{N}$, let $I^S(r)$ denote the set of all $m \in \mathbb{N}^S$ that have $\|m\|_1 = r$. The vectors m correspond bijectively to multisets of cardinality r of elements of S , thus we will say that m represents a choice $s_1, \dots, s_r \in S$. Bomze and De Klerk [19, Theorem 2.4] showed that, for all $r \geq 0$ and finite S ,

$$C_r(S) = \{ M \in \text{Sym}(S) : \langle mm^\top - \text{Diag } m, M \rangle \geq 0 \text{ for all } m \in I^S(r+2) \},$$

where $\text{Diag } m$ is the diagonal matrix with diagonal m .

If m corresponds to a choice $s_1, \dots, s_{r+2} \in S$, then $r!(mm^\top - \text{Diag } m)_{s_i s_j}$ is the multiplicity of $e_{s_i} \otimes e_{s_j}$ in $\mathcal{T}_r^*(e_{s_1} \otimes \dots \otimes e_{s_{r+2}})$, where e_{s_i} is the standard basis vector in $\mathbb{R}^S$ corresponding to s_i . Investigating expression (10) shows that $r!(mm^\top - \text{Diag } m)_{s_i s_j}$ is the number of permutations $\pi \in \mathfrak{S}_{r+2}$ such that $(s_i, s_j) = (s_{\pi 1}, s_{\pi 2})$.

We define a tensor analogue of the matrices $mm^\top - \text{Diag}(m)$. Let S be a finite set, $k \geq 2$ and $r \geq 0$ integers, and $m \in I^S(k+r)$. If m represents $t_1, \dots, t_{k+r}$, we define T_m as the k -tensor such that $r!(T_m)_{s_1, \dots, s_k}$ is the number of permutations $\pi \in \mathfrak{S}_{k+r}$ for which $(s_1, \dots, s_k) = (t_{\pi 1}, \dots, t_{\pi k})$. Although we do not need the exact values of T_m , for completeness's sake, if $m' \in I^S(k)$ represents $s_1, \dots, s_k \in S$, then

$$(T_m)_{s_1, \dots, s_k} = \begin{cases} \prod_{s \in \{s_1, \dots, s_k\}} \frac{m_s!}{(m_s - m'_s)!} & \text{if } m'_{s_i} \leq m_{s_i} \text{ for } 1 \leq i \leq k, \\ 0 & \text{otherwise.} \end{cases}$$

The notation for $\mathcal{T}_r$ and $A_{\mathcal{P}}$ does not specify the domains, meaning that the operators on $\text{Sym}(\mathcal{P}^+, k)$ get the same symbol as the corresponding operators on $L^2_{\text{sym}}(V, k)$. In Lemma 2.8 below, which ones are which should be derived from context.

LEMMA 2.8. *Let $(V, \mathcal{A}, \mu)$ be a finite measure space with finite partition $\mathcal{P}$, and let $k \geq 2$ and $r \geq 0$ be integers, then*

- (i) $\mathcal{T}_r A_{\mathcal{P}}^* = A_{\mathcal{P}}^* \mathcal{T}_r$, and
- (ii) if $P_1, \dots, P_{k+r} \in \mathcal{P}^+$ and $m \in I^{\mathcal{P}^+}(k+r)$ represents this choice, then

$$\mathcal{T}_r^* \frac{\mathbf{1}_{P_1, \dots, k+r}}{\mu^{k+r}(P_1, \dots, k+r)} = \frac{r!}{(k+r)!} B_{\mathcal{P}}^* T_m.$$

Proof. (i) Take $T \in (\mathbb{R}^{\mathcal{P}^+})^{k+r}$, $P_1, \dots, P_{k+r} \in \mathcal{P}^+$, and $w \in P_{1, \dots, k+r}$. Then

$$\begin{aligned} (A_{\mathcal{P}}^* \text{Av}_{\mathfrak{S}_{k+r}} T)(w) &= \sum_{Q_1, \dots, Q_{k+r} \in \mathcal{P}^+} \frac{1}{(k+r)!} \sum_{\pi \in \mathfrak{S}_{k+r}} T_{Q_{\pi 1}, \dots, Q_{\pi(k+r)}} \mathbf{1}_{Q_1, \dots, k+r}(w) \\ &= \frac{1}{(k+r)!} \sum_{\pi \in \mathfrak{S}_{k+r}} T_{P_{\pi 1}, \dots, P_{\pi(k+r)}} \\ &= \frac{1}{(k+r)!} \sum_{\pi \in \mathfrak{S}_{k+r}} \sum_{Q_1, \dots, Q_{k+r} \in \mathcal{P}^+} T_{Q_1, \dots, Q_{k+r}} \mathbf{1}_{Q_1, \dots, k+r}(\pi w) \\ &= (\text{Av}_{\mathfrak{S}_{k+r}} A_{\mathcal{P}}^* T)(w). \end{aligned}$$

Thus, for $T \in \text{Sym}(\mathcal{P}^+, k)$, $A_{\mathcal{P}}^* \text{Av}_{\mathfrak{S}_{k+r}}(T \otimes \mathbf{1}^{\otimes r}) = \text{Av}_{\mathfrak{S}_{k+r}} A_{\mathcal{P}}^*(T \otimes \mathbf{1}^{\otimes r})$. Moreover,

$$A_{\mathcal{P}}^*(T \otimes \mathbf{1}^{\otimes r}) = \sum_{P_1, \dots, P_k \in \mathcal{P}^+} T_{P_1, \dots, P_k} \mathbf{1}_{P_1, \dots, k} \otimes \left(\sum_{P \in \mathcal{P}^+} \mathbf{1}_P \right)^{\otimes r} = (A_{\mathcal{P}}^* T) \otimes \mathbf{1}^{\otimes r},$$

and the conclusion follows.

(ii) Let $P_1, \dots, P_{k+r} \in \mathcal{P}^+$, and let $m \in I^{\mathcal{P}^+}(k+r)$ be the vector representing this choice. For all $w \in V^k$

$$\begin{aligned} \mathcal{T}_r^* \frac{\mathbf{1}_{P_1, \dots, k+r}}{\mu^{k+r}(P_1, \dots, k+r)}(w) &= \frac{1}{(k+r)!} \int_{V^r} \sum_{\pi \in \mathfrak{S}_{k+r}} \frac{\mathbf{1}_{P_{\pi 1}, \dots, \pi(k+r)}}{\mu^{k+r}(P_1, \dots, k+r)}(w, v) d\mu(v) \\ &= \frac{1}{(k+r)!} \sum_{\pi \in \mathfrak{S}_{k+r}} \frac{\mathbf{1}_{P_{\pi 1}, \dots, \pi k}}{\mu^k(P_{\pi 1}, \dots, \pi k)}(w). \end{aligned}$$

By definition of T_m , for given $\pi \in \mathfrak{S}_{k+r}$, the number of $\sigma \in \mathfrak{S}_{k+r}$ such that $(P_{\sigma 1}, \dots, P_{\sigma k}) = (P_{\pi 1}, \dots, P_{\pi k})$ is $r!(T_m)_{P_{\pi 1}, \dots, P_{\pi k}}$. Group equal terms together to obtain

$$\begin{aligned} \mathcal{T}_r^* \frac{\mathbf{1}_{P_1, \dots, k+r}}{\mu^{k+r}(P_1, \dots, k+r)} &= \frac{r!}{(k+r)!} \sum_{Q_1, \dots, Q_k \in \mathcal{P}^+} (T_m)_{Q_1, \dots, Q_k} \frac{\mathbf{1}_{Q_1, \dots, k}}{\mu^k(Q_1, \dots, k)} \\ &= \frac{r!}{(k+r)!} B_{\mathcal{P}}^* T_m. \end{aligned} \quad \square$$

The next theorem is interesting, even if V is a finite set with the counting measure. In this case, taking $A_n = B_n = I$, it says that $C_r(V, k)^*$ is the conic hull of the tensors T_m with $m \in I^V(k+r)$. This is the k -tensor analogue of [19, Theorem 2.4].

THEOREM 2.9. *Let $(V, \mathcal{A}, \mu)$ be a finite measure space with a finite-rank approximation $(\mathcal{P}_n)_{n \in \mathbb{N}}$. For all integers $k \geq 2$ and $r \geq 0$,*

$$C_r(V, k) = \bigcap_{n \in \mathbb{N}} \{ K \in L_{\text{sym}}^2(V, k) : \langle K, B_n^* T_m \rangle \geq 0 \text{ for all } m \in I^{\mathcal{P}_n^+}(k+r) \}$$

and

$$C_r(V, k)^* = \overline{\text{cone}} \bigcup_{n \in \mathbb{N}} B_n^* \{ T_m : m \in I^{\mathcal{P}_n^+}(k+r) \}.$$

Proof. If $K \in C_r(V, k)$, then for all sets $P_1, \dots, P_{k+r} \in \mathcal{P}_n^+$ the inequality $\langle \mathcal{T}_r K, \mathbf{1}_{P_1, \dots, k+r} / \mu^k(P_{1, \dots, k+r}) \rangle \geq 0$ holds. So, using Lemma 2.8(ii), this implies $\langle K, B_n^* T_m \rangle \geq 0$ for all $m \in I^{\mathcal{P}_n^+}(k+r)$.

The martingale convergence theorem, Theorem 2.4, says that if K is in $L_{\text{sym}}^2(V, k)$, then

$$\mathcal{T}_r K = \lim_n \sum_{P_1, \dots, P_{k+r} \in \mathcal{P}_n^+} \frac{\langle \mathcal{T}_r K, \mathbf{1}_{P_1, \dots, k+r} \rangle}{\mu^k(P_{1, \dots, k+r})} \mathbf{1}_{P_1, \dots, k+r}$$

under the norm. Lemma 2.8(ii) shows that the condition $\langle K, B_n^* T_m \rangle \geq 0$ for all $m \in I^{\mathcal{P}_n^+}(k+r)$ is also sufficient for K to be in $C_r(V)$, proving the first statement of the theorem.

The equality

$$C_r(V, k)^* = \overline{\text{cone}} \bigcup_{n \in \mathbb{N}} B_n^* \{ T_m : m \in I^{\mathcal{P}_n^+}(k+r) \}$$

follows by taking the dual on both sides, and Theorem A.2(iv). $\square$

Proof of Theorem 2.7. This follows directly from Theorem 2.9. Indeed,

$$C_r(V, k)^* = \overline{\text{cone}} \bigcup_{n \in \mathbb{N}} B_n^* \{ T_m : m \in I^{\mathcal{P}_n^+}(k+r) \},$$

and it follows from the same theorem that

$$C_r(\mathcal{P}_n^+, k) = \overline{\text{cone}} \{ T_m : m \in I^{\mathcal{P}_n^+}(k+r) \}$$

by taking a trivial finite-rank approximation. $\square$

THEOREM 2.10. *If $(V, \mathcal{A}, \mu)$ is a countably generated finite measure space and $k \geq 2$ is an integer, then*

$$C_0(V, k)^* \supseteq C_1(V, k)^* \supseteq \dots \supseteq \text{CP}(V, k) \quad \text{and} \quad \text{CP}(V, k) = \bigcap_{r \in \mathbb{N}} C_r(V, k)^*.$$

Proof. The inclusions $C_r(V, k)^* \supseteq C_{r+1}(V, k)^*$ follow from Theorem 1.2 and Theorem 2.7.

Martingale convergence—Theorem 2.4—together with Lemma 2.5(i) says that $\lim_n B_n^* A_n = I$ under the weak operator topology. It suffices to check conditions (i)–(iii) of the projective approximation lemma for the cones $C_r(V, k)^*$ and $\text{CP}(V, k)$ and the operators A_n and B_n .

To this end, first prove that

$$A_n \text{CP}(V, k) = \text{CP}(\mathcal{P}_n^+, k) \quad \text{and} \quad A_n C_r(V, k)^* = C_r(\mathcal{P}_n^+, k)^*.$$

By Lemma 2.5(ii), for all n , $A_n B_n^* = I$ on $\text{Sym}(\mathcal{P}_n^+, k)$. Thus, it is enough to show that A_n maps the respective cones on V into their counterparts on $\mathcal{P}_n^+$, and that B_n^* does the opposite.

The maps A_n and B_n^* preserve pointwise nonnegativity. Hence, the equalities $A_n \text{CP}(V, k) = \text{CP}(\mathcal{P}_n^+, k)$ for all n follow from Lemma 2.5(iii): indeed, if $f \geq 0$ and $w \geq 0$, then $A_n f \geq 0$ and $B_n^* w \geq 0$, and $A_n(f^{\otimes k}) = (A_n f)^{\otimes k}$ and $B_n^*(w^{\otimes k}) = (B_n^* w)^{\otimes k}$.

Theorem 2.7 says that $B_n^* C_r(\mathcal{P}_n^+, k)^* \subseteq C_r(V, k)^*$. To prove

$$A_n C_r(V, k)^* \subseteq C_r(\mathcal{P}_n^+, k)^*,$$

by Theorem A.2(i) and (ii), it is equivalent to show that $C_r(\mathcal{P}_n^+, k)$ is contained in $A_n^{*-1} C_r(V, k)$. Lemma 2.8(i) says that if $T \in \text{Sym}(\mathcal{P}_n^+, k)$, then it follows that $\mathcal{T}_r A_n^* T = A_n^* \mathcal{T}_r T$. The latter is nonnegative if $T \in C_r(\mathcal{P}_n^+, k)$, whence it follows that $T \in A_n^{*-1} C_r(V, k)$. This proves $A_n C_r(V, k)^* = C_r(\mathcal{P}_n^+, k)^*$.

Next, prove conditions (i)–(iii) of the projective approximation lemma.

(i) The inclusion $B_n^* A_n C_r(V, k)^* \subseteq C_r(V, k)^*$ for all $n, r \in \mathbb{N}$ follows from $A_n C_r(V, k)^* = C_r(\mathcal{P}_n^+, k)^*$ and Theorem 2.7.

(ii) To prove $B_n^* A_n \text{CP}(V, k) \subseteq \text{CP}(V, k)$, it is enough to consider the generators $f^{\otimes k}$ with $f \in L^2(V)_{\geq 0}$. Lemma 2.5(iii) shows that the equality $B_n^* A_n(f^{\otimes k}) = (B_n^* A_n f)^{\otimes k}$ holds, and $B_n^* A_n f$ is nonnegative, so property (ii) follows.

(iii) The property $A_n \text{CP}(V, k) = \bigcap_r A_n C_r(V, k)^*$ follows from the fact that $A_n \text{CP}(V, k) = \text{CP}(\mathcal{P}_n^+, k)$, $A_n C_r(V, k) = C_r(\mathcal{P}_n^+, k)$, and Theorem 1.2. This concludes the proof. $\square$

4. An outer approximation on σ -finite measure spaces

We would like to extend Theorem 2.10 to σ -finite measure spaces by restricting to finite-measure subspaces. Indeed, we can recognize a completely positive kernel on a σ -finite space by showing it is completely positive on every finite-measure subspace. However, it is a well-known frustration that such a

procedure is ineffective for the Pólya-type cones, even in finite dimensions; this is related to the obstruction [68, Theorem 3].

Example 2.11. Fix an integer $n \geq 1$, and define $\text{Res} : \text{Sym}(n) \rightarrow \text{Sym}(n-1)$ as the restriction to the principal submatrix indexed by $[n-1]$. Its adjoint lifts a matrix by appending a row and a column of zeros.

Let $r \in \mathbb{N}$ and $A \in \text{Sym}(n-1)$, and assume there exists $A_{i,j} < 0$. Then,

$$(\mathcal{T}_r \text{Res}^* A)_{i,j,n,\dots,n} = (2r!/(r+2)!)A_{i,j} < 0.$$

So, the lifting $\text{Res}^* A$ is in $C_r([n])$ if and only if $A \geq 0$.

A similar problem occurs on the dual side. Take $A \in (C_r([n]))^*$ and B in $\text{Sym}(n-1)$. Then $\langle \text{Res} A, B \rangle = \langle A, \text{Res}^* B \rangle$, so, by the above and Theorem A.2, $\langle \text{Res} A, B \rangle \geq 0$ if and only if $B \geq 0$. It follows that $\text{Res} A \in (C_r([n-1]))^*$ if and only if $r = 0$.

A greater plight renders the objection raised by Example 2.11 irrelevant. Applied to an infinite measure space $(V, \mathcal{A}, \mu)$, except in degenerate cases, the Pólya-type cones fail to approximate anything more than the nonnegative orthant. If $\mathcal{P}$ is a partition of V into sets of finite measure and $|\mathcal{P}^+| = |\mathbb{N}|$, the codomain of the operator $A_{\mathcal{P}}$ on $L^2(V^2)$ is isomorphic to $l^2(\mathbb{N}^2)$. Example 2.12 shows we cannot simply copy-paste the definition of the Pólya-type cones to this setting.

Example 2.12. Consider $l^2_{\text{sym}}(\mathbb{N}^2)$ with the counting measure, and fix $r \in \mathbb{N}$. Define the operator $\mathcal{T}_r : l^2_{\text{sym}}(\mathbb{N}^2) \rightarrow l^\infty_{\text{sym}}(\mathbb{N}^{r+2})$ by

$$(\mathcal{T}_r a)_{i_1, \dots, i_{r+2}} = (1/(r+2)!) \sum_{\pi \in \mathfrak{S}_{r+2}} a_{i_{\pi 1}, i_{\pi 2}}.$$

Take a such that $\mathcal{T}_r a \geq 0$. Since a is square summable, for all $\epsilon > 0$ and all $j \in \mathbb{N}$ there is an n_j^ϵ such that if $i \geq n_j^\epsilon$, then $a_{i,j} \leq \epsilon$. Take $i, j \in \mathbb{N}$, $\epsilon > 0$ and $\epsilon' = 2\epsilon/(r(r+3))$. Choose $k_1, \dots, k_r \in \mathbb{N}$ such that

$$\begin{aligned} k_1 &\geq \max\{n_i^{\epsilon'}, n_j^{\epsilon'}\}, \\ k_2 &\geq \max\{n_i^{\epsilon'}, n_j^{\epsilon'}, n_{k_1}^{\epsilon'}\}, \\ &\vdots \\ k_r &\geq \max\{n_i^{\epsilon'}, n_j^{\epsilon'}, n_{k_1}^{\epsilon'}, \dots, n_{k_{r-1}}^{\epsilon'}\}. \end{aligned}$$

Then,

$$\begin{aligned}
0 &\leq ((r+2)!/(2r!))(\mathcal{T}_r a)_{i,j,k_1,\dots,k_r} \\
&= a_{i,j} + \sum_{l \in [r]} \left(a_{i,k_l} + a_{j,k_l} + \sum_{l < l' \leq r} a_{k_l,k_{l'}} \right) \\
&\leq a_{i,j} + (r(r+3)/2)\epsilon' = a_{i,j} + \epsilon,
\end{aligned}$$

so a itself is nonnegative. Thus, for all $r \geq 0$, $\mathcal{T}_r^{-1}l_{\text{sym}}^2(\mathbb{N}^{r+2})_{\geq 0} = l_{\text{sym}}^2(\mathbb{N}^2)_{\geq 0}$.

To overcome these difficulties, we define a slightly different set of cones, which offers an outer approximation of the completely positive cone. Compared with the approximation by Pólya-type cones on a finite measure space, the approximation defined in this section is strictly stronger.

Let $(V, \mathcal{A}, \mu)$ be a countably generated σ -finite measure space, and fix an integer $k \geq 2$. Denote the set of measurable subsets of nonzero and finite measure by $\mathcal{A}_{\text{fin}}$. Write for all $A \in \mathcal{A}_{\text{fin}}$ the restriction to A as

$$\text{Res}_A : L_{\text{sym}}^2(V, k) \rightarrow L_{\text{sym}}^2(A, k).$$

For $T \in L_{\text{sym}}^2(A, k)$, the tensor $\text{Res}_A^* T$ agrees with T on A and is 0 everywhere else.

For $r \in \mathbb{N}$, define the cone

$$\check{C}_r(V, k) = \overline{\text{cone}} \bigcup_{A \in \mathcal{A}_{\text{fin}}} \text{Res}_A^* C_r(A, k);$$

in words: $\check{C}_r(V, k)$ is the cone generated by tensors $T \in L_{\text{sym}}^2(V, k)$ for which there is an $A \in \mathcal{A}_{\text{fin}}$ and a $K \in C_r(A, k)$ such that $T(v) = K(v)$ for all $v \in A^k$, and $T(v) = 0$ everywhere else. Theorem A.2 shows

$$\check{C}_r(V, k)^* = \bigcap_{A \in \mathcal{A}_{\text{fin}}} \text{Res}_A^{-1} C_r(A, k)^*; \quad (11)$$

that is, $T \in \check{C}_r(V, k)^*$ if and only if $\text{Res}_A(T) \in C_r(A, k)^*$ for all $A \in \mathcal{A}_{\text{fin}}$. Since $\check{C}_r(V, k)$ is contained in $C_r(V, k)$ if $\mu(V) < \infty$, if $(V, \mathcal{A}, \mu)$ satisfies the conditions of Theorem 2.10, then $\bigcap_r \check{C}_r(V, k)^* = \text{CP}(V, k)$. So, on countably generated finite measure spaces the $\check{C}_r(V, k)$ give a tighter outer approximation of $\text{CP}(V, k)$ than the Pólya-type cones.

THEOREM 2.13. *If $(V, \mathcal{A}, \mu)$ is a countably generated σ -finite measure space, then, for every integer $k \geq 2$,*

$$\check{C}_0(V, k)^* \supseteq \check{C}_1(V, k)^* \supseteq \dots \supseteq \text{CP}(V, k) \quad \text{and} \quad \text{CP}(V, k) = \bigcap_{r \in \mathbb{N}} \check{C}_r(V, k)^*.$$

Proof. The inclusions $\check{C}_r(V, k)^* \supseteq \check{C}_{r+1}(V, k)^*$ for all $r \in \mathbb{N}$ follow from the corresponding inclusions $C_r(A, k)^* \supseteq C_{r+1}(A, k)^*$ for all $A \in \mathcal{A}_{\text{fin}}$.

By Lemma 2.6 and Theorem 2.10,

$$\bigcap_{r \in \mathbb{N}} \check{C}_r(V, k)^* = \bigcap_{A \in \mathcal{A}_{\text{fin}}} \text{Res}_A^{-1} \bigcap_{r \in \mathbb{N}} C_r(A, k)^* = \bigcap_{A \in \mathcal{A}_{\text{fin}}} \text{Res}_A^{-1} \text{CP}(A, k),$$

so, if $\text{CP}(V, k) = \bigcap_A \text{Res}_A^{-1} \text{CP}(A, k)$, the theorem follows.

If $\phi \in L^2(V^k)_{\geq 0}$, then for all $A \in \mathcal{A}_{\text{fin}}$, $\text{Res}_A(\phi^{\otimes k}) = (\text{Res}_A \phi)^{\otimes k}$. Since for all A , $\text{Res}_A \phi \in L^2(A)_{\geq 0}$, this shows that $\text{CP}(V, k) \subseteq \bigcap_A \text{Res}_A^{-1} \text{CP}(A, k)$.

On the other hand, $\text{Res}_A^*(\phi^{\otimes k}) = (\text{Res}_A^* \phi)^{\otimes k}$ for all $\phi \in L^2(A)_{\geq 0}$, so that $\text{Res}_A^* \text{CP}(A, k) \subseteq \text{CP}(V, k)$. Make $\mathcal{A}_{\text{fin}}$ into a directed set by equipping it with the opposite of the inclusion relation. Then, if $T \in \text{COP}(V, k)$ and if $K \in \bigcap_A \text{Res}_A^{-1} \text{CP}(A, k)$, then,

$$\langle K, T \rangle = \lim_{A \in \mathcal{A}_{\text{fin}}} \langle \text{Res}_A K, \text{Res}_A T \rangle = \lim_{A \in \mathcal{A}_{\text{fin}}} \langle \text{Res}_A^* \text{Res}_A K, T \rangle \geq 0.$$

Therefore, $\bigcap_A \text{Res}_A^{-1} \text{CP}(A, k) \subseteq \text{CP}(V, k)$, and thus these cones are equal, from which the theorem follows. $\square$

CHAPTER 3

Complete positivity and symmetry

Problems I–IV in the introduction have a lot of symmetry: for example, take $n \geq 1$ and integer and let $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ be the unit sphere in $\mathbb{R}^n$. If we take a set $S \subseteq S^{n-1}$ that contains no orthogonal pairs, then neither does TS for any $T \in O(n)$, the orthogonal group on $\mathbb{R}^n$. Moreover, this group action leaves the measure unchanged.

We expect therefore that the optimal value of such an optimization problem remains unchanged when we restrict the feasible region to feasible solutions that are invariant under a group action. Restricting the feasible region to invariant solutions decreases the size of such a problem significantly, and is essential for tractability. The goal of this section is to describe an outer approximation of the cone of group-invariant completely positive tensors.

This hinges on the existence of a linear averaging operator, one example of which we have already seen: the operator $\text{Av}_{\mathfrak{S}_{k+r}}$. Such operators deserve careful consideration. We start this chapter with some harmonic analysis and an in-depth investigation of these averaging operators.

1. Harmonic analysis

We largely follow Folland’s book on harmonic analysis [45]. See Section 5 of the appendix for more details on invariant measures.

When we denote a group multiplicatively, we call its unit 1. When we denote it additively, we call its unit 0. All groups we study are locally compact and Hausdorff. To avoid technicalities, we also assume they are σ -compact: a σ -compact group is a countable union of compact sets. Recall that any locally compact group has a left-invariant measure, the Haar measure, and is unimodular if and only if the Haar measure is also right-invariant. If Γ is a compact group, the Haar measure $\mu(\Gamma)$ is finite, and we normalize $\mu(\Gamma) = 1$, unless stated otherwise.

Homogeneous spaces. Let Γ be a σ -compact unimodular locally compact group. An *action* of Γ on a topological space V is a continuous function $\Gamma \times V \rightarrow V$, $(\gamma, v) \mapsto \gamma v$, such that $v \mapsto \gamma v$ is a homeomorphism of V for all $\gamma \in \Gamma$, and $\gamma(\zeta v) = (\gamma\zeta)v$ for all $\gamma, \zeta \in \Gamma$ and $v \in V$. A space with an action of Γ is called a Γ -space. For V a Γ -space and $v \in V$, call $\text{Orb}(v) = \{\gamma v : \gamma \in \Gamma\}$ the *orbit* of v and $\text{Stab}(v) = \{\gamma \in \Gamma : \gamma v = v\}$ the *stabilizer* of v . A Γ -space V is called *homogeneous* if for all $v, w \in V$ there exists $\gamma \in \Gamma$ such that $\gamma v = w$. In other words, V is a homogeneous space if and only if V has only one orbit.

A function $f : V \rightarrow W$ with V a Γ -space is Γ -invariant, or just *invariant*, if for all $\gamma \in \Gamma$, $f(\gamma v) = f(v)$. If W is also a Γ -space, the function f is called Γ -equivariant if, for all $\gamma \in \Gamma$ and $v \in V$, $f(\gamma v) = \gamma f(v)$.

Let V be a homogeneous Γ -space. A $v_0 \in V$ defines a Γ -equivariant map $\phi : \Gamma \rightarrow V$ through $\gamma \mapsto \gamma v_0$, and a quotient map $p : \Gamma \rightarrow \Gamma/\text{Stab}(v_0)$ through $\gamma \mapsto \gamma \text{Stab}(v_0)$. The stabilizer of v_0 is a closed subgroup of Γ . Then, ϕ induces a Γ -invariant homeomorphism $\Phi : \Gamma/\text{Stab}(v_0) \rightarrow V$ such that $\Phi \circ p = \phi$; this depends on Γ being σ -compact. Thus, for us, a homogeneous space is always the quotient of Γ by a closed subgroup H .

Let V be identified with Γ/H , with H closed and unimodular. Given a Haar measure μ on Γ and a Haar measure ν on H , Section 5 of the appendix describes a Γ -invariant measure ω on Γ/H . If H is compact, we can assume it is equal to the pushforward of μ under p , and we call it the *quotient measure* of μ under p . If H is not compact, we only have the equation

$$\int_{\Gamma} f(\gamma) d\mu(\gamma) = \int_{\Gamma/H} \int_H f(v\zeta) d\nu(\zeta) d\omega(p(v)), \quad (12)$$

for all $f \in L^1(\Gamma)$.

If H is compact and Γ/H is equipped with the quotient measure, then, for all $1 \leq p \leq \infty$ and integer $k \geq 1$, let $\gamma f(v_1, \dots, v_k) = f(\gamma^{-1}v_1, \dots, \gamma^{-1}v_k)$. This defines an action $(\gamma, f) \mapsto \gamma f$ on $L^p((\Gamma/H)^k)$, called the *diagonal action* of Γ on $L^2((\Gamma/H)^k)$. Indeed, for all $1 \leq p \leq \infty$, all $f \in L^p((\Gamma/H)^k)$, and all $\gamma \in \Gamma$, we have $\|f\|_p = \|\gamma f\|_p$; if $p = \infty$ this follows from homogeneity of Γ/H , and if $p < \infty$ this follows from invariance of the measure. So, for all $\gamma \in \Gamma$ the map $f \mapsto \gamma f$ is isometric, and in particular continuous with continuous inverse $f \mapsto \gamma^{-1}f$. Associativity, $(\zeta\gamma)f = \zeta(\gamma f)$, follows from a direct calculation.

Finally, the quotient map $p : \Gamma \rightarrow \Gamma/H$ is open. Indeed, let $U \subseteq \Gamma$ be open, then $p^{-1}(p(U)) = UH = \bigcup_{\zeta \in H} U\zeta$. Since U is open, each $U\zeta$ is open, and their union is as well, so $p(U)$ is open in Γ/H . Since open and continuous surjections send compact sets to compact sets, bases to bases and local bases to local bases, this implies that if Γ is locally compact, Γ/H is

also locally compact, if Γ is σ -compact, Γ/H is σ -compact, and if Γ is second countable, Γ/H is second countable.

Cross-correlations and functions of positive type. Let Γ be a σ -compact unimodular locally compact group with Haar measure μ , and let V be a homogeneous Γ -space equipped with the quotient map $p : \Gamma \rightarrow V$ such that $p(1) = v_0$. Let f and g be two measurable functions on V . Their *convolution* $f * g$ is given by

$$f * g(v) = \int_{\Gamma} f(\gamma v_0) g(\gamma^{-1}v) d\mu(\gamma).$$

Although it is standard to work with the convolution of f and g , in this exposition the *cross-correlation* is more natural. It is given by

$$f \star g(v) = \int_{\Gamma} f(\gamma^{-1}v_0) g(\gamma^{-1}v) d\mu(\gamma).$$

By invariance of the measure—i.e. by unimodularity of the group—if clarity demands so, we may drop the $^{-1}$ in the formula. Call the cross-correlation $f \star f$ of a function f with itself the *auto-correlation* of f .

For the convolution, we have the following inequalities.

- If $1 \leq p \leq \infty$, $f \in L^1(V)$, and $g \in L^p(V)$, then $\|f * g\|_p \leq \|f\|_1 \|g\|_p$, hence $f * g, g * f \in L^p(V)$.
- If $1 \leq p, q \leq \infty$ such that $1/p + 1/q = 1$, $f \in L^p(V)$, and $g \in L^q(V)$, then $f * g \in C_0(V)$ and $\|f * g\|_{\infty} \leq \|f\|_p \|g\|_q$. We will call this *Young's inequality*.

The convolution and the cross-correlation are related; let f and g be functions $V \rightarrow \mathbb{R}$ such that $f * g$ and $f \star g$ are well-defined. Denoting $\bar{f}(\gamma) = f(\gamma^{-1})$, it follows immediately that $f \star g = \bar{f} * g$. This means that the inequalities above also hold for the correlation.

Given a σ -compact unimodular locally compact group Γ with Haar measure μ , an *approximate identity* is a net $(\psi_i)_{i \in I}$ of compactly supported and bounded functions such that

- there exists a neighborhood basis $\mathcal{U}$ of 1 such that for every $U \in \mathcal{U}$ there is an i_0 such that for all $i \geq i_0$, $\text{supp } \psi_i \subseteq U$,
- $\psi_i \geq 0$ for all i , and
- $\int_{\Gamma} \psi_i(\gamma) d\mu(\gamma) = 1$.

We moreover always assume that $\psi_i(\gamma^{-1}) = \psi_i(\gamma)$ for all $i \in I$ and $\gamma \in \Gamma$. Then, for all $f \in L^p(\Gamma)$ with $1 \leq p < \infty$:

$$\lim_{i \in I} \|\psi_i * f - f\|_p = \lim_{i \in I} \|f * \psi_i - f\|_p = 0,$$

and likewise with the convolution replaced by the correlation. The same results hold if $p = \infty$ and f is uniformly continuous.

A *function of positive type* on Γ is a function $\phi \in L^\infty(\Gamma)$ such that for all $f \in L^1(\Gamma)$:

$$\int_{\Gamma} (f \star f)(\gamma) \phi(\gamma) d\mu(\gamma) \geq 0.$$

The set of positive-type functions is a closed convex cone in $L^\infty(\mathbb{R}^n)$, which we denote by $P(\Gamma)$. A function of positive type is continuous almost everywhere [45, Corollary 3.21], so we always assume them to be continuous. Moreover, if $f \in L^2(\Gamma)$, then $f \star f \in P(\Gamma)$ [45, Corollary 3.16].

2. Averaging

Fix a σ -compact unimodular locally compact group Γ with Haar measure μ . When Γ is compact, normalize $\mu(\Gamma) = 1$. Let H be a compact subgroup of Γ , $p : \Gamma \rightarrow \Gamma/H$ be the quotient map, $V = \Gamma/H$, and equip V with the quotient measure $\nu = p_*\mu$.

We will study two different operators that send a k -tensor on V to an invariant counterpart. The first exists for compact Γ and integer $k \geq 2$, and is an orthogonal projection of the Hilbert space $L^2(V^k)$. The second exists only for trace-class kernel, but is defined for any σ -compact unimodular locally compact Γ . We call these operators *averaging operators*, and denote them by Av_Γ . In the literature, the term *Reynolds operator* is also common.

To define such operators, for $T \in L^2(V^k)$, the integral formula

$$\text{Av}_\Gamma T(v) = \int_{\Gamma} T(g^{-1}v) d\mu(g) \tag{13}$$

is tempting, but it is a priori unclear whether this is well-defined. Even if Γ is compact, complications occur: the orbits of V^k under the diagonal action can have measure 0.

Example 3.1. Let H be a compact subgroup of Γ and $\mu(H) = 0$. Then, the orbit $\text{Orb}(H, H) = \{(gH, gH) : g \in \Gamma\} \subseteq (\Gamma/H)^2$ also has measure 0. Indeed,

$$\nu^2(\text{Orb}(H, H)) = \mu^2(\{(g, g') \in \Gamma^2 : (p(g), p(g')) \in \text{Orb}(H, H)\})$$

For $g \in \Gamma$, let $\xi^g(g') = 1$ if $(p(g), p(g')) \in \text{Orb}(H, H)$ and $\xi^g(g') = 0$ otherwise. Then, $\xi^g(g') = 1$ if and only if there are $h_1, h_2 \in H$ such that $g' = gh_1^{-1}h_2$, so its support is gH , which has measure 0. By Tonelli's theorem,

$$\mu^2(\{(g, g') \in \Gamma^2 : (p(g), p(g')) \in \text{Orb}(H, H)\}) = \int_{\Gamma} \mu(\xi^g) d\mu(g) = 0.$$

Let $k \geq 2$ be an integer, and denote the space of k -tensors on V that are invariant under the diagonal action by $L^2_{\text{sym}}(V, k)^\Gamma$. The map $T \mapsto gT$ is an isometry, so if T' is invariant and $\|T' - T\|_2 \leq \epsilon$, then

$$\|T' - gT\|_2 = \|gT' - gT\|_2 = \|T' - T\|_2 \leq \epsilon.$$

Thus, the subspace of invariant tensors is closed, so it comes with an orthogonal projection.

If Γ is compact, the continuous tensors lie dense in $L^2(V^k)$, and on $C(V^k)$ the integral formula (13) defines a bounded linear operator. The continuous extension to $L^2(V^k)$ of this map yields the first example of an averaging operator; it is the orthogonal projection onto $L^2(V^k)^\Gamma$.

THEOREM 3.2. *Let Γ be a compact group, H be a closed subgroup, and Γ/H be equipped with the quotient measure. The continuous extension of the operator on $C(V^k)$ defined by (13) to an operator on $L^2((\Gamma/H)^k)$ is the orthogonal projection $\text{Av}_\Gamma : L^2((\Gamma/H)^k) \rightarrow L^2((\Gamma/H)^k)^\Gamma$.*

Proof. Let $V = \Gamma/H$ and denote the quotient measure on V by ν . It is enough to show that the two operators coincide on $C(V^k)$. If $T \in C(V^k)$ and $K \in L^2(V^k)$, then, by Fubini-Tonelli,

$$\begin{aligned} \int_{V^k} \int_\Gamma T(g^{-1}v) d\mu(g) K(v) d\nu^k(v) &= \int_\Gamma \int_{V^k} T(g^{-1}v) K(v) d\nu^k(v) d\mu(g) \\ &\leq \|T\|_\infty \|K\|_2 < \infty. \end{aligned}$$

Therefore, by the Riesz representation theorem, $v \mapsto \int_\Gamma T(g^{-1}v) d\mu(g)$ is in $L^2(V^k)$. The map $C(V^k) \rightarrow L^2(V^k)$ thus defined is moreover bounded, because both measures are finite.

The image of this map lies in $L^2(V^k)^\Gamma$, and when T is already invariant, the integral only adds a factor $\mu(\Gamma) = 1$. So, the extension to $L^2(V^k)$ is indeed the identity on $L^2(V^k)^\Gamma$ and equal to the orthogonal projection onto $L^2(V^k)^\Gamma$. $\square$

If Γ is not compact, the integral formula applied to a continuous function on Γ is not necessarily square-integrable, even when the function has compact support. So, even though the orthogonal projection $L^2(V^k) \rightarrow L^2(V^k)^\Gamma$ exists, it is not defined by (13).

In general, we will be mainly interested in trace-class kernels. On a rank-one operator $\phi \otimes \psi \in \mathcal{B}^1(L^2(\Gamma))$, Equation (13) produces the correlation $\phi \star \psi$ by

$$\phi \star \psi(h) = \int_\Gamma (\phi \otimes \psi)(g^{-1}, g^{-1}h) d\mu(g).$$

Young's inequality tells us that it is bounded: $\|\phi \star \psi\|_\infty \leq \|\phi\|_2 \|\psi\|_2 < \infty$. The extension of this map to $\mathcal{B}^1(L^2(V))$ is the second averaging operator we

define. This operator is different from the projection $L^2(\Gamma^2) \rightarrow L^2(\Gamma^2)^\Gamma$: its range is contained in $C_0(\Gamma)$, the space of continuous functions that vanish outside compact sets, and the map $(g, h) \mapsto (\phi \star \psi)(g^{-1}h)$ is in general not in $L^2(\Gamma^2)$.

THEOREM 3.3. *Let Γ be a σ -compact unimodular locally compact group, H be a compact subgroup, and Γ/H be equipped with the quotient measure. The continuous linear extension of the map $\phi \otimes \psi \mapsto \phi \star \psi$ induces a bounded linear operator $\text{Av}_\Gamma : \mathcal{B}^1(L^2(\Gamma/H)) \rightarrow C_0(\Gamma/H)$.*

The proof depends on the following Lemma, which will be of use to us more often.

LEMMA 3.4. *Let Γ be a σ -compact unimodular locally compact group, H be a compact subgroup, and Γ/H be equipped with the quotient measure. For all $\phi_1, \dots, \phi_k \in L^k(\Gamma/H)$, the function*

$$(v_1, \dots, v_k) \mapsto \int_\Gamma \phi_1(g^{-1}v_1) \cdots \phi_k(g^{-1}v_k) d\nu(g)$$

is continuous and vanishes at infinity.

Proof. From the definition of the quotient measure it follows that $\phi \in L^k(\Gamma/H)$ if and only if $\phi \circ p \in L^k(\Gamma)$. The function

$$(g_1, \dots, g_k) \mapsto \int_\Gamma \phi_1(p(g^{-1}g_1)) \cdots \phi_k(p(g^{-1}g_k)) d\mu(g)$$

is constant on the cosets of H , so if it is in $C_0(\Gamma)$, it defines an element of $C_0(\Gamma/H)$. It is therefore sufficient to take $H = \{1\}$ by replacing ϕ_i by $\phi_i \circ p$ for all i . For the sake of readability, we will integrate over g instead of g^{-1} , as this is equivalent.

If $\phi_1, \dots, \phi_k \in C_c(\Gamma)$, each ϕ_i is left uniformly continuous by [45, Proposition 2.6], therefore $(v_1, \dots, v_k) \mapsto \int_\Gamma \phi_1(gv_1) \cdots \phi_k(gv_k) d\mu(g)$ is continuous.

If $\phi_1, \dots, \phi_k \in L^k(\Gamma)$, each ϕ_i is approximated in the norm topology by elements of $C_c(\Gamma)$. The strategy is to prove that a choice of $\epsilon > 0$ and $\psi_1, \dots, \psi_k \in C_c(\Gamma)$ for which $\|\phi_i - \psi_i\|_k \leq \epsilon$, for all i , give a bound on

$$\left| \int_\Gamma \phi_1(gv_1) \cdots \phi_k(gv_k) - \psi_1(gv_1) \cdots \psi_k(gv_k) d\mu(g) \right| \quad (14)$$

that goes to 0 as ϵ goes to 0 and is uniform in $v_1, \dots, v_k$. The conclusion then follows.

Apply the triangle inequality to see that (14) is at most

$$\begin{aligned}
& \int_{\Gamma} |\phi_1(gv_1) \cdots \phi_k(gv_k) - \phi_1(gv_1) \cdots \phi_{k-1}(gv_{k-1})\psi_k(gv_k) \\
& \quad + \phi_1(gv_1) \cdots \phi_{k-1}(gv_{k-1})\psi_k(gv_k) - \psi_1(gv_1) \cdots \psi_k(gv_k)| d\mu(g) \\
& \leq \int_{\Gamma} |\phi_1(gv_1) \cdots \phi_{k-1}(gv_{k-1})| |\phi_k(gv_k) - \psi_k(gv_k)| d\mu(g) \\
& \quad + \int_{\Gamma} |\phi_1(gv_1) \cdots \phi_{k-1}(gv_{k-1}) - \psi_1(gv_1) \cdots \psi_{k-1}(gv_{k-1})| |\psi_k(gv_k)| d\mu(g).
\end{aligned}$$

Repeatedly applying this method shows that (14) is bounded from above by a sum of k terms of the form

$$\int_{\Gamma} |\phi_i(gv_i) - \psi_i(gv_i)| |\xi_1(g)| \cdots |\xi_{k-1}(g)| d\mu(g), \quad (15)$$

where each ξ_j is one of the functions $g \mapsto \phi_l(gv_l)$ or $g \mapsto \psi_l(gv_l)$.

Recursive application of Hölder's inequality reveals that each term of the form (15) is at most

$$\|\phi_i - \psi_i\|_k \|\xi_1\|_k \cdots \|\xi_{k-1}\|_k \leq \epsilon M^{k-1},$$

where $M = \epsilon + \max\{\|f_1\|_k, \dots, \|f_k\|_k\}$. Since there are k such terms, (14) is bounded from above by $\epsilon k M^{k-1}$. This proves

$$(v_1, \dots, v_k) \mapsto \int_{\Gamma} \phi_1(g^{-1}v_1) \cdots \phi_k(g^{-1}v_k) d\mu(g)$$

is a uniform limit of continuous functions with compact support and hence lies in $C_0(\Gamma)$. $\square$

Proof of Theorem 3.3. As in the proof of Lemma 3.4, it is enough to prove the theorem for $H = \{1\}$.

Define Av_{Γ} on finite-rank operators by $\text{Av}_{\Gamma}(\sum_{i=1}^n \phi_i \otimes \psi_i) = \sum_{i=1}^n \phi_i \star \psi_i$. This is linear, since the correlation is linear in both its arguments. Young's inequality for convolutions shows it is separately continuous in the ϕ_i and ψ_i under the supremum norm on the codomain. Then, Av_{Γ} is independent of an expansion of the ϕ_i and ψ_i over a basis, and so well-defined, and Av_{Γ} extends linearly and continuously to $\mathcal{B}^1(L^2(\Gamma))$.

If $K \in \mathcal{B}^1(L^2(\Gamma))$ with singular value decomposition $K = \sum_{i \in \mathbb{N}} \lambda_i \phi_i \otimes \psi_i$, then

$$\|\text{Av}_{\Gamma} K\|_{\infty} \leq \sum_{i \in \mathbb{N}} |\lambda_i| \|\phi_i \star \psi_i\|_{\infty} \leq \sum_{i \in \mathbb{N}} |\lambda_i| \|\phi_i\|_2 \|\psi_i\|_2 = \sum_{i \in \mathbb{N}} |\lambda_i| < \infty.$$

Applying the same argument to the tails $\sum_{i \geq n} \lambda_i \phi_i \otimes \psi_i$ shows that the series $\sum_{i \in \mathbb{N}} \lambda_i \phi_i \star \psi_i$ converges absolutely and uniformly. Lemma 3.4 implies that $\text{Av}_\Gamma(K)$ is continuous and vanishes at infinity, which proves the theorem. $\square$

When Γ is compact, the two averaging operators we defined coincide on the trace class. Every continuous positive semidefinite kernel on a compact space is of trace class, which is a corollary of Mercer's theorem. So, in this case, the image of the trace class under Av_Γ contains all invariant continuous positive semidefinite kernels.

After doing all this work, it is perfectly fine to interpret (13) in the almost-everywhere sense. Indeed, under the conditions of Theorem 3.2 or Theorem 3.3, it is valid to use this formula in expressions of the form $\langle \text{Av}_\Gamma T, F \rangle$.

My main objection to the integral formula (13) was that $T \in L^p(V^k)$ might not be well-defined on orbits, as they might have measure 0. However, even without the above discussion, there is some sense to the integral formula when T is a trace-class kernel and Γ is second countable. Picking points $v, w \in V$ defines the kernel $K(g, h) = T(gv, hw)$ almost everywhere on Γ^2 . It is trace class, for example because

$$\|K\|_{\mathcal{B}^1} = \sup_{(\phi_n)_{n \in \mathbb{N}}, (\psi_n)_{n \in \mathbb{N}}} \|(\langle K \phi_n, \psi_n \rangle)_{n \in \mathbb{N}}\|_1 = \|T\|_{\mathcal{B}^1},$$

with the supremum over pairs of orthonormal bases of $L^2(V)$ [96, Proposition 2.6]. Brislawn showed [20]—using martingales—that there exists a function $\tilde{K}$ almost-everywhere equal to K for which $\text{Tr}(K) = \int_\Gamma \tilde{K}(g, g) d\mu(g)$, which is as close to (13) as we can wish for.

3. A group-invariant outer approximation

We have gathered all we need to discuss group-invariant analogues of the completely positive cone and its approximations. Our approach is by reduction to Theorems 2.10 and 2.13. The Borel algebra is countably generated if and only if the topology is second countable, and together with local compactness this implies the space is σ -compact, so second countable unimodular locally compact groups satisfy the conditions of Theorems 2.10 and 2.13 and the results in Section 2. Parallel to Chapter 2, we first discuss compact groups, then we extend to second countable unimodular locally compact groups.

Recall that if Γ is second countable and locally compact, then for every subgroup H , Γ/H is second countable and locally compact, see Section 1.

THEOREM 3.5. *If Γ is a second countable compact group, if H is a compact subgroup, and if $V = \Gamma/H$ is equipped with the quotient measure, then*

$$\text{Av}_\Gamma \text{CP}(V, k) = \bigcap_{r \in \mathbb{N}} \text{Av}_\Gamma C_r(V, k)^*,$$

and all cones in the expression are closed.

Proof. Recall Theorem 3.2: the operator $\text{Av}_\Gamma : L^2_{\text{sym}}(V, k) \rightarrow L^2_{\text{sym}}(V, k)^\Gamma$ is an orthogonal projection. If $\mathcal{C} \subseteq L^2_{\text{sym}}(V, k)$ is a closed convex cone, then $\text{Av}_\Gamma \mathcal{C}$ is a closed convex cone in $L^2_{\text{sym}}(V, k)^\Gamma$, and its dual in $L^2_{\text{sym}}(V, k)^\Gamma$ under the inner product is $\text{Av}_\Gamma \mathcal{C}^*$. Indeed, for k -tensors T, T' , $\langle T, T' \rangle = \langle \text{Av}_\Gamma T, \text{Av}_\Gamma T' \rangle$, so if $T \in \mathcal{C}$ then $\langle \text{Av}_\Gamma T, \text{Av}_\Gamma T' \rangle \geq 0$ if and only if $T' \in \mathcal{C}^*$.

This shows that the dual of $\text{Av}_\Gamma \text{CP}(V, k)$ is $\text{Av}_\Gamma \text{COP}(V, k)$. Moreover, by Theorem 2.10 and Av_Γ being continuous, closed and linear,

$$\text{Av}_\Gamma \text{COP}(V, k) = \text{Av}_\Gamma \overline{\text{cone}} \bigcup_{r \in \mathbb{N}} C_r(V, k) = \overline{\text{cone}} \bigcup_{r \in \mathbb{N}} \text{Av}_\Gamma C_r(V, k).$$

Taking the dual, by the discussion above and Theorem A.2,

$$\text{Av}_\Gamma \text{CP}(V, k) = (\text{Av}_\Gamma \text{COP}(V, k))^* = \bigcap_{r \in \mathbb{N}} \text{Av}_\Gamma C_r(V, k)^*. \quad \square$$

The situation is more nuanced when Γ is not compact. We restrict the discussion to trace-class kernels on Γ , with Γ a second countable unimodular locally compact group.

By Theorem 3.3, Av_Γ is a bounded operator $\mathcal{B}^1(L^2(\Gamma)) \rightarrow C_0(\Gamma)$. The space $C_0(\Gamma)$ can be difficult to work with, as it has too few compact sets; we extend the codomain of Av_Γ to $L^\infty(\Gamma)$. Since Av_Γ is bounded, it is continuous with respect to the weak topologies on $\mathcal{B}^1(L^2(\Gamma))$ and $L^\infty(\Gamma)$, and so also with respect to the weak* topology on $L^\infty(\Gamma)$ under the duality with $L^1(\Gamma)$, which is weaker.

A warning: although in the compact setting the operator Av_Γ is an orthogonal projection, in the current setting, the range $\text{Av}_\Gamma(\mathcal{B}^1(L^2(\Gamma)))$ is not even closed in $L^\infty(V)$. Indeed, by taking an approximate identity $(\psi_i)_{i \in I}$ and a compact neighborhood basis $(V_j)_{j \in J}$, it can be shown that $\text{Av}_\Gamma(\psi_i \otimes f \mathbf{1}_{V_j})$ converges to f under the weak* topology.

Let $\text{Av}_\Gamma : \mathcal{B}^1(L^2(\Gamma)) \rightarrow L^\infty(\Gamma)$, and define the completely positive cone on $L^\infty(\Gamma)$ by

$$\text{CP}(\Gamma)_{\text{inv}} = \text{cl Av}_\Gamma(\text{CP}(\Gamma, 2) \cap \mathcal{B}^1(L^2(\Gamma))) = \overline{\text{cone}}\{\phi * \phi : \phi \in L^2(\Gamma)_{\geq 0}\},$$

with closure under the weak* topology. The latter equality follows directly from the definition of $\text{CP}(V, 2)$. Define the copositive cone on $L^1(\Gamma)$ by

$$\text{COP}(\Gamma)_{\text{inv}} = \{\rho \in L^1(\Gamma) : \langle \phi * \rho, \phi \rangle \geq 0 \text{ for all } \phi \in L^2(\Gamma)_{\geq 0}\}.$$

By a direct calculation and Theorem A.2, $\text{COP}(\Gamma)_{\text{inv}} = \text{CP}(\Gamma)_{\text{inv}}^*$. Let

$$\check{C}_r(\Gamma)_{\text{inv}}^* = \text{cl Av}_\Gamma\left(\check{C}_r(\Gamma, 2)^* \cap \mathcal{B}^1(L^2(\Gamma))\right),$$

with closure under the weak* topology.

Denote the Borel algebra of Γ by $\mathcal{B}$, and the set of Borel sets with finite and nonzero measure by $\mathcal{B}_{\text{fin}}$. For any function $f : \Gamma \rightarrow \mathbb{R}$ and $A \in \mathcal{B}_{\text{fin}}$, let $K_A f(g, h) = f(h^{-1}g)$ for all $g, h \in A$. If $f \in L^\infty(\Gamma)$, this is the kernel of the operator $\phi \mapsto \phi * f$ restricted to A^2 . Its adjoint is $K_A^* = \text{Av}_\Gamma \circ \text{Res}_A^*$.

We need one more property to obtain the results we want: *amenability*. Amenable σ -compact locally compact groups are those σ -compact locally compact groups that have an *averaging sequence* [86, Proposition 16.14 and 16.16]: a sequence of sets $(A_n)_{n \in \mathbb{N}}$ with $A \in \mathcal{B}_{\text{fin}}$ such that for every finite subset $S \subseteq \Gamma$

$$\lim_{n \rightarrow \infty} \frac{\mu(\bigcap_{s \in S} sA_n)}{\mu(A_n)} = 1.$$

The convergence is uniform in S when the sets S considered are subsets of a fixed compact set. The second statement of the following lemma will be useful later on.

LEMMA 3.6. *If Γ is a second countable unimodular locally compact group and $r \in \mathbb{N}$, then*

$$\check{C}_r(\Gamma)_{\text{inv}}^* \subseteq \bigcap_{A \in \mathcal{B}_{\text{fin}}} K_A^{-1} C_r(A, 2)^*.$$

If Γ is moreover amenable, then

$$\check{C}_r(\Gamma)_{\text{inv}}^* \cap \text{P}(\Gamma) = \bigcap_{A \in \mathcal{B}_{\text{fin}}} K_A^{-1} C_r(A, 2)^* \cap \text{P}(\Gamma),$$

that is, the cones contain the same positive-type functions.

Proof. Let $f = \text{Av}_\Gamma K$ with $K = \sum_{\phi \in \Phi} \lambda_\phi \phi \otimes \phi$ trace class, and such that for all $A \in \mathcal{B}_{\text{fin}}$, $\text{Res}_A K \in C_r(A, 2)^*$. If $B \in C_r(A, 2)$ and $\gamma \in \Gamma$, then γB is

in $C_r(\gamma A, 2)$. By Fubini's theorem and invariance of the Haar measure, and a change of variables,

$$\begin{aligned}
\langle K_A f, B \rangle &= \int_{A^2} \text{Av}_\Gamma K(h^{-1}g) B(g, h) d\mu(g) d\mu(h) \\
&= \int_{A^2} \sum_{\phi \in \Phi} \lambda_\phi \int_\Gamma \phi(\gamma^{-1}) \phi(\gamma^{-1} h^{-1} g) B(g, h) d\mu(\gamma) d\mu(g) d\mu(h) \\
&= \int_\Gamma \int_{A^2} \sum_{\phi \in \Phi} \lambda_\phi \phi(\gamma^{-1}) \phi(\gamma^{-1} h^{-1} g) B(g, h) d\mu(g) d\mu(h) d\mu(\gamma) \\
&= \int_\Gamma \langle \text{Res}_{\gamma A} K, \gamma B \rangle d\mu(\gamma) \geq 0,
\end{aligned}$$

and $\gamma A \in \mathcal{B}_{\text{fin}}$. So, indeed, $K_A f \in C_r(A, 2)^*$ for all $A \in \mathcal{B}_{\text{fin}}$.

Since Av_Γ is continuous, it follows that $\check{C}_r(\Gamma)_{\text{inv}}^* \subseteq \bigcap_A K_A^{-1} C_r(A, 2)^*$ if the latter is weak* closed. For this, it is enough that the maps K_A are bounded linear operators from $L^\infty(\Gamma)$ to $L^2(A^2)$, since then, for all A , $K_A^{-1} C_r(A, 2)^*$ is closed. For $f \in L^\infty(\Gamma)$, $\|K_A f\|_2 \leq \|f\|_\infty \mu(A)^2$, and the result follows.

To prove the second statement of the lemma, it is enough to prove that any $f \in \bigcap_A K_A^{-1} C_r(A, 2)^*$ is the limit of a sequence of averages of trace-class kernels in $\check{C}(\Gamma, 2)^*$.

Let $f \in \bigcap_A K_A^{-1} C_r(A, 2)^* \cap P(\Gamma)$; we may assume that f is continuous. Since we assume that Γ is amenable, take an averaging sequence $(A_n)_n$, and define for all $n \in \mathbb{N}$: $K_n = \text{Res}_{A_n}^* K_{A_n} f$, so K_n is the extension by zeros of $K_{A_n} f$ to Γ . First show that $K_n \in C_r(\Gamma, 2)$ for all n and that it is trace class.

To see that $K_n \in \check{C}(V, 2)^*$, note that for every $A \subseteq A_n$,

$$\text{Res}_A K_n = K_A f \in C_r(A, 2)^*$$

by the first part of this proof. Moreover, by a direct calculation, the extension by zeros of an element of $C_r(A_n, 2)^*$ to a set $A \supseteq A_n$ is in $C_r(A, 2)^*$, so indeed $K_n \in \check{C}(\Gamma, 2)^*$.

Since f is continuous, $K_{A_n} f$ is so as well. Since f is positive type, K_n is positive semidefinite, see for example [45, Proposition 3.35]. Thus,

$$\|K_n\|_{\mathcal{B}^1} = \text{Tr} K_n = \text{Tr} K_{A_n} f = \mu(A_n) f(0) < \infty,$$

and K_n is trace class, so $K_n \in \check{C}(\Gamma, 2)^* \cap \mathcal{B}^1(L^2(\Gamma))$.

Define $f_n = \text{Av}_\Gamma K_n / \mu(A_n)$. An application of the Fubini-Tonelli theorem shows that if K is trace class and positive semidefinite, then $\text{Av}_\Gamma K$ is positive type, so f_n is positive type.

It is left to show that the weak* limit $\lim_n f_n = f$. Take $\rho \in C_c(\Gamma)$ and let $C_\rho = \text{supp } \rho$, then

$$\begin{aligned} |\langle f_n - f, \rho \rangle| &\leq \int_{\Gamma} |\text{Av}_{\Gamma} K_n(g) / \mu(A_n) - f(g)| |\rho(g)| d\mu(g) \\ &\leq \int_{\Gamma} f(g) \left| \int_{\Gamma} \frac{\mathbf{1}_{A_n}(\gamma^{-1}) \mathbf{1}_{A_n}(\gamma^{-1}g) d\mu(\gamma)}{\mu(A_n)} - 1 \right| |\rho(g)| d\mu(g) \\ &\leq \|f\|_{\infty} \|\rho\|_1 \int_{C_\rho} \left| \frac{\mu(A_n \cap gA_n)}{\mu(A_n)} - 1 \right| d\mu(g). \end{aligned}$$

Since $\lim_n \mu(A_n \cap gA_n) / \mu(A_n) = 1$ uniformly for $g \in C_\rho$, and $|\langle f_n - f, \rho \rangle| \rightarrow 0$ follows. Since $C_c(\Gamma)$ is dense in $L^1(\Gamma)$, this proves the theorem. $\square$

THEOREM 3.7. *If Γ is an amenable unimodular second countable locally compact group, then*

$$\text{CP}(\Gamma)_{\text{inv}} = \bigcap_{A \in \mathcal{B}_{\text{fin}}} K_A^{-1} \text{CP}(A, 2).$$

Moreover,

$$\check{C}_0(\Gamma)_{\text{inv}} \supseteq \check{C}_1(\Gamma)_{\text{inv}} \supseteq \cdots \supseteq \text{CP}(\Gamma)_{\text{inv}} \quad \text{and} \quad \text{CP}(\Gamma)_{\text{inv}} = \bigcap_r \check{C}_r(\Gamma)_{\text{inv}}.$$

Proof. By Lemma 3.6 it is enough to prove $\text{CP}(\Gamma)_{\text{inv}} = \bigcap_r \bigcap_A K_A^{-1} C_r(A, 2)^*$. Since

$$\bigcap_{r \in \mathbb{N}} \bigcap_{A \in \mathcal{B}_{\text{fin}}} K_A^{-1} C_r(A, 2)^* = \bigcap_{A \in \mathcal{B}_{\text{fin}}} K_A^{-1} \bigcap_{r \in \mathbb{N}} C_r(A, 2)^* = \bigcap_{A \in \mathcal{B}_{\text{fin}}} K_A^{-1} \text{CP}(A, 2),$$

the result follows from $\text{CP}(\Gamma)_{\text{inv}} = \bigcap_A K_A^{-1} \text{CP}(A, 2)$, which is the subject of the remainder of this proof. In what follows, denote $\mathcal{C} = \bigcap_A K_A^{-1} \text{CP}(A, 2)$.

Let $\phi \in L^2(\Gamma)_{\geq 0}$, $A \in \mathcal{B}_{\text{fin}}$, and $K \in \text{COP}(A, 2)$. Then, by a change of variables and Fubini-Tonelli,

$$\begin{aligned} \langle K_A(\phi \star \phi), K \rangle &= \int_{A^2} \int_{\Gamma} \phi(\gamma^{-1}) \phi(\gamma^{-1}g^{-1}h) d\mu(\gamma) K(g, h) d\mu(g) d\mu(h) \\ &= \int_{A^2} \int_{\Gamma} \phi(\gamma^{-1}g) \phi(\gamma^{-1}h) d\mu(\gamma) K(g, h) d\mu(g) d\mu(h) \\ &= \int_{\Gamma} \int_{A^2} \phi(\gamma^{-1}g) \phi(\gamma^{-1}h) K(g, h) d\mu(g) d\mu(h) d\mu(\gamma) \\ &\geq 0. \end{aligned}$$

Thus, $\phi \star \phi$ lies in $\mathcal{C}$, which is closed, so $\text{CP}(\Gamma)_{\text{inv}} \subseteq \mathcal{C}$ follows.

To prove the other inclusion, show

$$\text{COP}(\Gamma)_{\text{inv}} \subseteq \bigcap_{A \in \mathcal{B}_{\text{fin}}} K_A^{-1} \text{COP}(A, 2) \subseteq \mathcal{C}^*. \quad (16)$$

Indeed, the result then follows from $\mathcal{C}^* \subseteq \text{CP}(\Gamma)_{\text{inv}}^*$ and Theorem A.2.

For the first inclusion of (16), take $\rho \in \text{COP}(\Gamma)_{\text{inv}}$, take $A \in \mathcal{B}_{\text{fin}}$ and take $\phi \in L^2(A)_{\geq 0}$. Then, by a change of variables and Fubini-Tonelli,

$$\begin{aligned} \langle K_A \rho, \phi \otimes \phi \rangle &= \int_{\Gamma^2} \rho(h^{-1}g) \text{Res}_A^* \phi(g) \text{Res}_A^* \phi(h) d\mu(g) d\mu(h) \\ &= \int_{\Gamma} \int_{\Gamma} \rho(g) \text{Res}_A^* \phi(hg) \text{Res}_A^* \phi(h) d\mu(g) d\mu(h) \\ &= \langle \rho, \text{Res}_A^* \phi \star \text{Res}_A^* \phi \rangle \geq 0, \end{aligned}$$

and $K_A \rho \in \text{COP}(A, 2)$.

For the second inclusion in (16), note that for $\rho \in L^1(\Gamma)$ and $A \in \mathcal{B}_{\text{fin}}$,

$$\int_{A^2} K_A \rho(g, h) d\mu(g) d\mu(h) = \int_{\Gamma} \rho(g) (\mathbf{1}_A \star \mathbf{1}_A)(g) d\mu(g).$$

Take an averaging sequence $(A_n)_{n \in \mathbb{N}}$ of Γ , and let $\rho_n = ((\mathbf{1}_{A_n} \star \mathbf{1}_{A_n}) / \mu(A_n)) \rho$ almost everywhere. Then, $\lim_n \rho_n(g) = \rho(g)$ almost everywhere, and for all n , $|\rho_n| \leq |\rho|$. So, by the Lebesgue dominated convergence theorem

$$\lim_{n \rightarrow \infty} \frac{1}{\mu(A_n)} \int_{A_n^2} K_{A_n} \rho(g, h) d\mu(g) d\mu(h) = \int_{\Gamma} \rho(g) d\mu(g).$$

Take $\rho \in \bigcap_A K_A^{-1} \text{COP}(A, 2)$ and $f \in \mathcal{C}$. For all A , $K_A(\rho f) = (K_A \rho)(K_A f)$, so apply the above to ρf to obtain

$$\langle \rho, f \rangle = \lim_{n \rightarrow \infty} \frac{1}{\mu(A_n)} \langle K_{A_n} \rho, K_{A_n} f \rangle \geq 0.$$

In summary,

$$\text{CP}(\Gamma)_{\text{inv}}^* = \text{COP}(\Gamma)_{\text{inv}} \subseteq \mathcal{C}^* \subseteq \text{CP}(\Gamma)_{\text{inv}}^*,$$

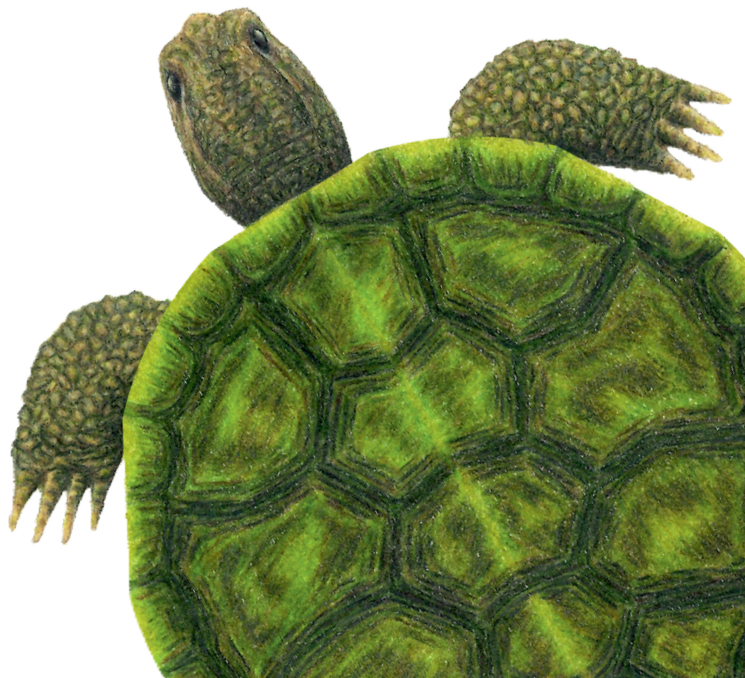
and since all cones are closed, $\mathcal{C} = \text{CP}(\Gamma)_{\text{inv}}$. □

The work by Pier [86] contains many characterizations of amenable groups. All locally compact Abelian groups and all compact groups are amenable, and so are products of amenable groups.

The special orthogonal group $\mathrm{SO}(2)$ is also amenable under the discrete topology, which Matolcsi, Ruzsa, Varga, and Zsámboi recently exploited to prove that the fractional chromatic number of the plane is at least 4 [72]. They could not extend their method to dimensions larger than 2, because the group $\mathrm{SO}(n)$ is not amenable under the discrete topology if $n > 2$. This is closely related to the Banach-Tarski paradox, which prompted the study of amenability.

Part 2

Finite measure spaces



Completely positive formulations

In Part 2 we will study optimization problems of Class I from the introduction. They all have the following form: given a finite measure space $(V, \mathcal{A}, \mu)$, an integer $k \geq 2$, and a symmetric set $E \subseteq V^k$, find

$$\sup\{\mu(I) : I \text{ measurable and } I^k \cap E = \emptyset\}. \quad (17)$$

For general V and E , there is little we can say about such a problem. We treat two special—but still quite broad—subcategories. In this chapter, for these subcategories, we present exact completely positive formulations of (17). These results build heavily on the arguments developed by DeCorte, Oliveira, and Vallentin [36].

We will call the first subcategory the *thick setting*. A problem is in the thick setting if E is a thick set; as the definition is a bit technical, we postpone it until Section 2. Informally, a set in V^k is thick if it is adequately described by sets with positive μ^k -measure.

Example 4.1. For finite V equipped with the counting measure, any $E \subseteq V^k$ is thick, and we find ourselves practicing combinatorics. Thus, the thick setting extends the finite-graph setting.

Example 4.2. An example of a graph in the thick setting that has an infinite vertex set is as follows. If $V = [0, 1]$ with the Lebesgue measure λ , then the set

$$E_{\geq} = \{(x, y) \in [0, 1]^2 : |x - y| \geq 1/2\}$$

is a thick set. It is thick because to prove that a point is not in $E_{\geq}$, it is enough to find an open neighborhood whose intersection with $E_{\geq}$ has λ^2 -measure 0. Figure 1 demonstrates that this is in contrast with a set like

$$E_{=} = \{(x, y) \in [0, 1]^2 : |x - y| = 1/2\}.$$

Indeed, $\mu^2(E_{=}) = 0$, so there is no hope of a similar argument.

A problem is in the second subcategory if (V, E) displays sufficient symmetry; we will call this the *homogeneous setting*. More precisely, V is a compact

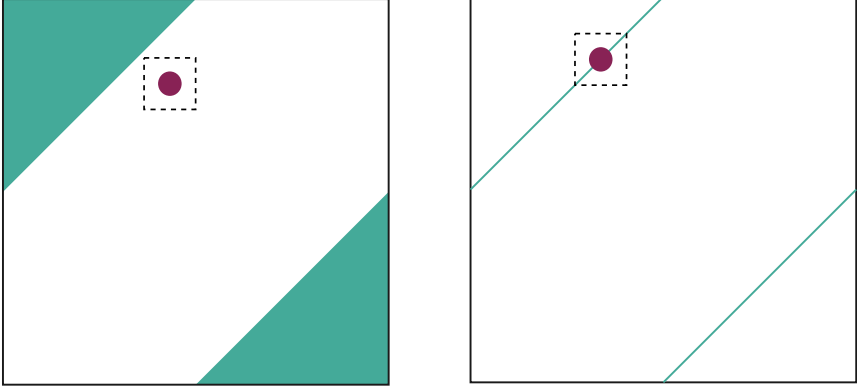


FIGURE 1. On the left, the square $[0, 1]^2$ with $E_{\geq}$ shaded. To see that the point is not in E , it suffices to find an open neighborhood such that its intersection with E has measure 0. On the right, $E_{=}$ is shown, and we see that although the point is in E , the intersection of any open neighborhood with E has measure 0.

space that is homogeneous under a compact metrizable group Γ , $\Gamma E \subseteq E$ under the naturally induced action on V^k , and V is equipped with the quotient of the Haar measure. See Section 3.1 and the appendix for more background. The metric on Γ should furthermore meet technical assumptions which will be explained as we encounter them.

Our main example of an optimization problem in the homogeneous setting is *Witsenhausen's problem*: what is the largest surface measure of a set on the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$ that does not contain a pair of orthogonal points? This is exactly (17) with V the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$, μ the standard uniform surface measure, $E = \{(x, y) \in S^{n-1} : x^\top y = 0\}$, and $\Gamma = O(n)$ —the group of orthogonal $(n \times n)$ -matrices.

Since the orthogonal group preserves inner products, indeed $O(n)E \subseteq E$. Moreover, S^{n-1} is homeomorphic to $O(n)/\text{Stab}(e)$ with $\text{Stab}(e)$ the stabilizer subgroup of an arbitrary point $e \in S^{n-1}$, thus the $(n - 1)$ -dimensional unit sphere is $O(n)$ -homogeneous. Witsenhausen's problem is the subject of the paper [12], and will be treated in depth in Chapter 6.

We can modify Witsenhausen's problem to produce homogeneous examples with $k > 2$. Take V and μ the same, but consider E to be the set of all k -tuples $\{v_1, \dots, v_k\}$ such that $v_i^\top v_j = 0$ for all $i \neq j$. Expression (17) then

gives the maximal measure of a set of unit vectors not containing an orthonormal set of size k . Castro-Silva, Oliveira, Slot, and Vallentin [23] present an upper bound to (17) for such problems, which is a recursive version of the Lovász ϑ -number. These questions are relevant as they are compact versions of questions in Euclidean Ramsey theory; see [23] and references therein.

The example with the thick edge set $\{(x, y) \in [0, 1]^2 : |x - y| \geq 1/2\}$ is not homogeneous, and the examples in the homogeneous setting given above are not thick. However, the two settings have much overlap. For instance, the latter example with the condition $v_i^\top v_j = 0$ replaced by $v_i^\top v_j \in [-1/2, 1/2]$ displays the required symmetry, but also has a thick edge set.

We will use the language of hypergraphs to present completely positive formulations of (17). The proofs of exactness stick closely to the proof for $k = 2$ by DeCorte, Oliveira, and Vallentin, but hold in greater generality: Theorem 5.1 of [36] is the special case of Theorem 4.12 in Section 4 where $k = 2$ and E is a closed set. We reproduce their proof for $k \geq 2$ and without the requirement that E is closed.

1. Upper bounds on the measurable independence number

Let $k \geq 2$ an integer. We say $H = ((V, \mathcal{A}, \mu), E)$ is a *k-uniform measurable hypergraph* if (V, E) is a k -uniform hypergraph, $(V, \mathcal{A}, \mu)$ is a measure space, and $E \subseteq V^k$ is measurable. A *uniform measurable hypergraph* is a k -uniform hypergraph for some integer $k \geq 2$. If H is a uniform measurable hypergraph with measure μ , the *measurable independence number* of H is

$$\alpha(H) = \sup\{\mu(I) : I \subseteq V \text{ measurable and independent}\}.$$

A set $I \subseteq V$ is independent if and only if $I^k \cap E = \emptyset$, thus $\alpha(H)$ is exactly (17).

The measure is part of the data of H , so there can be no confusion with the traditional independence number, and we will often just call $\alpha(H)$ the independence number of H . Having said that, we usually abuse notation, and say $H = (V, E)$ is a uniform measurable hypergraph if the measure on V is clear.

Similarly, we say $H = (V, E, \Gamma)$ is a *k-uniform homogeneous hypergraph under Γ* if (V, E) is a k -uniform hypergraph, $\Gamma \subseteq \text{Aut}((V, E))$ is a σ -compact unimodular locally compact group, V is a homogeneous Γ -space that admits a Γ -invariant Radon measure, and $E \subseteq V^k$ is Borel. A *uniform homogeneous hypergraph under Γ* is a k -uniform homogeneous hypergraph under Γ for some k . We again abuse notation and just say that (V, E) is a uniform homogeneous hypergraph under a group Γ .

In particular, we consider a uniform homogeneous hypergraph to be a vertex-transitive hypergraph and a uniform measurable hypergraph with a Γ -invariant measure. When Γ is in addition compact, the Γ -invariant Radon measure on V is always the pushforward of a Haar measure of Γ under the quotient map; recall that we named it the quotient measure. See Sections 3.1 and A.5 for more details on harmonic analysis and invariant measures.

Finally, if there is no group action, we need another condition that ensures compatibility between a topology, a measure, and an edge set. We say $H = (V, E)$ is a *k-uniform locally independent hypergraph* if V is a Hausdorff space and under the induced topology on V^k , E is Borel and every compact independent set is contained in an open independent set. We say $H = ((V, \mathcal{B}, \mu), E)$ is a *k-uniform measurable locally independent hypergraph* if $((V, \mathcal{B}, \mu), E)$ is a *k-uniform measurable hypergraph*, (V, E) is locally independent, $\mathcal{B}$ is the Borel algebra of V , and μ is a Borel measure. Again, denote such hypergraphs just by (V, E) if all other data is clear.

In all the definitions above, we replace “hypergraph” by “graph” and drop the adjective *k-uniform* when $k = 2$.

We first describe a family of programs that give upper bounds on the independence number of a uniform measurable hypergraph. Let $k \geq 2$ be an integer and $H = (V, E)$ be a *k-uniform measurable hypergraph* with measure μ . Recall the operators

$$\mathcal{T}_{k-2}^* : L_{\text{sym}}^2(V, k) \rightarrow L_{\text{sym}}^2(V, 2), \quad \mathcal{T}_{k-2}^* T(x, y) = \int_{V^{k-2}} T(x, y, v) d\mu^{k-2}(v).$$

Say that a symmetric *k*-tensor $T \in L_{\text{sym}}^2(V, k)$ is *slice positive* if and only if for all $G \in L^2(V^{k-2})_{\geq 0}$ and $f \in L^2(V)$

$$\int_{V^{k-2}} \int_V \int_V T(x, y, v) G(v) f(x) f(y) d\mu^{k-2}(v) d\mu(x) d\mu(y) \geq 0.$$

When V is additionally a topological space, T is continuous, and μ is regular and has full support, slice positivity is equivalent to $T(\cdot, \cdot, v)$ being a positive-semidefinite kernel for all $v \in V^{k-2}$. In particular, completely positive tensors are slice positive.

For a convex cone $\mathcal{C} \subseteq L_{\text{sym}}^2(V, k)$, let

$$\begin{aligned} \vartheta_{\text{big}}(H, \mathcal{C}) &= \sup \langle A, \mathbf{1}^{\otimes k} \rangle \\ &\quad \text{Tr}(\mathcal{T}_{k-2}^* A) = 1, \\ &\quad A(v_1, \dots, v_k) = 0 \quad \text{for all } v_1 \cdots v_k \in E, \\ &\quad A \in \mathcal{C} \text{ is slice positive.} \end{aligned} \tag{18}$$

The normalization constraint should be interpreted as: “the trace of $\mathcal{T}_{k-2}^* A$ exists and is equal to 1”. The edge constraints on A are pointwise; this means that only certain combinations of cones and edge sets produce useful optimization problems, as the support of A as an L^2 class is only defined up to a set of measure 0.

On the other hand, the requirements for $\vartheta_{\text{big}}(H, \mathcal{C}) \geq \alpha(H)$ to hold are mild. For an integer $k \geq 2$ and $H = (V, E)$ a k -uniform measurable hypergraph with $0 < \alpha(H) < \infty$, it is not so hard to prove that, if I is a measurable independent set with $\mu(I) > 0$, then $A = \mathbf{1}_I^{\otimes k} / \mu(I)^{k-1}$ has $\text{Tr}(\mathcal{T}_{k-2}^* A) = 1$ and $\langle A, \mathbf{1}^{\otimes k} \rangle = \mu(I)$. Thus, any $\mathcal{C}$ containing these tensors results in an upper bound.

However, in applications, we usually ask tensors to be continuous. One of the reasons is that the most practical way to obtain feasible solutions is by use of polynomials. Another reason is that the edge set might have measure 0, in which case the constraint on edges does not affect the optimal value, and the upper bound given by ϑ_{big} is trivial.

DeCorte, Oliveira and Vallentin [36] showed that when $G = ((V, \mathcal{B}, \mu), E)$ is a measurable locally independent graph equipped with μ an inner regular Borel measure and V compact such that $0 < \alpha(G) < \infty$, then Urysohn’s Lemma gives continuous kernels $f^{\otimes k}$ with $0 \leq f \leq 1$ and objective value arbitrarily close to $\alpha(G)$. Thus, in this case, the independence number is indeed bounded from above by continuous kernels satisfying the constraints of $\vartheta_{\text{big}}(H, \mathcal{C})$ for a suitable cone $\mathcal{C}$. Their proof goes through without issue in the k -uniform setting.

LEMMA 4.3. *If $k \geq 2$ is an integer, $H = ((V, \mathcal{B}, \mu), E)$ is a k -uniform measurable locally independent hypergraph with μ a finite inner regular Borel measure such that $\alpha(H) > 0$, V is compact, and $\mathcal{C} \subseteq C_{\text{sym}}(V, k)$ is a cone containing all continuous tensors of the form $f^{\otimes k}$ with $f \in L^2(V)_{\geq 0}$, then*

$$\vartheta_{\text{big}}(H, \mathcal{C}) \geq \alpha(H).$$

Lemma 3.4 offers a second way of obtaining continuous tensors. It says that when $H = (V, E)$ is a k -uniform homogeneous hypergraph under a compact group Γ , then for every measurable independent set I with $\mu(I) > 0$,

the map $A = \text{Av}_\Gamma(\mathbf{1}_I^{\otimes k})/\mu(I)^{k-1}$ is continuous, and we can repeat the arguments above for the tensor A , obtaining an upper bound on $\alpha(H)$ if $\mathcal{C}$ contains all tensors of this form. However, we will see that, in the homogeneous setting, for a proof of exactness of the completely positive formulation we introduce later, it is necessary to replace the condition $\text{Tr}(\mathcal{T}_{k-2}^* A) = 1$ in (18) by $\int_V A(v, \dots, v) d\mu(v) = 1$, obtaining an upper bound on $\alpha(H)^{k-1}$. For a convex cone $\mathcal{C} \subseteq C_{\text{sym}}(V, k)$, let

$$\begin{aligned} \vartheta_{\text{small}}(H, \mathcal{C}) = \sup \quad & \langle A, \mathbf{1}^{\otimes k} \rangle \\ & \int_V A(v, \dots, v) d\mu(v) = 1, \\ & A(v_1, \dots, v_k) = 0 \quad \text{for all } v_1 \dots v_k \in E, \\ & A \in \mathcal{C} \text{ is slice positive.} \end{aligned} \quad (19)$$

For $k = 2$ and $\mathcal{C} \subseteq C_{\text{sym}}(V, k)$, we have $\vartheta_{\text{big}}(H, \mathcal{C}) = \vartheta_{\text{small}}(H, \mathcal{C})$.

LEMMA 4.4. *If $k \geq 2$ is an integer, $H = (V, E, \Gamma)$ is a k -uniform homogeneous hypergraph with $0 < \alpha(H) < \infty$, and $\mathcal{C} \subseteq C_{\text{sym}}(V, k)$ is a cone containing all tensors of the form $\text{Av}_\Gamma(f^{\otimes k})$ with $f \in L^k(V)_{\geq 0}$, then*

$$\vartheta_{\text{small}}(H, \mathcal{C}) \geq \alpha(H)^{k-1}.$$

Proof. See the discussion above. □

2. Density and thick sets

The strategies for proving sharpness of a completely positive bound on the independence number for the thick and the homogeneous setting are similar: given a feasible solution A , find a function $f \geq 0$ such that its support is independent and whose measure gives an upper bound on the integral of A over the space. Since the suitable space for the optimization variables is a space of square-integrable functions, by functional-analytic arguments we can only describe values of integrals; we have little control over pointwise properties. We introduce the concept of density to deal with this.

In the thick setting, although the definitions are technical, this approach is rather straightforward. We obtain a zero-measure-removal lemma, Lemma 4.6. It implies that if we find a function $f \geq 0$ such that $\mu^k((\text{supp } f)^k \cap E) = 0$, we can remove a subset of $\text{supp } f$ of measure 0 to obtain an independent set. In the homogeneous setting, things are a little more involved. We also obtain a removal lemma there, though we do not explicitly state it.

Let $(V, \mathcal{A}, \mu)$ be a measure space, and denote the family of all measurable sets with nonzero measure by $\mathcal{A}^+$. Given a subfamily $\mathcal{D} \subseteq \mathcal{A}^+$ that is a

directed set under the opposite of the inclusion relation, define for measurable $A \in \mathcal{A}^+$ and $v \in V$,

$$\text{dens}_v(A) = \lim_{v \in D \in \mathcal{D}} \frac{\mu(D \cap A)}{\mu(D)},$$

when the limit exists. We call $\text{dens}_v(A)$ the *density* of v in A with respect to $\mathcal{D}$. Denote the set of points $v \in V$ such that $\text{dens}_v(A) = 1$ by $D(A)$. Note that $D(A)$ is not necessarily contained in A .

Such a directed set $\mathcal{D} \subseteq \mathcal{A}^+$ is called a *density system* if $\mu(A \triangle D(A)) = 0$ for all $A \in \mathcal{A}^+$, where $\triangle$ is the symmetric difference of sets. If $\mathcal{D}$ is a density system, we call a point $v \in D(A)$ a *density point* of A . In particular, for every $A \in \mathcal{A}^+$, almost every $v \in A$ is a density point of A ; on the other hand, almost every density point of A is in A , from which we conclude $\mu(A) = \mu(D(A)) = \mu(A \cap D(A))$.

For us, there are three main examples of a measure space with a density system. The first is that of a set with a discrete measure, that is, combinations of point measures. The set of singletons forms a density system for such a measure.

The second example is $\mathbb{R}^n$ with the Lebesgue measure. The Lebesgue density theorem states that the set of open balls forms a density system of $\mathbb{R}^n$.

The third example is when the measure space has a finite-rank approximation. Let $(V, \mathcal{A}, \mu)$ be a measure space with a finite-rank approximation $(\mathcal{P}_n)_n$ with associated conditional expectation operators E_n . Let $\mathcal{D} = \bigcup_n \mathcal{P}_n^+$, the set of the parts with positive measure of all partitions in $\mathcal{P}_n$. For all $v \in V$, denote the set in $\mathcal{P}_n$ containing v by $P_n(v)$. Let $A \in \mathcal{A}^+$. We have

$$\lim_n E_n \mathbf{1}_A(v) = \lim_n \sum_{P \in \mathcal{P}_n} \frac{\langle \mathbf{1}_P, \mathbf{1}_A \rangle}{\mu(P)} \mathbf{1}_P(v) = \lim_n \frac{\mu(P_n(v) \cap A)}{\mu(P_n(v))} = \text{dens}_v(A).$$

The martingale convergence theorem, Theorem 2.4, then implies that, indeed, $\mu(D(A) \triangle A) = 0$.

Let $k \geq 1$ be an integer. A set $E \subseteq V^k$ is called $\mathcal{D}$ -*thick* if and only if for all $(v_1, \dots, v_k) \in E$ we have

$$\lim_{\substack{D_1, \dots, D_k \in \mathcal{D} \\ (v_1, \dots, v_k) \in D_1 \times \dots \times D_k}} \frac{\mu^k((D_1 \times \dots \times D_k) \cap E)}{\mu(D_1) \cdots \mu(D_k)} > 0. \quad (20)$$

We simply call E a *thick set* if there exists a density system $\mathcal{D}$ for which E is $\mathcal{D}$ -thick. We also call the quantity (20) the *density* of $(v_1, \dots, v_k)$ in E with respect to $\mathcal{D}$.

A measure space may have many density systems $\mathcal{D}$; which sets are $\mathcal{D}$ -thick depends on the choice of $\mathcal{D}$. This means that, although a finite-rank

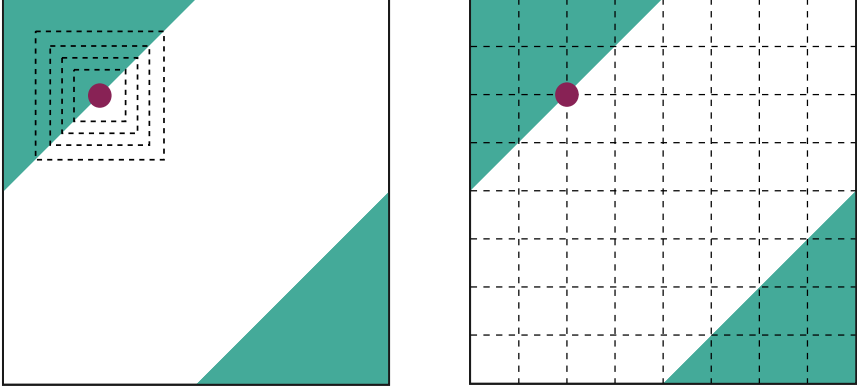


FIGURE 2. The square $[0, 1]^2$ with the set $E_{\geq}$ shaded. The figure on the left shows that if we take $\mathcal{D}_I$ for the density system, the density in $E_{\geq}$ of the point is $1/2$. On the right, the point coincides with the corner of a dyadic square that contains no other points of $E_{\geq}$. This remains true for any subdivision of the grid, so the point has density 0 in $E_{\geq}$ under $\mathcal{D}_D$.

approximation defines a density system, the density system thus obtained is not necessarily the one we want to use, as the next example shows.

Example 4.5. Take the set $E_{\geq} = \{(x, y) : |x - y| \geq 1/2\}$ in the square $[0, 1]^2$ from the introduction. By the Lebesgue density theorem, the set of open intervals $\mathcal{D}_I$ forms a density system of $[0, 1]$. Under this system, $E_{\geq}$ is $\mathcal{D}_I$ -thick. On the other hand, the dyadic decomposition of $[0, 1]$ provides a finite-rank approximation of $[0, 1]$, that is, $(\mathcal{P}_n)_n$ with

$$P_n = \{[2^{-n}i, 2^{-n}(i+1)) : 0 \leq i \leq 2^n - 1\} \cup \{\{1\}\}.$$

The set $\mathcal{D}_D = \bigcup_n \mathcal{P}_n \setminus \{\{1\}\}$ is a density system, but E is not $\mathcal{D}_D$ -thick. For example, the point in Figure 2 has density 0 under $\mathcal{D}_D$, because it is unlucky enough to fall on the corner of a dyadic square that contains no other points of $E_{\geq}$.

The usefulness of density systems and thick sets stems from the following removal lemma, which is reminiscent of the graph removal lemma from combinatorics [43]. It implies that if S is a set of vertices such that the set of edges in S^k has measure 0, then $D(S)$ is independent.

LEMMA 4.6 (Removal lemma). *Let $k \geq 1$ be an integer, $(V, \mathcal{A}, \mu)$ be a measure space with a density system $\mathcal{D}$, and $E \subseteq V^k$ be a $\mathcal{D}$ -thick set. If $S \in \mathcal{A}^+$, then $D(S)^k \cap E \neq \emptyset$ if and only if $\mu^k(S^k \cap E) > 0$.*

Proof. Let $S \in \mathcal{A}^+$ and $E \subseteq V^k$ be $\mathcal{D}$ -thick. Note first that if $S_1, S_2 \subseteq V$ such that $\mu(S_1 \triangle S_2) = 0$, then $\mu^k(S_1^k \triangle S_2^k) = 0$. Indeed, to see this, note that $S_1^k \triangle S_2^k \subseteq (S_1 \cup S_2)^k \setminus (S_1 \cap S_2)^k$, so

$$\begin{aligned} \mu^k(S_1^k \triangle S_2^k) &\leq \mu^k((S_1 \cup S_2)^k \setminus (S_1 \cap S_2)^k) \\ &= \mu^k((S_1 \cup S_2)^k) - \mu^k((S_1 \cap S_2)^k) = \mu(S_1)^k - \mu(S_1)^k = 0. \end{aligned}$$

Thus, if $D(S)^k \cap E = \emptyset$, then $\mu^k(D(S)^k \cap E) = 0$, and since $\mu^k(S^k \triangle D(S)^k) = 0$, we have $\mu^k(S^k \cap E) = 0$.

On the other hand, suppose there is a $(v_1, \dots, v_k) \in D(S)^k \cap E$. Since E is $\mathcal{D}$ -thick, there exists a constant $1 \geq c > 0$ and for each i a set $D_{i,1} \in \mathcal{D}$ such that for all choices $D_1, \dots, D_k \in \mathcal{D}$ with $v_i \in D_i \subseteq D_{i,1}$,

$$\frac{\mu^k((D_1 \times \dots \times D_k) \cap E)}{\mu(D_1) \cdots \mu(D_k)} \geq c.$$

Moreover, each v_i is a density point of S , so for each i there exists a set $D_{i,2} \in \mathcal{D}$ such that for all D with $v_i \in D \subseteq D_{i,2}$,

$$\frac{\mu(D \cap S)}{\mu(D)} > (1 - c/2)^{1/k}.$$

The set $\mathcal{D}$ is directed, so take for each i a set $D_{i,0} \in \mathcal{D}$ contained in $D_{i,1} \cap D_{i,2}$.

For each i , choose $D_i \in \mathcal{D}$ such that $D_i \subseteq D_{i,0}$. Use the following version of the inclusion-exclusion principle. For a measure ν and ν -measurable sets S_1, S_2 , and S_3 ,

$$\begin{aligned} \nu(S_1 \cap S_2 \cap S_3) &= \nu((S_1 \cap S_2) \cap (S_1 \cap S_3)) = \nu(S_1 \cap S_2) + \nu(S_1 \cap S_3) - \\ &\quad \nu((S_1 \cap S_2) \cup (S_1 \cap S_3)) \geq \nu(S_1 \cap S_2) + \nu(S_1 \cap S_3) - \nu(S_1). \end{aligned}$$

Apply this with $S_1 = D_1 \times \cdots \times D_k$, $S_2 = S^k$ and $S_3 = E$, to obtain

$$\begin{aligned}
\frac{\mu^k(S^k \cap E)}{\mu(D_1) \cdots \mu(D_k)} &\geq \frac{\mu^k((D_1 \times \cdots \times D_k) \cap S^k \cap E)}{\mu(D_1) \cdots \mu(D_k)} \\
&\geq \frac{\mu(D_1 \cap S) \cdots \mu(D_k \cap S)}{\mu(D_1) \cdots \mu(D_k)} + \\
&\quad \frac{\mu^k((D_1 \times \cdots \times D_k) \cap E)}{\mu(D_1) \cdots \mu(D_k)} - \frac{\mu(D_1) \cdots \mu(D_k)}{\mu(D_1) \cdots \mu(D_k)} \\
&> (1 - c/2) + c - 1 \\
&= c/2.
\end{aligned}$$

It follows that $\mu^k(S^k \cap E) > 0$. □

3. The thick setting

We first prove $\vartheta_{\text{big}}(H, \text{CP}(V, k)) = \alpha(H)$ for uniform measurable hypergraphs H with a thick edge set. In computational implementations of upper bounds, V is usually equipped with a topology and a Borel measure, and we only consider the continuous tensors in each cone. However, in our treatment of the thick setting, leaving out the continuity condition simplifies matters a little. This suffices, because the programs $\vartheta_{\text{big}}(H, \mathcal{C})$ defined in (18) are maximization problems, so restricting the feasible region decreases the optimal value, and Lemma 4.3 ensures that restricting to continuous tensors still produces a valid bound, which is also sharp.

A key tool in the proof of Theorem 5.1 in [36] is the identification of a compact region of the completely positive cone, its tip, whose extreme points have a simple description. We extend this definition here to k -tensors in a weak sense. An oddity of this extension is that the tip is not necessarily a part of the cone, but once we accept this, everything works as it should.

Let

$$\mathcal{E}_{\text{big}} = \{0\} \cup \{f^{\otimes k} : f \in L^{k/(k-1)}(V)_{\geq 0} \text{ and } \|f\|_{k/(k-1)} = 1\}$$

and define the *big tip* of $\text{CP}(V, k)$ by $\mathcal{T}(\text{CP}(V, k)) = \overline{\text{con}} \mathcal{E}_{\text{big}}$, with closure in the weak topology on $L^{k/(k-1)}$. The appearance of the exponent $k/(k-1)$ —the L^p exponent conjugate to k —is a bit mysterious. As we will see in the proof of the lemma below, it comes from the following interpolation trick. Let m be a natural number, let $\{k_i\}_{i \in [m]}$ be positive real numbers such that $\sum_i k_i = 1$, let $1 \leq p_1, \dots, p_m \leq \infty$ be real numbers, and define p by $1/p = \sum_i (k_i/p_i)$.

For all measurable f ,

$$\|f\|_p \leq \prod_{i=1}^m \|f\|_{p_i}^{k_i}, \quad (21)$$

which follows from repeated application of Hölder's inequality.

The proof of the lemma makes use of the conditional expectation operators associated to a finite-rank approximation. Recall Lemma 2.6, which says in particular that a finite countably generated measure space has a finite-rank approximation.

LEMMA 4.7. *If $(V, \mathcal{A}, \mu)$ is a finite countably generated measure space, $k \geq 2$ is an integer, and $A \in \text{CP}(V, k)$ with $\text{Tr}(\mathcal{T}_{k-2}^* A) \leq 1$, then $A \in \mathcal{T}(\text{CP}(V, k))$.*

Proof. Let $(\mathcal{P}_n)_{n \in \mathbb{N}}$ be a finite-rank approximation of V with the conditional expectation operators E_n and the associated operators A_n as in Section 2.2. Since μ is finite, using Hölder's inequality, $\|A\|_{k/(k-1)}$ is finite. Indeed, if $k \geq 2$ then $k/(k-1) \leq 2$, and for any $1 \leq p \leq q \leq \infty$,

$$\|A\|_p = \langle |A|^p, \mathbf{1} \rangle^{1/p} \leq c \| |A|^p \|_{q/p}^{1/p} = c \|A\|_q < \infty,$$

where $c = \mu^k(V)^{(q/p)/((q/p)-1)}$. By taking $p = k/(k-1)$ and $q = 2$, it follows that $A \in L^{k/(k-1)}(V^k)$, and $A = \lim_n E_n A$ in the $L^{k/(k-1)}$ norm. Hence, it is enough to prove that each $E_n A \in \mathcal{T}(\text{CP}(V, k))$.

Recall that $A_n A \in \text{CP}(\mathcal{P}_n^+, k)$. It has finite expansion $A_n A = \sum_m \lambda_m x_m^{\otimes k}$ with $\lambda_m \in \mathbb{R}_{\geq 0}$ and $x_m \in \mathbb{R}_{\geq 0}^{\mathcal{P}_n^+}$. Denote $f_m = \sum_P (x_m)_P \mathbf{1}_P / \mu(P)$ so that the equality $E_n A = \sum_m \lambda_m f_m^{\otimes k}$ holds, and rescale each λ_m and f_m such that $\|f_m\|_{k/(k-1)} = 1$. Because $0 \in \mathcal{T}(\text{CP}(V, k))$, if $\sum_m \lambda_m \leq 1$ the lemma follows.

By Parseval's identity, for any orthonormal basis Φ ,

$$\text{Tr} E_n K = \sum_{P_1, P_2 \in \mathcal{P}_n^+} \frac{\langle K \mathbf{1}_{P_1}, \mathbf{1}_{P_2} \rangle}{\sqrt{\mu^2(P_{1,2})}} \sum_{\phi \in \Phi} \frac{\langle \mathbf{1}_{P_1}, \phi \rangle \langle \phi, \mathbf{1}_{P_2} \rangle}{\sqrt{\mu(P_1) \mu(P_2)}} \leq \text{Tr}(K).$$

The inequality uses that the set of functions $\mathbf{1}_P / \sqrt{\mu(P)}$ with $P \in \mathcal{P}_n^+$ is also orthonormal.

Since the sum over m is finite and again using Parseval's identity,

$$\begin{aligned} 1 &\geq \text{Tr}(E_n \mathcal{T}_{k-2}^* A) = \text{Tr}(\mathcal{T}_{k-2}^* E_n A) \\ &= \sum_m \lambda_m \|f_m\|_1^{k-2} \sum_{\phi \in \Phi} \langle f_m, \phi \rangle^2 = \sum_m \lambda_m \|f_m\|_1^{k-2} \|f_m\|_2^2. \end{aligned}$$

Now use the interpolation inequality (21) with $p_1 = 1$, $p_2 = 2$, $k_1 = (k-2)/k$ and $k_2 = 2/k$. Then $1/p = (k-1)/k$, so that

$$\sum_m \lambda_m (\|f_m\|_1^{(k-2)/k} \|f_m\|_2^{2/k})^k \geq \sum_m \lambda_m \|f_m\|_{k/(k-1)}^k = \sum_m \lambda_m.$$

The conclusion $A = \lim_n E_n A \in \mathcal{T}(\text{CP}(V, k))$ follows. $\square$

LEMMA 4.8. *If $(V, \mathcal{A}, \mu)$ is a finite countably generated measure space and $k \geq 2$ is an integer, the extreme points of $\mathcal{T}(\text{CP}(V, k))$ lie in $\mathcal{E}_{\text{big}}$.*

Proof. Since every element of $\mathcal{T}(\text{CP}(V, k))$ is nonnegative, 0 certainly is an extreme point.

Milman's theorem [97, Theorem 9.4] says that the extreme points of the big tip $\mathcal{T}(\text{CP}(V, k))$ lie in the weak closure $\text{cl } \mathcal{E}_{\text{big}}$. Suppose that $(f_i^{\otimes k})_{i \in I}$ is a convergent net in $\mathcal{E}_{\text{big}}$. The space $L^{k/(k-1)}(V)$ is reflexive, so by the Banach-Alaoglu theorem its unit ball is compact [32, Theorem V.4.2]. So, since $\|f_i\|_{k/(k-1)} = 1$ for all i , the sequence $(f_i)_i$ has a weakly converging subnet; without loss of generality assume it converges with limit f . It follows that $\|f\|_{k/(k-1)} \leq 1$.

The weak limit of $(f_i^{\otimes k})_i$ is $f^{\otimes k}$. Indeed, the linear span of the set of functions $h_1 \otimes \cdots \otimes h_k \in L^k(V^k)$ is dense in $L^k(V^k)$ —for example, use the martingale convergence theorem. So it is enough to prove that

$$|\langle f_i^{\otimes k} - f^{\otimes k}, h_1 \otimes \cdots \otimes h_k \rangle| \rightarrow 0$$

with $h_j \in L^k(V^k)$ arbitrary. For $i \in I$ and $h_1, \dots, h_k \in L^k(V)$, use the triangle inequality to obtain

$$\begin{aligned} |\langle f_i^{\otimes k} - f^{\otimes k}, h_1 \otimes \cdots \otimes h_k \rangle| &\leq |\langle f_i^{\otimes(k-1)} \otimes (f_i - f), h_1 \otimes \cdots \otimes h_k \rangle| + \\ &\quad |\langle (f_i^{\otimes(k-1)} - f^{\otimes(k-1)}) \otimes f, h_1 \otimes \cdots \otimes h_k \rangle|. \end{aligned}$$

By repeated application of this inequality, $|\langle f_i^{\otimes k} - f^{\otimes k}, h_1 \otimes \cdots \otimes h_k \rangle|$ is bounded from above by k terms of the form

$$|\langle f_i - f, h_j \rangle| \prod_{l \neq j} |\langle \xi_l, h_l \rangle| \leq |\langle f_i - f, h_j \rangle| M^{k-1},$$

where each ξ_l is either f_i or f , and the final inequality is Hölder's inequality with $\|\xi_l\|_{k/(k-1)} \leq 1$ and $M = \max_{1 \leq i \leq k} \|h_i\|_k$. Then, $\lim_i f_i^{\otimes k} = f^{\otimes k}$ with $\|f\|_{k/(k-1)} \leq 1$. Because $0 \in \mathcal{T}(\text{CP}(V, k))$, it follows that $f^{\otimes k}$ is an extreme point if and only if $\|f\|_{k/(k-1)} = 1$ or $f = 0$; conclude that $\mathcal{E}_{\text{big}}$ contains all extreme points of $\mathcal{T}(\text{CP}(V, k))$. $\square$

We have now gathered everything we need to state and prove the main theorem of this section. The following result is independent on the choice of the density system $\mathcal{D}$.

THEOREM 4.9. *Let $k \geq 2$ an integer, $H = ((V, \mathcal{A}, \mu), E)$ be a k -uniform measurable hypergraph with $(V, \mathcal{A}, \mu)$ finite and countably generated measure space. If $\alpha(H) > 0$ and if E is thick, then*

$$\alpha(H) = \vartheta_{\text{big}}(H, \text{CP}(V, k)).$$

Proof. The inequality $\vartheta_{\text{big}}(H, \text{CP}(V, k)) \geq \alpha(H)$ is established in Section 1. The remaining objective of this proof is to show that given a feasible solution A for $\vartheta_{\text{big}}(H, \text{CP}(V, k))$, there exists a measurable independent set I such that $\langle A, \mathbf{1}^{\otimes k} \rangle \leq \mu(I)$.

By Lemma 4.7, a feasible solution A of $\vartheta_{\text{big}}(H, \text{CP}(V, k))$ is in the big tip of $\text{CP}(V, k)$. As $L^{k/(k-1)}(V)$ is reflexive, $L^k(V^k)$ is separable, and $\mathcal{T}(\text{CP}(V, k))$ is closed and bounded, the big tip is weakly metrizable and compact [32, Theorem V.5.1 and Theorem V.4.2], thus Choquet's theorem [97, Theorem 10.7] says that there exists a probability measure ρ on $\mathcal{E}_{\text{big}}$ so that for all T in $L^k(V^k)$ the equality $\langle A, T \rangle = \int_{\mathcal{E}_{\text{big}}} \langle f^{\otimes k}, T \rangle d\rho(f^{\otimes k})$ holds.

Since $A(v) = 0$ for all $v \in E$ and E is measurable,

$$0 = \langle A, \mathbf{1}_E \rangle = \int_{\mathcal{E}_{\text{big}}} \langle f^{\otimes k}, \mathbf{1}_E \rangle d\rho(f^{\otimes k}).$$

Hence,

$$\rho(\{f^{\otimes k} \in \mathcal{E}_{\text{big}} : \langle f^{\otimes k}, \mathbf{1}_E \rangle \neq 0\}) = 0.$$

Furthermore, $\langle A, \mathbf{1}^{\otimes k} \rangle = \int_{\mathcal{E}_{\text{big}}} \langle f^{\otimes k}, \mathbf{1}^{\otimes k} \rangle d\rho(f^{\otimes k})$ and ρ is a probability measure, so there exists a function $f \in L^{k/(k-1)}(V)_{\geq 0}$ with $\|f\|_{k/(k-1)} = 1$ such that $\langle f^{\otimes k}, \mathbf{1}_E \rangle = 0$ and

$$\langle f, \mathbf{1} \rangle^k = \langle f^{\otimes k}, \mathbf{1}^{\otimes k} \rangle \geq \langle A, \mathbf{1}^{\otimes k} \rangle.$$

Take a density system $\mathcal{D}$ on $(V, \mathcal{A}, \mu)$ with respect to which E is $\mathcal{D}$ -thick. Denote the set of density points of $\text{supp } f$ by $D(\text{supp } f)$. Then

$$\langle A, \mathbf{1}^{\otimes k} \rangle \leq \langle f, \mathbf{1} \rangle^k = \langle f, \mathbf{1}_{\text{supp } f} \rangle^k \leq \|f\|_{k/(k-1)}^k \|\mathbf{1}_{\text{supp } f}\|_k^k = \mu(D(\text{supp } f)).$$

It remains to prove that $D(\text{supp } f)$ is independent. Since $\int_E f^{\otimes k}(v) d\mu^k(v) = 0$ and $f \geq 0$, $\mu^k((\text{supp } f)^k \cap E) = 0$. The removal lemma, Lemma 4.6, says that $D(\text{supp } f)^k \cap E = \emptyset$. Conclude that $D(\text{supp } f)$ is an independent set with $\mu(D(\text{supp } f)) \geq \langle A, \mathbf{1}^{\otimes k} \rangle$. $\square$

4. The homogeneous setting

We move on to the homogeneous setting. These results rely on a compact Hausdorff space being metrizable if and only if it is second countable.

We need a final result on martingales. For a finite measure space $(V, \mathcal{A}, \mu)$ with finite-rank approximation $(\mathcal{P}_n)_n$ and conditional expectations E_n , define the *martingale maximal function* of $f \in L^p(V)$ with $1 \leq p < \infty$ by

$$Mf(v) = \sup_n E_n f(v),$$

and call M the *martingale maximal operator*. A proof of the following can be found in [41, Theorem 5.2.7], the important implication being that the maximal function of a p -integrable function is again p -integrable.

THEOREM 4.10 (Martingale maximal theorem). *Let $(V, \mathcal{A}, \mu)$ be a finite measure space with finite-rank approximation $(\mathcal{P}_n)_{n \in \mathbb{N}}$. The martingale maximal function M defines a bounded operator $L^p(V) \rightarrow L^p(V)$ for $1 < p < \infty$, and a continuous operator $L^1(V) \rightarrow L^1(V)$ under the weak topology.*

For V a compact metric space with finite Borel measure μ and $k \geq 2$ an integer, let

$$\mathcal{E}_{\text{small}} = \{0\} \cup \{f^{\otimes k} : f \in L^k(V)_{\geq 0} \text{ and } \|f\|_k = 1\},$$

and define the *small tip* of $\text{CP}(V, k)$ by $\tau(\text{CP}(V, k)) = \overline{\text{conv}} \mathcal{E}_{\text{small}}$ with closure under the weak topology of L^k . With a proof similar to that for the big tip, the extreme points of the small tip are contained in $\mathcal{E}_{\text{small}}$. The proof of an analogue of Lemma 4.7 is slightly different.

For a set $S \subseteq L^2(V^k)$, let $S_c = S \cap C(V^k)$.

LEMMA 4.11. *If V is a compact metrizable space with a finite Borel measure μ , $k \geq 2$ is an integer, $T \in \text{CP}(V, k)_c$, and $\int_V T(x, \dots, x) d\mu(x) \leq 1$, then $T \in \tau(\text{CP}(V, k))$.*

Proof. The space V is compact, second countable, and Hausdorff, so take a finite-rank approximation of V with conditional expectation operators E_n . Let $T \in \text{CP}(V, k)_c$. Since T is continuous and V is compact, T is a bounded function, thus $\lim_n E_n T = T$ under the L^k norm. Similar to the proof of Lemma 4.7, write $E_n T = \sum_m \lambda_{m,n} f_{m,n}^{\otimes k}$, which is a finite sum with every $\lambda_{m,n} \in \mathbb{R}_{\geq 0}$ and $f_{m,n} \in L^k(V)_{\geq 0}$ such that $\|f_{m,n}\|_k = 1$. As opposed to the proof of Lemma 4.7, it is not necessarily true that $\sum_m \lambda_m \leq 1$, but it can be renormalized such that this follows.

Indeed, by the martingale maximal theorem—Theorem 4.10—combined with the dominated convergence theorem,

$$1 \geq \int_V T(x, \dots, x) d\mu(x) = \int_V \lim_n (E_n T(x, \dots, x)) d\mu(x) = \lim_n \sum_m \lambda_{m,n}.$$

Hence, for all $\epsilon > 0$ there is an $n_\epsilon \in \mathbb{N}$ such that $\sum_m \lambda_{m,n} \leq 1 + \epsilon$ for all $n \geq n_\epsilon$. Vice versa, there is a sequence $(\epsilon_n)_n$ with limit 0 such that each $\epsilon_n \geq 0$ and $(1/(1 + \epsilon_n)) \sum_m \lambda_{m,n} \leq 1$. Thus, $(1/(1 + \epsilon_n)) E_n T \in \tau(\text{CP}(V, k))$ for all $n \in \mathbb{N}$ large enough, and since $\lim_n 1/(1 + \epsilon_n) = 1$, $\lim_n (1/(1 + \epsilon_n)) E_n T = T$ under the L^k norm, which concludes the proof. $\square$

As in the thick setting, we make use of a density system; see Section 2 for more details. The difference is that the edge set is not necessarily thick with respect to this density system. This is replaced by the metric being compatible with the group action.

Let V be a metric space with a metric d , $\mathcal{B}$ be its family of Borel sets, and μ be a Borel measure with full support. Then d is called a *density metric* if the set of open balls forms a density system of $(V, \mathcal{B}, \mu)$. In particular, for all $A \in \mathcal{B}^+$ and almost all $v \in A$ we have

$$\text{dens}_v(A) = \lim_{\delta \rightarrow 0} \frac{\mu(A \cap B_\delta(v))}{\mu(B_\delta(v))} = 1,$$

where $B_\delta(v) = \{u \in V : d(u, v) \leq \delta\}$. The most important example of this is the Euclidean metric on $\mathbb{R}^n$, which is a density metric by the Lebesgue density theorem.

Call a metric on a locally compact group Γ *right-invariant* if for all $g \in \Gamma$ and $v_1, v_2 \in V$ we have $d(v_1, v_2) = d(v_1 g, v_2 g)$. This implies that for all $v \in V$ and $\delta \geq 0$, $B_\delta(v g) = B_\delta(v)$. The metric on the orthogonal group $O(n) \subseteq \mathbb{R}^{n \times n}$ inherited from the Euclidean metric on $\mathbb{R}^{n \times n}$ is an example of a right-invariant density metric on a compact group.

Recall that if a group Γ is compact and second countable, then every Hausdorff Γ -homogeneous space is compact and second countable as well.

THEOREM 4.12. *If $k \geq 2$ is an integer, if $H = (V, E, \Gamma)$ is a k -uniform homogeneous hypergraph with Γ a compact group that is metrizable by a right-invariant density metric, and if $\alpha(H) > 0$, then*

$$\alpha(H)^{k-1} = \vartheta_{\text{small}}(H, \text{CP}(V, k)_c).$$

Proof. The inequality $\alpha(H)^{k-1} \leq \vartheta_{\text{small}}(H, \text{CP}(V, k)_c)$ holds by Lemma 4.4.

For the other inequality, the strategy is the same as that of the proof of Theorem 4.9, closely following the steps in [36]: denoting the quotient measure

on V by μ , and given a feasible solution A of $\vartheta_{\text{small}}(H, \text{CP}(V, k)_c)$, prove that there is a measurable independent set I such that $\mu(I)^{k-1} \geq \langle A, \mathbf{1}^{\otimes k} \rangle$.

As opposed to when E is thick, the linear functional $\langle \cdot, \mathbf{1}_E \rangle$ is identically 0. The first objective is to find a workaround for this problem, which goes as follows. The set E has a countable dense subset under the relative topology: for a countable basis $\mathcal{U}$ of the topology of V^k , let S be a set that contains for every $U \in \mathcal{U}$ such that $U \cap E \neq \emptyset$ exactly one point $v \in U \cap E$. Since $\mathcal{U}$ is a countable basis for the topology, S is a countable dense subset of E . Order it by $S = \{(s_{1,1}, \dots, s_{1,k}), (s_{2,1}, \dots, s_{2,k}), \dots\}$ with each $s_{i,j} \in V$.

Since V is metrizable, fix a metric on V producing the topology, and, for fixed n , using compactness, let $\mathcal{U}_n$ be a finite cover of V by open balls of radius $1/n$ with respect to this metric. Choose for all $n \in \mathbb{N}$ and $v \in V$ a set in $\mathcal{U}_n$ that contains v , and denote it $U_n(v)$.

Define for $n \in \mathbb{N}$ the tensor

$$T_n = \sum_{i \in \mathbb{N}} 2^{-i} \frac{\mathbf{1}_{U_n(s_{i,1}) \times \dots \times U_n(s_{i,k})}}{\mu(U_n(s_{i,1})) \cdots \mu(U_n(s_{i,k}))}.$$

It is $k/(k-1)$ -integrable, for if $M_n = \max\{\mu(U) : U \in \mathcal{U}_n\}$, then by Minkowski's inequality

$$\begin{aligned} \|T_n\|_{k/(k-1)} &\leq \sum_{i \in \mathbb{N}} \left(\int_{V^k} \left(2^{-i} \frac{\mathbf{1}_{U_n(s_{i,1}) \times \dots \times U_n(s_{i,k})}(v)}{\mu(U_n(s_{i,1})) \cdots \mu(U_n(s_{i,k}))} \right)^{\frac{k}{k-1}} d\mu^k(v) \right)^{\frac{k-1}{k}} \\ &= \sum_{i \in \mathbb{N}} 2^{-i} \frac{(\mu(U_n(s_{i,1})) \cdots \mu(U_n(s_{i,k})))^{(k-1)/k}}{\mu(U_n(s_{i,1})) \cdots \mu(U_n(s_{i,k}))} \\ &= \sum_{i \in \mathbb{N}} 2^{-i} (\mu(U_n(s_{i,1})) \cdots \mu(U_n(s_{i,k})))^{-1/k} \\ &\leq 2/M_n. \end{aligned}$$

Since any continuous $A : V^k \rightarrow \mathbb{R}$ is uniformly continuous, for every $\epsilon > 0$ there is an $n_\epsilon \in \mathbb{N}$ so that for all $n \geq n_\epsilon$

$$|A(v_1, \dots, v_k) - A(s_{i,1}, \dots, s_{i,k})| \leq \epsilon \quad \text{for all } i \in \mathbb{N} \text{ and } v_j \in U_n(s_{i,j}).$$

Whence, if $A : V^k \rightarrow \mathbb{R}$ is continuous, then

$$\lim_n \langle T_n, A \rangle = \sum_{i \in \mathbb{N}} 2^{-i} A(s_{i,1}, \dots, s_{i,k}). \quad (22)$$

Let A be a feasible solution of $\vartheta_{\text{small}}(H, \text{CP}(V, k)_c)$. Since H is homogeneous and all constraints and the objective are invariant under the action

of Γ on A , A can be assumed invariant under the diagonal action of Γ , that is, $\text{Av}_\Gamma A = A$.

Since $\int_V A(x, \dots, x) d\mu(x) = 1$, by Lemma 4.11, $A \in \tau(\text{CP}(V, k))$. Because the Banach space $L^k(V^k)$ is reflexive, $L^{k/(k-1)}(V^k)$ is separable, and the small tip is closed and bounded in $L^k(V^k)$, it is compact, and the weak topology on the small tip is metrizable [32, Theorem V.4.2 and Theorem V.5.1]. Then, Choquet's theorem [97, Theorem 10.7] says that there exists a probability measure ρ on $\mathcal{E}_{\text{small}}$ such that for all $F \in L^{k/(k-1)}(V^k)$,

$$\langle A, F \rangle = \int_{\mathcal{E}_{\text{small}}} \langle f^{\otimes k}, F \rangle d\rho(f^{\otimes k}). \quad (23)$$

Since A is 0 on E , by Γ -invariance of A and Fatou's lemma,

$$\begin{aligned} 0 &= \lim_{n \in \mathbb{N}} \langle \text{Av}_\Gamma A, T_n \rangle \\ &= \lim_{n \in \mathbb{N}} \langle A, \text{Av}_\Gamma T_n \rangle \\ &\geq \int_{\mathcal{E}_{\text{small}}} \liminf_{n \in \mathbb{N}} \langle f^{\otimes k}, \text{Av}_\Gamma T_n \rangle d\rho(f^{\otimes k}) \\ &= \int_{\mathcal{E}_{\text{small}}} \liminf_{n \in \mathbb{N}} \langle \text{Av}_\Gamma f^{\otimes k}, T_n \rangle d\rho(f^{\otimes k}). \end{aligned}$$

Thus, since in the above T_n and all f s are nonnegative, the set

$$\{f^{\otimes k} : \liminf_{n \in \mathbb{N}} \langle \text{Av}_\Gamma f^{\otimes k}, T_n \rangle \neq 0\}$$

has measure 0 under ρ .

This, together with Equation (23) applied to $\langle A, \mathbf{1}^{\otimes k} \rangle$, shows that there exists an $f \in L^k(V)_{\geq 0}$ with $\|f\|_k = 1$ such that $\langle f^{\otimes k}, \mathbf{1}^{\otimes k} \rangle \geq \langle A, \mathbf{1}^{\otimes k} \rangle$ and

$$\liminf_{n \in \mathbb{N}} \langle \text{Av}_\Gamma f^{\otimes k}, T_n \rangle = 0.$$

Since $\text{Av}_\Gamma f^{\otimes k}$ is continuous by Lemma 3.4, Equation (22), and density of S in E imply that $\text{Av}_\Gamma f^{\otimes k}(v_1, \dots, v_k) = 0$ if $v_1 \cdots v_k \in E$.

Identify V with a quotient of Γ and let $p : \Gamma \rightarrow V$ be the quotient map. Let $I = D(p^{-1}(\text{supp } f)) \cap p^{-1}(\text{supp } f)$, where $D(p^{-1}(\text{supp } f))$ is the set of density points of $p^{-1}(\text{supp } f)$ with respect to the right-invariant density metric on Γ . Denote the Haar measure on Γ by ν . Then, since $p(I) \subseteq \text{supp } f$,

$$\mu(\text{supp } f) = \nu(p^{-1}(\text{supp } f)) = \nu(I) \leq \nu(p^{-1}(p(I))) = \mu(p(I)) \leq \mu(\text{supp } f),$$

so that the inequalities are equalities, and

$$\langle A, \mathbf{1}^{\otimes k} \rangle \leq \langle f, \mathbf{1} \rangle^k \leq \|f\|_k^k \|\mathbf{1}_{\text{supp } f}\|_{k/(k-1)}^k = \mu(p(I))^{k-1}.$$

If $p(I)$ is independent, the theorem follows.

Let $h_1, \dots, h_k \in I$ be distinct points. Because $\text{Av}_\Gamma f^{\otimes k}$ is 0 on edges, it suffices to show $\text{Av}_\Gamma(f \circ p)^{\otimes k}(h_1, \dots, h_k) > 0$. The function $f \circ p$ is strictly positive on I , and

$$\text{Av}_\Gamma(f \circ p)^{\otimes k}(h_1, \dots, h_k) = \int_{\bigcap_{i=1}^k I h_i^{-1}} (f \circ p)^{\otimes k}(g h_1, \dots, g h_k) d\nu(g),$$

so that it suffices to show that $\nu(\bigcap_{i=1}^k I h_i^{-1}) > 0$.

Since all h_i are density points of $p^{-1}(\text{supp } f)$, and so also of I , there exists a $\delta > 0$ such that for all i

$$\frac{\nu(I \cap B_\delta(h_i))}{\nu(B_\delta(h_i))} \geq \frac{k}{k+1}.$$

Let $N_i = I h_i^{-1}$, then

$$\nu\left(\bigcap_{i=1}^k N_i \cap B_\delta(1)\right) \geq \sum_{i=1}^k \nu(N_i \cap B_\delta(1)) - (k-1)\nu(B_\delta(1)).$$

Indeed, for two subsets $S_1, S_2 \subseteq V$, $\nu(S_1 \cap S_2) = \nu(S_1) + \nu(S_2) - \nu(S_1 \cup S_2)$ by the inclusion-exclusion principle. For $\{S_i\}_{i \in [l]}$ with $l > 2$, apply induction by

$$\nu\left(S_l \cap \bigcap_{i < l} S_i\right) = \nu(S_l) + \nu\left(\bigcap_{i < l} S_i\right) - \nu\left(S_l \cup \bigcap_{i < l} S_i\right),$$

and the claim follows by taking $S_i = N_i \cap B_\delta(1)$ and $N_i \cap B_\delta(1) \subseteq B_\delta(1)$ for all i . Since the metric and the Haar measure are right-invariant, it follows that

$$\begin{aligned} \nu\left(\bigcap_{i=1}^k N_i \cap B_\delta(h_i)\right) &\geq \sum_{i=1}^k \nu(I \cap B_\delta(h_i)) - (k-1)\nu(B_\delta(1)) \\ &\geq \nu(B_\delta(1)) \left(\frac{k^2}{k+1} - (k-1) \right) = \frac{\nu(B_\delta(1))}{k+1} > 0, \end{aligned}$$

which completes the proof. $\square$

CHAPTER 5

Convergence through completely positive programming

We work from the general to the specific in this chapter. First, we define a completely positive hierarchy and versions of the moment hierarchy and the block moment hierarchy for the measurable independence number of certain uniform measurable hypergraphs, which we then compare to each other. The proofs are algebraic in nature, and we can state them in great generality.

We also prove that the moment and block moment hierarchies give an upper bound on the measurable independence number, but only for uniform homogeneous hypergraphs. It is unclear whether a similar result holds without the presence of a group.

After that, we investigate convergence of these hierarchies to the measurable independence number, based on the exactness results of the previous chapter. These proofs depend on functional-analytic specifics, and in the homogeneous setting even on details of the representation theory of the group, limiting our results to only some measurable graphs. Most results in this chapter are based on [12], but the extension to hypergraphs is new.

1. Spaces of subsets

We use the convention $V^0 = \{\emptyset\}$. If $v \in V^r$, denote by $\llbracket v \rrbracket$ the set of its coefficients; that is, $\llbracket \emptyset \rrbracket = \emptyset$, and $\llbracket (v_1, \dots, v_r) \rrbracket = \{v_1, \dots, v_r\}$. This defines a map $\llbracket \cdot \rrbracket : V^r \rightarrow \text{Sub}_r(V)$. We omit the index r from the notation, but to avoid ambiguity, we do include it in the notation for the inverse image:

$$\llbracket S \rrbracket_r^{-1} = \{v \in V^r : \llbracket v \rrbracket = S\},$$

for all $S \in \text{Sub}_r(V)$.

If $(V, \mathcal{A}, \mu)$ is a measure space, we can turn $\text{Sub}_r(V) \setminus \{\emptyset\}$ into a measure space with measure μ_{sub}^r by taking the pushforward of μ^r under $\llbracket \cdot \rrbracket$. That is, $S \subseteq \text{Sub}_r(V) \setminus \{\emptyset\}$ is measurable if and only if $\llbracket S \rrbracket_r^{-1} \in \mathcal{A}^r$, and its measure is $\mu_{\text{sub}}^r(S) = \mu(\llbracket S \rrbracket_r^{-1})$. Define furthermore $\mu_{\text{sub}}^r(\{\emptyset\}) = 1$. In other

words, if $f : \text{Sub}_r(V) \rightarrow \mathbb{R}$ is a measurable function, then

$$\int_{\text{Sub}_r(V)} f(S) d\mu_{\text{sub}}^r(S) = f(\emptyset) + \int_{V^r} f(\llbracket v \rrbracket) d\mu(v).$$

Therefore, f is integrable if and only if $f \circ \llbracket \cdot \rrbracket$ is.

Let V be a Hausdorff space. For a topology on $V^0 = \text{Sub}_0(V) = \{\emptyset\}$ there is no choice. Using the map $\llbracket \cdot \rrbracket$, a topology on V induces a topology on $\text{Sub}_r(V)$, called the *standard topology* of $\text{Sub}_r(V)$; indeed, for $r \geq 1$, $\llbracket \cdot \rrbracket$ is surjective onto $\text{Sub}_r(V) \setminus \{\emptyset\}$, so we equip it with the quotient topology: a collection $S \subseteq \text{Sub}_r(V)$ is open if and only if $\llbracket S \rrbracket_r^{-1}$ is open in V^r . We then take the disjoint union with $\{\emptyset\}$ to obtain a topology on all of $\text{Sub}_r(V)$. The spaces $\text{Sub}_r(V)$ are Hausdorff, and if V is compact, so are all $\text{Sub}_r(V)$ [51].

If $1 \leq r \leq k$ and $U_1, \dots, U_r \subseteq V$ are open sets, then the set

$$(U_1, \dots, U_r)_k = \{A \in \text{Sub}_k(V) : A \cap U_i \neq \emptyset \text{ for all } i \text{ and } A \subseteq U_1 \cup \dots \cup U_r\}$$

is open in $\text{Sub}_k(V)$. If the U_i are pairwise disjoint, call such a set *basic open*. Handel showed [51, Proposition 2.11] that the collection of all basic open sets forms a basis for the topology on $\text{Sub}_k(V)$. For more background on this topology, see Handel [51]. The spaces $\text{Sub}_{=r}(V)$ inherit measures and topologies from $\text{Sub}_r(V)$ by restriction.

If V is equipped with a topology, the union map

$$\cup : \text{Sub}_r(V) \times \text{Sub}_s(V) \rightarrow \text{Sub}_{r+s}(V), \quad (S, T) \mapsto S \cup T$$

is continuous. So, it induces a bounded linear operator

$$M : C(\text{Sub}_{r+s}(V)) \rightarrow C(\text{Sub}_r(V) \times \text{Sub}_s(V)), \quad Mf(S, T) = f(S \cup T)$$

for all $r, s \geq 0$ [51, Proposition 2.14].

Let $k \geq 2$ be an integer and $H = (V, E)$ be a k -uniform measurable, or locally independent, or measurable locally independent hypergraph. Denote the set of all independent subsets of H with cardinality at most r by $\mathcal{I}_r$. A subset K of V is called a *clique* of H if every k -subset of K is an edge. Denote the set of all cliques of cardinality at most r by $\mathcal{K}_r$. If $|K| < k$, then K vacuously satisfies this definition, so if $r < k$, $\mathcal{K}_r = \text{Sub}_r(V)$. Define the spaces $\mathcal{I}_{=r} = \mathcal{I}_r \cap \text{Sub}_{=r}(V)$ and $\mathcal{K}_{=r} = \mathcal{K}_r \cap \text{Sub}_{=r}(V)$.

We will often switch between the identification $E = \mathcal{K}_{=k}$ and the identification of E with a subset of V^k . Call $\bar{E} = \text{Sub}_{=k}(V) \setminus E = \mathcal{I}_{=k}$ the set of *nonedges* of H , and define the *complement hypergraph* $\bar{H}$ of H by $\bar{H} = (V, \bar{E})$.

2. Three hierarchies and how they compare

We now formulate analogues of the moment hierarchy and the block moment hierarchy from Section 2 of the introduction for the independence number of a uniform measurable hypergraph, and show that for uniform homogeneous hypergraphs, these hierarchies are stronger than the completely positive hierarchy.

Let $k \geq 2$ be an integer, V be a compact Hausdorff space, and $H = (V, E)$ a k -uniform measurable hypergraph with Borel measure μ . For an integer $r \geq 1$, let $M : C(\text{Sub}_{2r}(V)) \rightarrow C_{\text{sym}}(\text{Sub}_r(V))$ be the operator induced by the union map:

$$(M\phi)(S, T) = \phi(S \cup T).$$

The r th level of the *moment hierarchy* for H is

$$\begin{aligned} M_r(H) = \sup \int_{\text{Sub}_{=1}(V)} \phi(S) d\mu_{\text{sub}}^1(S) \\ \phi(\emptyset) = 1, \\ \phi(S) = 0 \quad \text{if } S \in \text{Sub}_{2r}(V) \text{ not independent,} \\ \phi \in C(\text{Sub}_{2r}(V)), \quad M\phi \in C_{\text{sym}}(\text{Sub}_r(V))_{\geq 0}. \end{aligned}$$

For $r \geq 2$ and $Q \in \text{Sub}_{r-2}(V)$, let $M_Q : C(\text{Sub}_r(V)) \rightarrow C_{\text{sym}}(\text{Sub}_1(V))$ be the operator

$$M_Q K(S, T) = K(Q \cup S \cup T).$$

The r th level of the *block moment hierarchy* is the program

$$\begin{aligned} \text{block}M_r(H) = \sup \int_{\text{Sub}_{=1}(V)} \phi(S) d\mu_{\text{sub}}^1(S) \\ \phi(\emptyset) = 1, \\ \phi(S) = 0 \quad \text{if } S \in \text{Sub}_r(V) \text{ not independent,} \\ \phi \in C(\text{Sub}_r(V)), \\ M_Q \phi \in C_{\text{sym}}(\text{Sub}_1(V))_{\geq 0} \quad \text{for all } Q \in \mathcal{I}_{r-2}. \end{aligned}$$

The restriction of a feasible solution of $M_{r+1}(H)$ to $\text{Sub}_{2r}(H)$ is continuous [51, Proposition 2.4], hence is a feasible solution of $M_r(H)$. A similar statement holds for $\text{block}M_{r+1}(H)$ and $\text{block}M_r(H)$. Thus,

$$M_1(H) \geq M_2(H) \geq \cdots \quad \text{and} \quad \text{block}M_1(H) \geq \text{block}M_2(H) \geq \cdots.$$

Moreover, if we have $M\phi \in C_{\text{sym}}(\text{Sub}_r(V))_{\geq 0}$, then for all sets $Q \in \mathcal{I}_{r-1}$ and $S, T \in \text{Sub}_1(V)$,

$$M_Q \phi(S, T) = \phi(Q \cup S \cup T) = M\phi(Q \cup S, Q \cup T),$$

so $M_Q \phi \geq 0$ for all $Q \in \mathcal{I}_{r-1}$, and $M_r(H) \leq \text{block}M_{r+1}(H)$.

THEOREM 5.1. *Let $k \geq 2$ be an integer, and $H = (V, E, \Gamma)$ be a k -uniform homogeneous hypergraph with Γ a compact group. If $\alpha(H) > 0$ and $r \geq k - 1$, then $M_r(H) \geq \alpha(H)$ and $\text{block}M_{r+1}(H) \geq \alpha(H)$.*

Proof. In light of the inequality $M_r(H) \leq \text{block}M_{r+1}(H)$, it suffices to prove that $M_r(H) \geq \alpha(H)$ for all $r \geq k - 1$.

Suppose H is homogeneous under a compact group Γ and μ is the quotient of the Haar measure of Γ . Fix $r \geq k - 1$ an integer and $I \subseteq V$ an independent set with nonzero measure. Define $F(v) = \text{Av}_\Gamma \mathbf{1}_I^{\otimes 2r}$, which by Lemma 3.4 is continuous. Since $\mathbf{1}_I$ only takes values in $\{0, 1\}$, the function F only depends on $\llbracket v \rrbracket$, that is, if $v, w \in V^{2r}$ such that $\llbracket v \rrbracket = \llbracket w \rrbracket$, then $F(v) = F(w)$.

Define $\phi \in C(\text{Sub}_{2r}(V))$ by $\phi(\emptyset) = 1$ and $\phi(\llbracket v \rrbracket) = F(v)$ for all $v \in V^r$. To show that ϕ is indeed continuous, it is enough to show it is continuous on $\text{Sub}_{2r}(V) \setminus \{\emptyset\}$. This is true, because F is continuous and $\text{Sub}_{2r}(V)$ has the quotient topology under $\llbracket \cdot \rrbracket$. It then follows that $\phi \in M^{-1}C_{\text{sym}}(\text{Sub}_r(V))_{\geq 0}$, for example by looking at the restrictions of $M\phi$ to finite principle submatrices.

Now, $\int_{\text{Sub}_{=1}(V)} \phi(S) d\mu_{\text{sub}}^1(S) = \int_V F(\llbracket v \rrbracket) d\mu(v) = \mu(I)$. Moreover, Γ preserves edges, so $\phi(S) = 0$ if S is not independent, which concludes the proof. $\square$

Recall definitions (18) of $\vartheta_{\text{big}}(H, \mathcal{C})$ and (19) of $\vartheta_{\text{small}}(H, \mathcal{C})$, and recall that $S_c = S \cap C(V, k)$ for any $S \subseteq L^2(V, k)$.

THEOREM 5.2. *Suppose $k \geq 2$ is an integer and $H = ((V, \mathcal{A}, \mu), E)$ is a k -uniform measurable hypergraph. If V is a compact Hausdorff space and μ is a finite Borel measure on V , then, for every $r \geq 1$,*

$$M_{k+r}(H) \leq \text{block}M_{k+r+1}(H) \leq \vartheta_{\text{small}}(H, C_r(V, k)_c^*)^{1/(k-1)}$$

and

$$M_{k+r}(H) \leq \text{block}M_{k+r+1}(H) \leq \vartheta_{\text{big}}(H, C_r(V, k)_c^*).$$

Proof. Fix an integer $r \geq 1$. First show the inequality

$$\text{block}M_{k+r+1}(H) \leq \vartheta_{\text{small}}(H, C_r(V)_c^*)^{1/(k-1)}.$$

Let $\phi \in C(\text{Sub}_{k+r+1}(V))$ be a feasible solution of $\text{block}M_{k+r+1}(H)$ with nonzero objective value. Define the function $F : V^{k+r} \rightarrow \mathbb{R}$ by

$$F(v, w) = \phi(\llbracket v \rrbracket \cup \llbracket w \rrbracket) = M_{\{v_3, \dots, v_k\} \cup \llbracket w \rrbracket}(v_1, v_2),$$

where $v = (v_1, \dots, v_k) \in V^k$ and $w \in V^r$. Then, F is a continuous and symmetric $(k+r)$ -tensor, so $\mathcal{T}_r^* F$ is continuous and for $v = (v_1, \dots, v_k) \in V^k$,

$$\begin{aligned} (\mathcal{T}_r^* F)(v) &= \int_{V^r} F(v, w) d\mu^r(w) \\ &= \int_{V^r} \phi(\llbracket v \rrbracket \cup \llbracket w \rrbracket) d\mu^r(w) \\ &= \int_{V^r} M_{\{v_3, \dots, v_k\} \cup \llbracket w \rrbracket} \phi(v_1, v_2) d\mu(w). \end{aligned}$$

Since $M_{\{v_3, \dots, v_k\} \cup \llbracket w \rrbracket} \phi$ is positive semidefinite for every $w \in V^r$ and $v_i \in V$ for $3 \leq i \leq k$, $\mathcal{T}_r^* F$ is slice positive. To see that $\mathcal{T}_r^* F \in C_r(V, k)_c^*$, it is enough to see that $F \geq 0$. This is true, because if $\llbracket v \rrbracket \cup \llbracket w \rrbracket$ is not independent, $F(v, w) = \phi(\llbracket v \rrbracket \cup \llbracket w \rrbracket) = 0$; otherwise, $F(v, w) = M_{\{v_2, \dots, v_k\} \cup \llbracket w \rrbracket}(v_1, v_1)$ is a diagonal entry of a continuous positive-semidefinite kernel, thus nonnegative.

Let $A = \mathcal{T}_r^* F$. The above also shows that $A(v) = 0$ for all $v \in E$. Let $\tau = \int_V A(v, \dots, v) d\mu(v)$. The objective of the remainder of the proof is to show that $\tau > 0$, and that $\tau^{-1}A$ is a feasible solution of $\vartheta(H, C_r(V, k)^*)$ with objective value at least $\int_{\text{Sub}=1(v)} \phi(S) d\mu_{\text{sub}}^1(S)$.

For an integer $0 \leq t \leq k+r+1$, write

$$\Phi_t = \int_{V^t} \phi(\llbracket v \rrbracket) d\mu^t(v).$$

First, show that the matrix

$$\begin{pmatrix} \Phi_t & \Phi_{t+1} \\ \Phi_{t+1} & \Phi_{t+2} \end{pmatrix} \quad (24)$$

is positive semidefinite for all integers $0 \leq t \leq k+r-1$.

Indeed, fix an integer t so that $0 \leq t \leq k+r-1$ and let $B : \text{Sub}_1(V)^2 \rightarrow \mathbb{R}$ be such that

$$B(S, T) = \int_{V^t} \phi(\llbracket v \rrbracket \cup S \cup T) d\mu^t(v) = \int_{V^t} (M_{\llbracket v \rrbracket})(S, T) d\mu^t(v).$$

Then B is a positive-semidefinite kernel, for example by Fubini's theorem, and $B(S, T)$ only depends on $S \cup T$. Moreover,

$$\begin{aligned} B(\emptyset, \emptyset) &= \int_{V^t} \phi(\llbracket v \rrbracket) d\mu(v) = \Phi_t, \\ \int_V B(\emptyset, \{x\}) d\mu(x) &= \int_V \int_{V^t} \phi(\llbracket v \rrbracket \cup \{x\}) d\mu^t(v) d\mu(x) = \Phi_{t+1}, \text{ and} \\ \int_{V^2} B(\{x\}, \{y\}) d\mu^2(x, y) &= \int_{V^2} \int_{V^t} \phi(\llbracket v \rrbracket \cup \{x, y\}) d\mu^t(v) d\mu^2(x, y) = \Phi_{t+2}, \end{aligned}$$

so the matrix in (24) is positive semidefinite. Indeed, for all x_1 and $x_2 \in \mathbb{R}$,

$$(x_1, x_2) \begin{pmatrix} \Phi_t & \Phi_{t+1} \\ \Phi_{t+1} & \Phi_{t+2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \langle B, (x_1 \mathbf{1}_{\text{Sub}=0(V)} + x_2 \mathbf{1}_{\text{Sub}=1(V)})^{\otimes 2} \rangle \geq 0.$$

Since $\Phi_0 = 1$ and since ϕ has objective value $\Phi_1 > 0$, it follows that $\Phi_2 > 0$ as well. Repeating the argument, it follows that $\Phi_t > 0$ for all t . Hence, for fixed t , $\Phi_t \Phi_{t+2} - \Phi_{t+1}^2 \geq 0$, whence $\Phi_{t+2} \Phi_{t+1}^{-1} \geq \Phi_{t+1} \Phi_t^{-1}$. Apply the inequality repeatedly to get

$$\Phi_{k+r+1} \Phi_{r+1}^{-1} = \Phi_{k+r+1} \Phi_{k+r}^{-1} \Phi_{k+r} \Phi_{k+r-1}^{-1} \Phi_{k+r-1} \cdots \Phi_{r+2} \Phi_{r+1}^{-1} \geq (\Phi_1 \Phi_0^{-1})^{k-1}.$$

Moreover, $\Phi_{r+1} = \tau$ and $\Phi_{k+r+1} = \langle A, \mathbf{1}^{\otimes k} \rangle$, hence $\tau^{-1}A$ is a feasible solution of $\vartheta_{\text{small}}(H, C_r(V, k)_c^*)$ with objective value at least Φ_1^{k-1} , proving the inequality $\text{blockM}_{k+r+1}(H) \leq \vartheta_{\text{small}}(H, C_r(V, k)_c^*)^{1/(k-1)}$.

The proof of the inequality $\text{blockM}_{k+r+1}(H) \leq \vartheta_{\text{big}}(H, C_r(V, k)_c^*)$ is entirely the same, but with the final step suitably adjusted. This completes the proof. $\square$

3. Convergence of the completely positive hierarchy

The main objective of this section is to show that for certain measurable graphs G the completely positive hierarchy converges to $\alpha(G)$. These results are based on Bekker, Kuryatnikova, Oliveira, and Vera [12, Theorems 5.5 and 5.6], with the difference that the result for thick edge sets no longer depends on a group action, and is complemented by Theorem 4.9 to show the hierarchy indeed converges to the independence number. Another small difference with [12] is that convergence in the $O(n)$ -homogeneous setting on S^{n-1} no longer requires the edge set to be closed.

The ideal situation for convergence of the hierarchy is when the feasible regions of each level lie in a mutual compact set and the constraints and objective are continuous. These kinds of problems are rare, as compactness of the feasible region and continuity of the constraints are competing properties: the more continuous linear functionals a topological vector space has, the fewer

compact sets there are. It turns out, however, that we get a little lucky in this regard. The situation is particularly favorable when the edge set is thick.

As everything in this section happens on a graph, write $\text{CP}(V) = \text{CP}(V, 2)$ and $C_r(V) = C_r(V, 2)$. A slice-positive symmetric kernel is just a positive-semidefinite kernel. Moreover, $\vartheta_{\text{small}}(G, \mathcal{C})$ and $\vartheta_{\text{big}}(G, \mathcal{C})$ coincide for $k = 2$, so we denote both by $\vartheta(G, \mathcal{C})$.

THEOREM 5.3. *Let $G = ((V, \mathcal{A}, \mu), E)$ be a measurable graph with $(V, \mathcal{A}, \mu)$ a finite countably generated measure space. If $\alpha(G) > 0$ and E is thick, then $\alpha(G) = \lim_r \vartheta(G, C_r(V)^*)$.*

Proof. That $\lim_r \vartheta(G, C_r(V)^*) \geq \alpha(G)$ follows from the discussion in § 4.1.

For the other inequality, take for all r a feasible solution $A_r \in L_{\text{sym}}^2(V)$ of $\vartheta(G, C_r(V)^*)$ with objective value at least $\alpha(G)/2$. Each A_r is positive semidefinite, so that $\|A_r\|_2 \leq \|A_r\|_{\mathcal{B}^1} = \text{Tr}(A_r) = 1$ for all r , hence the sequence $(A_r)_r$ lies in the unit ball of $L_{\text{sym}}^2(V)$. The unit ball in $L_{\text{sym}}^2(V)$ is compact since $L_{\text{sym}}^2(V)$ is reflexive [32, Theorem V.4.2]. By Lemma 2.6, $L_{\text{sym}}^2(V)$ is separable, and its unit ball is closed and bounded, so it is metrizable [32, Theorem V.5.1]. Thus, $(A_r)_r$ has a weakly converging subsequence; assume the sequence itself has weak limit A .

Each $C_r(V)^*$ is weakly closed, thus Theorem 2.10 says that $A \in \bigcap_r C_r(V)^*$ and $\bigcap_r C_r(V)^* = \text{CP}(V)$. In particular, since $\langle A, \mathbf{1}_E \rangle = \lim_r \langle A_r, \mathbf{1}_E \rangle = 0$ and $A \geq 0$, we can consider A to be the representant of its L^2 -equivalence class such that $A(x, y) = 0$ for all $xy \in E$.

Finally, A is nonzero: $\langle A, J \rangle = \lim_r \langle A_r, J \rangle = \alpha(G)/2 > 0$. Take the spectral decomposition $A = \sum_n \lambda_n f_n \otimes f_n$, then for all $N \in \mathbb{N}$

$$\sum_{n=1}^N \lambda_n = \langle A, \sum_{n=1}^N f_n \otimes f_n \rangle = \lim_{r \rightarrow \infty} \langle A_r, \sum_{n=1}^N f_n \otimes f_n \rangle \leq \lim_r \text{Tr}(A_r) = 1,$$

so $0 \neq \text{Tr}(A) \leq 1$. Therefore, $\text{Tr}(A)^{-1}A$ is a feasible solution of $\vartheta(G, \text{CP}(V))$, with $\langle \text{Tr}(A)^{-1}A, J \rangle \geq \lim_r \langle A_r, J \rangle$, as required. Conclude by Theorem 4.9. $\square$

Combining Theorem 5.3 above with Lemma 4.3, we obtain the following theorem which restricts the feasible region to only continuous kernels.

THEOREM 5.4. *Let $G = ((V, \mathcal{B}, \mu), E)$ be a measurable locally independent graph with V compact and second countable with μ a finite inner regular Borel measure. If $\alpha(G) > 0$ and E is thick, then*

$$\alpha(G) = \lim_{r \rightarrow \infty} \vartheta(G, C_r(V)_c^*).$$

Proof. This follows immediately from Theorem 5.3 and Lemma 4.3. $\square$

For nonthick edge sets it is difficult to make a general statement, even under the assumption that V is homogeneous under a compact subgroup of $\text{Aut}(G)$. Indeed, it seems that all we have at our disposal is weak compactness of the unit ball in $L^2_{\text{sym}}(V)$, but weak limits do not necessarily preserve zeros on sets that are not thick.

In the following case this can be salvaged through Schoenberg's theorem. For $D \subseteq [-1, 1]$, define $E(D) \subseteq \text{Sub}_{=2}(S^{n-1})$ by $xy \in E(D)$ if and only if $x^\top y \in D$. We always understand S^{n-1} to be equipped with the uniform surface measure ω , which is the quotient of a Haar measure of the orthogonal group $O(n)$ on $\mathbb{R}^n$. That this graph satisfies all conditions of Theorem 4.12 is explained in [36, §5] and references therein.

THEOREM 5.5. *If $n \geq 3$ is integer, $D \subseteq (-1, 1)$, and $G = (S^{n-1}, E(D), O(n))$ is a measurable homogeneous graph with the uniform surface measure ω such that $\alpha(G) > 0$, then*

$$\alpha(G) = \lim_r \vartheta(G, C_r(V)_c).$$

The following details are found in Andrews, Askey, and Roy [3]. Denote by ω_n the total measure $\omega(S^{n-1})$.

For $k \geq 0$, let $H_k^n \subseteq C(S^{n-1})$ be the space of n -variable spherical harmonics of degree k ; denote the dimension of H_k^n by h_k^n , and let $S_{k,1}^n, \dots, S_{k,h_k^n}^n$ be a complete orthogonal system of H_k^n . The set $\{S_{k,i}^n\}_{k,i}$ forms an orthonormal basis of $L^2(S^{n-1})$.

Let P_k^n be the Jacobi polynomial of degree k with the parameters $\alpha = \beta = (n-3)/2$ normalized by $P_k^n(1) = 1$. The addition formula [3, Theorem 9.6.3] states that

$$P_k^n(x^\top y) = \frac{\omega_n}{h_k^n} \sum_{i=1}^{h_k^n} S_{k,i}^n(x) S_{k,i}^n(y)$$

for all $x, y \in S^{n-1}$.

Write $E_k^n(x, y) = P_k^n(x^\top y)$. The kernels E_k^n are $O(n)$ -invariant and form a complete orthogonal system of the space $L^2_{\text{sym}}(S^{n-1})^{O(n)}$ of $O(n)$ -invariant kernels. We have $\langle E_k^n, E_l^n \rangle = 0$ if $k \neq l$ and $\langle E_k^n, E_k^n \rangle = \omega_n^2/h_k^n$. So if we write $A = \sum_k f(k)E_k^n$ and $B = \sum_k g(k)E_k^n$, then

$$\langle A, B \rangle = \sum_{k=0}^{\infty} f(k)g(k)\omega_n^2/h_k^n. \quad (25)$$

By the addition formula, each of the kernels E_k^n is positive semidefinite, hence $\sum_k f(k)E_k^n$ is positive semidefinite if and only if $f(k) \geq 0$ for all k . Schoenberg's theorem [93] states that if a kernel $K \in L_{\text{sym}}^2(S^{n-1})$ is continuous, $O(n)$ -invariant, and positive semidefinite, then there is a nonnegative sequence $f \in \ell^1$ such that

$$K(x, y) = \sum_{k \in \mathbb{N}} f(k) E_k^n(x, y) = \sum_{k \in \mathbb{N}} f(k) P_k^n(x^\top y)$$

with absolute and uniform convergence on $S^{n-1} \times S^{n-1}$.

Proof of Theorem 5.5. By Lemma 4.4, for each r , $\vartheta(G, C_r(S^{n-1})_c^*) \geq \alpha(G)$, hence $\lim_r \vartheta(G, C_r(S^{n-1})_c^*) \geq \alpha(G)$.

To prove the other inequality, see that $\text{Av}_{O(n)} C_r(S^{n-1})^* \subseteq C_r(S^{n-1})^*$ and that each feasible solution A of $\vartheta(G, C_r(S^{n-1})_c^*)$ is trace class, so by Theorem 3.3, $\text{Av}_{O(n)} A \in C_r(S^{n-1})_c^*$. Hence, we may restrict ourselves to $O(n)$ -invariant feasible solutions, which are automatically continuous. We show that $\lim_r \vartheta(G, C_r(S^{n-1})_c^*) \leq \vartheta(G, \text{CP}(S^{n-1})_c)$.

For every integer $r \geq 1$, let A_r be an $O(n)$ -invariant feasible solution of $\vartheta(G, C_r(S^{n-1})_c^*)$ with objective value at least $\alpha(G)/2 > 0$. Use Schoenberg's theorem to obtain a nonnegative sequence $f_r \in \ell^1$ such that $A_r = \sum_k f_r(k) E_k^n$. Then,

$$\omega_n \sum_{k \in \mathbb{N}} f_r(k) = \int_{S^{n-1}} A_r(x, x) d\omega(x) = 1. \quad (26)$$

Let c_0 be the space of all real-valued sequences that vanish at infinity. Under the supremum norm, c_0 is a Banach space and its dual is ℓ^1 by the duality $(f, g) = \sum_k f(k)g(k)$.

By (26) it follows that $\omega_n f_r$ is in the unit ball of ℓ^1 for all r . By Banach-Alaoglu [32, Theorem V.4.2], the unit ball of ℓ^1 is weak* compact, and the weak* topology on ℓ^1 is separable, thus the unit ball is weak* metrizable [32, Theorem V.5.1]. Therefore, the sequence $(f_r)_r$ has a weak*-converging subsequence; without loss of generality, assume the sequence itself converges with limit f .

It follows that f is nonnegative and $\omega_n \|f\|_1 \leq 1$. By Schoenberg's theorem, the kernel $A = \sum_k f(k) E_k^n$ is continuous, $O(n)$ -invariant and positive semidefinite. Moreover, if $B \in L_{\text{sym}}^2(S^{n-1})^{O(n)}$, then $B = \sum_k g(k) E_k^n$ for some real-valued sequence g , and since the sequence $(g(k) \omega_n^2 / h_k^n)_k$ vanishes at infinity, from (25) $\lim_r \langle A_r, B \rangle = \langle A, B \rangle$ follows.

Thus, $\langle A, J \rangle = \lim_r \langle A_r, J \rangle > 0$, so that A is nonzero and

$$0 < \tau = \int_V A(x, x) d\omega(x) = \omega_n \|f\|_1 \leq 1.$$

It is left to show that A is completely positive and vanishes on the edges, thus showing that $\tau^{-1}A$ is a feasible solution of $\vartheta(G, \text{CP}(V)_c)$ and that the inequality $\vartheta(G, \text{CP}(V)) \geq \lim_r \vartheta(G, C_r(V)_c^*)$ holds, finishing the proof by [36, Theorem 5.1].

Since the sequence A_r converges under the weak- L^2 topology to A , and each $C_r(V)^*$ is weakly closed, Theorem 2.10 states that $A \in \text{CP}(V)_c$. Finally, for $n \geq 3$ the asymptotic formula for the Jacobi polynomials [98, Theorem 8.21.8] implies that $(P_k^n(t))_k$ vanishes at infinity for all $t \in (-1, 1)$. So, for all $xy \in E(D)$ with $D \subseteq (-1, 1)$,

$$A(x, y) = \sum_{k \in \mathbb{N}} f(k) P_k^n(x^\top y) = \lim_{r \rightarrow \infty} \sum_{k \in \mathbb{N}} f_r(k) P_k^n(x^\top y) = 0,$$

which concludes the proof by Theorem 4.12. $\square$

4. Discussion and future directions

The results on convergence of the completely positive hierarchy are very incomplete. In the thick setting, it is clear that the square-integrable setting is insufficient for k -uniform hypergraphs with $k > 2$, as the normalization $\text{Tr} \mathcal{T}_{k-2}^* A = 1$ only bounds the $L^{k/(k-1)}$ -norm, and not the L^2 -norm. It is also clear that the space $L_{\text{sym}}^{k/(k-1)}(V, k)$ is too large; even for $k = 2$ we really only work in the trace class. Thus, a well-structured study of ϑ_{big} comprises a study of k -tensor analogues of the trace norm. To extend the convergence of the completely positive hierarchy to k -uniform homogeneous hypergraphs on the unit sphere, we have to look for a k -tensor version of Schoenberg's theorem. Tensor analogues of the necessary arguments are known; in particular, Castro-Silva [22] seems to have produced the important details. This is discussed in more detail in the concluding chapter of this thesis.

The most challenging problem in the homogeneous setting is extending the convergence result to other groups. There are several compact groups whose representation theory is similar to that of the sphere, so that the proof of Theorem 5.5 goes through with minimal changes. These are the continuous, compact, two-point homogeneous spaces with real dimension at least 2: the sphere, the real, complex and quaternionic projective spaces, and the octonionic projective plane. For these spaces, a theorem like Schoenberg's theorem is available [83, Theorem 3.1].

Extension beyond these spaces requires a different argument. For example, on the real circle S^1 the sequence of continuous, $O(2)$ -invariant, and positive-semidefinite kernels

$$((x, y) \mapsto e^{inx^\top y})_{n \in \mathbb{N}}$$

has weak* limit 0, but for orthogonal x and y the sequence $(e^{inx^\top y})_{n \in \mathbb{N}}$ is the constant-1 sequence. Since the proof of Theorem 5.5 hinges on weak* convergence, even though it is known that $\vartheta((S^1, E(0)), C_1(V)_c) = \alpha((S^1, E(0)))$, the method fails!

It is therefore necessary to study positive type functions of compact groups more systematically. This requires moving away from the generic setting of L^k functions. It seems that the Fourier algebra offers a suitable setting, as it is spanned by the positive type functions. Again, see the concluding chapter of this thesis for more details.

CHAPTER 6

Application: Witsenhausen’s problem

In this chapter, we apply completely positive-programming methods to find upper bounds on the maximum fraction of a unit sphere that can be covered by a set containing no orthogonal pairs—Problem I from the introduction. It is an example of an independent-set problem on a homogeneous graph under a compact but infinite group. We will see how the completely positive-programming bounds can be implemented using semidefinite programming. This offers the first bounds for this class of problems that improve on the linear programming bounds introduced by Bachoc, Nebe, Oliveira, and Vallentin [5], Oliveira [80], DeCorte, Oliveira, and Vallentin [36], and DeCorte [35], and results in the best bounds known. The results and exposition are taken from Bekker, Kuryatnikova, Oliveira, and Vera [12].

A subset of the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ *avoids orthogonal pairs* if it does not contain pairs of orthogonal vectors. *Witsenhausen’s problem* [102] asks for the maximum density that a measurable subset of S^{n-1} can have if it avoids orthogonal pairs. That is, we want to know the value of

$$\alpha_n = \{\omega(I)/\omega_n : I \subseteq S^{n-1} \text{ is measurable and avoids orthogonal pairs}\},$$

where ω denotes the standard surface measure of S^{n-1} , and $\omega_n = \omega(S^{n-1})$. We refer to the quantity $\omega(I)/\omega_n$ for a measurable $I \subseteq S^{n-1}$ as the *density* of I .

Fix $d \in S^{n-1}$. Witsenhausen [102] observed that the union of two open antipodal spherical caps of spherical radius $\pi/4$, i.e. the set

$$\{x \in S^{n-1} : |e^T x| > \cos \pi/4\},$$

avoids orthogonal pairs, hence α_n is at least the density of this set, which is $O(n^{-1/2}2^{-n/2})$. Kalai [52, Conjecture 2.8] conjectured that this construction is optimal, that is, that α_n is exactly the density of these two caps; this is known as the *double-cap conjecture*. A version of the double-cap conjecture for the complex unit sphere has an interpretation in quantum information theory [74].

The canonical basis vectors of $\mathbb{R}^n$ are n pairwise-orthogonal unit vectors. Any set that avoids orthogonal pairs can contain at most one of them. It then follows from a simple averaging argument that $\alpha_n \leq 1/n$. This upper bound was also given by Witsenhausen [102]; it is quite far from the lower bound of the double-cap conjecture for all $n \geq 3$. For $n = 2$, the lower and upper bounds coincide. Frankl and Wilson [46] were the first to give an asymptotic upper bound for α_n that decreases exponentially with the dimension n .

On the unit sphere, distance and inner product are related; a set of points on the sphere avoids orthogonal pairs if it avoids pairs of points at distance $\sqrt{2}$. More generally, let V be metric space with metric d and let $D \subseteq (0, \infty)$ be a set of *forbidden* distances. We say that a set $I \subseteq V$ *avoids* the distances in D or that it is a *D-avoiding set* if $d(x, y) \notin D$ for all $x, y \in I$. In these terms, Witsenhausen's problem asks for the maximum density of a $\sqrt{2}$ -avoiding set on the sphere equipped with the Euclidean distance.

Distance-avoiding sets can be modeled as independent sets of graphs. Given a metric space V with metric d and a set D of forbidden distances, let G be the graph with vertex set V in which $x, y \in V$ are adjacent if and only if $d(x, y) \in D$; call such a graph G a *distance graph*. The independent sets of G are exactly the D -avoiding sets.

Denote the distance graph with vertex set S^{n-1} and forbidden distance $\sqrt{2}$ by G_n . Witsenhausen's problem can be seen as an independent-set problem on G_n with the measure ω/ω_n , i.e. $\alpha_n = \alpha(G_n)$. We will use the hierarchies derived in Chapter 5 to bound α_n from above.

There is a special role for the completely positive hierarchy: although Theorem 5.2 states that $\text{blockM}_3(G_n)$ should give a bound at least as good as $\vartheta(G_n, C_1(V, 2)_c^*)$, the usual way of implementing this three-point bound as a polynomial optimization problem does not result in a rigorous upper bound. However, a particular restriction of the completely positive hierarchies does, and this leads to the first use of optimization hierarchies for Witsenhausen's problem, giving the best upper bounds known.

1. Invariant positive-semidefinite kernels

The orthogonal group on $\mathbb{R}^n$ is denoted $O(n)$. Witsenhausen's problem is invariant under the action of $O(n)$, in the sense that it is defined on a graph that is vertex transitive under the action of $O(n)$. We exploit this to significantly reduce the size of the optimization problems. The thesis of De Muinck Keizer [77] gives a more structured and detailed account of the contents of this section. For us, though, a straight-forward ad-hoc approach suffices.

We have already seen that averaging a feasible solution of a program in the block moment hierarchy and the completely positive hierarchy preserves feasibility and the objective value. We may thus assume all tensors to be invariant under $O(n)$. As we will see, we can describe $O(n)$ -invariant 3-tensors by $O(n)$ - and $\text{Stab}(e)$ -invariant kernels, where $e \in S^{n-1}$ is an arbitrary point.

Positive-semidefinite kernels on S^{n-1} invariant under $O(n)$ are easily described by Schoenberg's theorem [93]; also see the discussion in §5.3. Let $n \geq 1$ and $k \geq 0$. If P_k^n is the Jacobi polynomial of degree k with parameters $\alpha = \beta = (n-3)/2$ normalized by $P_k^n(1) = 1$, write $E_k^n(x, y) = P_k^n(x^\top y)$. Schoenberg's theorem [93] states that a kernel $K \in L_{\text{sym}}^2(V, 2)$ is continuous, $O(n)$ -invariant, and positive semidefinite, if and only if there is a nonnegative sequence $f \in \ell^1$ such that

$$K(x, y) = \sum_{k \in \mathbb{N}} f(k) E_k^n(x, y) = \sum_{k \in \mathbb{N}} f(k) P_k^n(x^\top y) \quad (27)$$

with absolute and uniform convergence on $S^{n-1} \times S^{n-1}$.

Fix $e \in S^{n-1}$ and let $\text{Stab}(e)$ be its stabilizer under the action of $O(n)$; see §3.1 for more. The kernels invariant under this action were described by Musin [78] and Bachoc and Vallentin [6] as follows.

With P_k^n as in (27), for integers $n \geq 2$ and $k \geq 0$ consider the polynomial

$$Q_k^n(u, v, t) = (1 - u^2)^{k/2} (1 - v^2)^{k/2} P_k^n \left(\frac{t - uv}{(1 - u^2)^{1/2} (1 - v^2)^{1/2}} \right)$$

and let $Y_{k,d}^n$ be the $(d - k + 1) \times (d - k + 1)$ matrix given by

$$(Y_{k,d}^n)_{i,j} = u^i v^j Q_k^{n-1}(u, v, t) \quad (28)$$

for $0 \leq i, j \leq d - k$. If $K \in C((S^{n-1})^2)$ is $\text{Stab}(e)$ -invariant, then $K(x, y)$ depends only on the inner products $e^\top x$, $e^\top y$, and $x^\top y$. Bachoc and Vallentin showed that, for any d and any choice of positive semidefinite matrices $F_k \in$

$\mathbb{R}^{(d-k+1) \times (d-k+1)}$, the kernel

$$K(x, y) = \sum_{k=0}^d \langle F_k, Y_{k,d}^n(e^\top x, e^\top y, x^\top y) \rangle \quad (29)$$

is $\text{Stab}(e)$ -invariant and positive semidefinite. It is continuous by construction, since it is a polynomial on the three inner products.

Although the above decomposition looks similar to that in Schoenberg's theorem, there is no guarantee that every kernel K is a pointwise converging sum of the form (29) with $d = \infty$ and the F_k infinite matrices. However, we may uniformly approximate a continuous, positive-semidefinite, and $\text{Stab}(e)$ -invariant kernel by kernels of the form (29).

2. The failure of the block moment hierarchy

The three-point bound $\text{blockM}_3(G_n)$ fails to give an implementable bound. For simplicity, we use an alternative normalization, i.e.

$$\begin{aligned} & \sup \int_{\text{Sub}=2(S^{n-1})} \phi(S) d\omega_{\text{sub}}^2(S) \\ & \int_{\text{Sub}=1(S^{n-1})} \phi(S) d\omega_{\text{sub}}^1(S) = 1, \\ & \phi(S) = 0 \quad \text{if } S \in \text{Sub}_3(S^{n-1}) \text{ is not independent,} \\ & M_{\{u\}}\phi \succeq 0 \quad \text{for all } u \in S^{n-1}, \\ & M_\emptyset \phi \succeq 0, \\ & \phi \in C(\text{Sub}_3(S^{n-1})). \end{aligned}$$

In the proof of Theorem 5.2 we showed that the matrix

$$\begin{pmatrix} \phi(\emptyset) & \int_{\text{Sub}=1(S^{n-1})} \phi(S) d\omega_{\text{sub}}^1(S) \\ \int_{\text{Sub}=1(S^{n-1})} \phi(S) d\omega_{\text{sub}}^1(S) & \int_{\text{Sub}=2(S^{n-1})} \phi(S) d\omega_{\text{sub}}^2(S) \end{pmatrix}$$

is positive semidefinite, which implies that $\text{blockM}_3(G_n)$ is bounded from above by this program. We further relax the bound by changing the operators M_Q for $|Q| \leq 1$ so that it sends ϕ to the kernel $K \in C((S^{n-1})^2)$ such that $K(x, y) = \phi(Q \cup \{x, y\})$, i.e., we disregard the empty set.

We may restrict to $O(n)$ -invariant functions ϕ by applying an averaging operator, because $\text{Sub}_3(V)$ comes with a natural group action and an invariant measure.

Let $e \in S^{n-1}$. Given any $u \in S^{n-1}$, there is an orthogonal transformation T such that $Tu = e$, hence if ϕ is an $O(n)$ -invariant feasible solution, then

$$(M_{\{u\}}\phi)(x, y) = \phi(\{u, x, y\}) = \phi(\{e, Tx, Ty\}) = (M_{\{e\}}\phi)(Tx, Ty).$$

It follows that if $M_{\{e\}}\phi$ is positive semidefinite, so is $M_{\{u\}}\phi$ for every $u \in S^{n-1}$. Since ϕ is $O(n)$ -invariant, $M_{\{e\}}\phi$ is $\text{Stab}(e)$ -invariant. This allows us to rewrite the problem by considering two kernels $A = M_{\emptyset}\phi$ and $K = M_{\{e\}}\phi$:

$$\begin{aligned}
& \sup \int_{\text{Sub}=2(S^{n-1})} \phi(S) d\omega_{\text{sub}}^2(S) \\
& \int_{\text{Sub}=1(S^{n-1})} \phi(S) d\omega_{\text{sub}}^1(S) = 1, \\
& A(x, y) = K(e, Ty) \quad \text{for all } x, y \in S^{n-1} \\
& \quad \quad \quad \text{and } T \in O(n) \text{ with } Tx = e, \\
& K(e, x) = K(x, x) \quad \text{for all } x \in S^{n-1}, \\
& K(x, y) = 0 \quad \text{if } \{e, x, y\} \text{ is not independent,} \\
& A \in C_{\text{sym}}((S^{n-1}))_{\geq 0} \text{ is } O(n)\text{-invariant,} \\
& K \in C_{\text{sym}}((S^{n-1}))_{\geq 0} \text{ is } \text{Stab}(e)\text{-invariant.}
\end{aligned} \tag{30}$$

Since A is $O(n)$ -invariant, Schoenberg's theorem can be used to express A in terms of Jacobi polynomials as in (27). The kernel K is invariant under $\text{Stab}(e)$, so the expansion (29) parametrizes a large class of the required kernels.

However, as pointed out, even though the kernels of the form (29) approximate the $\text{Stab}(e)$ -invariant kernels uniformly, a pointwise converging sum of the same form with $d = \infty$ is in general not guaranteed. The constraint “ $K(x, y) = 0$ if $\{e, x, y\}$ is not independent” can therefore not be written equivalently in terms of such an expansion, hence it is unclear that the resulting problem would give an upper bound to α_n .

Even when we replace this constraint by a relaxation, for example requiring that $K(x, y) \in [-\epsilon, \epsilon]$ for some fixed $\epsilon > 0$, it remains difficult to get a rigorous upper bound on α_n .

Indeed, to solve the modified problem (30) we have to fix the degrees of the polynomials at some point. To get an upper bound, we have to solve a problem of this form to optimality. To do so rigorously we have to use polynomials of high degree, and since K is parametrized by 3-variable polynomials, the variable matrices become prohibitively large.

3. The fix: another hierarchy

Recall the definition of slice positive tensors from Section 4.1. If V is a compact Hausdorff space equipped with a Radon measure with full support, a function $F \in C(V^r)$ with $r \geq 2$ integer is slice positive if and only if for all $v \in (S^{n-1})^{r-2}$ the kernel $(x, y) \mapsto F(x, y, v)$ is a positive-semidefinite kernel.

For an integer $r \geq 1$, let

$$\mathcal{Q}_r(V, k) = \{ A \in L_{\text{sym}}^2(V, k) : \text{Av}_{\mathfrak{S}_{k+r}}(A \otimes \mathbf{1}^{\otimes r} - F) \geq 0 \\ \text{for some slice-positive } F \in L^2(V^{k+r}) \}.$$

Immediately we see that $\mathcal{Q}_r(V, k) \subseteq \mathcal{C}_r(V, k)$ for all k and r . Moreover, $\mathcal{Q}_r(V, k) \subseteq \text{COP}(V, k)$ for all r and k . Indeed, take $A \in \mathcal{Q}_r(V, k)$ and $F \in L^2(V^{k+r})$ with $\text{Av}_{\mathfrak{S}_{k+r}}(A \otimes \mathbf{1}^{\otimes r} - F) \geq 0$. Given a nonnegative $f \in L^2(V)$ with $\langle \mathbf{1}, f \rangle \geq 0$ we have

$$\begin{aligned} 0 &\leq \langle \text{Av}_{\mathfrak{S}_{k+r}}(A \otimes \mathbf{1}^{\otimes r} - F), f^{\otimes(k+r)} \rangle \\ &= \langle A \otimes \mathbf{1}^{\otimes r}, \text{Av}_{\mathfrak{S}_{k+r}} f^{\otimes(k+r)} \rangle - \langle F, \text{Av}_{\mathfrak{S}_{k+r}} f^{\otimes(k+r)} \rangle \\ &= \langle A, f^{\otimes k} \rangle \langle \mathbf{1}, f \rangle^r - \langle F, f^{\otimes(k+r)} \rangle. \end{aligned}$$

Since F is slice positive,

$$\begin{aligned} \langle F, f^{\otimes(k+r)} \rangle &= \int_{V^{k+r-2}} \int_{V^2} F(v, w) f^{\otimes 2}(v) d\omega^2(v) f^{\otimes k+r-2}(w) d\omega^{k+r-2}(w) \\ &\geq 0, \end{aligned}$$

and we see that $\langle A, f^{\otimes k} \rangle \geq 0$, so A is copositive.

One shows, as for example in the proof of [12, Theorem 4.1] that

$$\mathcal{Q}_1(V, k) \subseteq \mathcal{Q}_2(V, k) \subseteq \cdots \subseteq \text{COP}(V, k),$$

is a hierarchy of inner approximation of $\text{COP}(V, k)$ stronger than the $\mathcal{C}_r(V, k)$ hierarchy; it was proposed by Peña, Vera, and Zuluaga [85] and extended to the infinite-dimensional setting by Kuryatnikova and Vera [56].

Given a graph $G = (V, E)$ such that V is compact and Hausdorff and equipped with a Radon measure, write $\mathcal{Q}_r(V) = \mathcal{Q}_r(V, 2)$, and consider the programs $\vartheta(G, \mathcal{Q}_r(V)_c^*)$. Under the conditions of Chapters 4 and 5, this gives a hierarchy of bounds for the measurable independence number, namely

$$\vartheta(G, \mathcal{Q}_1(V)_c^*) \geq \vartheta(G, \mathcal{Q}_2(V)_c^*) \geq \cdots \geq \alpha(G),$$

that is at least as strong as the hierarchy $\vartheta(G, C_r(V, 2)_c^*)$. In particular, the convergence results from Chapter 5 hold.

We will implement a version of the bound $\vartheta(G_n, \mathcal{Q}_1(V)^*)$. For this, we first have to figure out how to describe slice-positive 3-tensors effectively.

Let V be a Γ -space with Γ a compact group. Let p be the quotient map onto the set of orbits $p : V^k \rightarrow V^k/\Gamma$ and $R : V^k/\Gamma \rightarrow V^k$ be a section of p .

Suppose $F \in C(V^{k+2})$ is slice-positive and Γ -invariant. Consider the function $K : (V^k/\Gamma) \times V^2 \rightarrow \mathbb{R}$ such that

$$K(\xi, x, y) = F(x, y, R(\xi))$$

and for every orbit ξ let $K_\xi(x, y) = K(\xi, x, y)$; note that K depends on the choice of R . The kernel K_ξ is continuous and positive semidefinite for every ξ . Moreover, since F is Γ -invariant, K_ξ is $\text{Stab}(R(\xi))$ -invariant. If we equip V^k/Γ with the quotient topology, K is continuous if R is.

Conversely, say $K : (V^k/\Gamma) \times V^2 \rightarrow \mathbb{R}$ is a continuous function such that K_ξ is a positive-semidefinite $\text{Stab}(R(\xi))$ -invariant kernel for all ξ . Then we may define a function $F : V^{k+2} \rightarrow \mathbb{R}$ by

$$F(x, y, v) = K(p(v), \sigma x, \sigma y)$$

for all $\sigma \in \Gamma$ such that $\sigma v = R(v)$. Such F is well-defined: if $\tau v = R(\xi)$, then $\sigma\tau^{-1}R(\xi) = R(\xi)$, so from the $\text{Stab}(R(\xi))$ -invariance of K_ξ we obtain $K(\xi, \tau x, \tau y) = K(\xi, \sigma x, \sigma y)$.

By construction, F is slice-positive and Γ -invariant. To see the latter, given $\sigma \in \Gamma$, let $\tau \in \Gamma$ be such that $\tau\sigma v = R(p(v))$. Then,

$$F(\sigma x, \sigma y, \sigma v) = K(p(v), \tau\sigma x, \tau\sigma y) = F(x, y, v).$$

If there exists a continuous function $s : V^k \rightarrow \Gamma$ such that $s(v)v = R(p(v))$ for all $v \in V^k$, then F is continuous.

We will now see how $\vartheta(G_n, \mathcal{Q}_1(S^{n-1})_c^*)$ behaves better than $\text{blockM}_3(G_n)$ with respect to approximation by the kernels from (29); indeed, we will see that restricting to kernels of this form gives a subset of $\mathcal{Q}_1(S^{n-1})$ and therefore a superset of $\mathcal{Q}_1(S^{n-1})^*$, thus relaxing the program.

If $Z \in \mathcal{Q}_1(S^{n-1})$, then there is a continuous slice positive $F : (S^{n-1})^3 \rightarrow \mathbb{R}$ such that $\text{Av}_{\mathfrak{S}_3}(Z \otimes \mathbf{1} - F) \geq 0$. If Z is $O(n)$ -invariant, we can assume that F is $O(n)$ -invariant as well, otherwise we simply take $\text{Av}_{O(n)}F$, which is continuous and slice positive.

There is only one orbit for the action of $O(n)$ on S^{n-1} ; pick $e \in S^{n-1}$ as its representative. The invariant function F is continuous and slice positive if and only if there is a continuous, positive-semidefinite, and $\text{Stab}(e)$ -invariant kernel $K : (S^{n-1})^2 \rightarrow \mathbb{R}$ such that $F(x, y, z) = K(Tx, Ty)$, where T is any orthogonal matrix such that $Tz = e$. So the value of $F(x, y, z)$ depends only on $e^\top Tx = x^\top z$, $e^\top Ty = y^\top z$, and $(Tx)^\top Ty = x^\top y$.

The kernels (29) are positive semidefinite and $\text{Stab}(e)$ -invariant. Fix an integer $d \geq 1$. Say Z is the $O(n)$ -invariant kernel given by

$$Z(x, y) = \sum_{k=0}^{2d} f(k) P_k^n(x^\top y). \quad (31)$$

Let $\bar{Y}_{k,d}^n = \text{Av}_{\mathfrak{S}_3} Y_{k,d}^n$ be the matrix obtained from $Y_{k,d}^n$ of (28) by averaging over all permutations of (u, v, t) .

If there are positive-semidefinite matrices $F \in \mathbb{R}^{(d-k+1) \times (d-k+1)}$ for $k = 0, \dots, d$, such that

$$\sum_{k=0}^{2d} f(k) (1/3) (P_k^n(u) + P_k^n(v) + P_k^n(t)) - \sum_{k=0}^d \langle F_k, \bar{Y}_{k,d}^n(u, v, t) \rangle \geq 0 \quad (32)$$

for all $(u, v, t) \in \Delta = \{ (x^\top z, y^\top z, x^\top y) : x, y, z \in S^{n-1} \}$, then $Z \in \mathcal{Q}_1(S^{n-1})$.

Indeed, for $x, y, z \in S^{n-1}$ with $u = x^\top z$, $v = y^\top z$, and $t = x^\top y$ we have

$$\text{Av}_{\mathfrak{S}_r}(Z \otimes 1)(x, y, z) = \sum_{k=0}^{2d} f(k) (1/3) (P_k^n(u) + P_k^n(v) + P_k^n(t)).$$

The function F given by

$$F(x, y, z) = \sum_{k=0}^d \langle F_k, Y_{k,d}^n(u, v, t) \rangle$$

is slice positive and continuous and

$$(\text{Av}_{\mathfrak{S}_3} F)(x, y, z) = \sum_{k=0}^d \langle F_k, \bar{Y}_{k,d}^n(u, v, t) \rangle.$$

Putting it all together, $Z \in \mathcal{Q}_1(S^{n-1})$.

The left-hand side of (32) is a polynomial $p \in \mathbb{R}[u, v, t]$ of degree at most $2d$ that should be nonnegative on Δ . The polynomial p is invariant under the permutation action of $\mathfrak{S}_3$ on the variables. The domain Δ is also invariant under $\mathfrak{S}_3$; it is a semi-algebraic set:

$$\Delta = \{ (u, v, t) : g_i(u, v, t) \geq 0 \text{ for } i = 1, \dots, 4 \},$$

where

$$\begin{aligned} g_1 &= g(u) + g(v) + g(t), & g_2 &= g(u)g(v) + g(u)g(t) + g(v)g(t), \\ g_3 &= g(u)g(v)g(t), & g_4 &= 1 + 2uv t - u^2 - v^2 - t^2, \end{aligned}$$

where $g(w) = 1 - w^2$. So, if there are sums-of-squares polynomials $q_0, \dots, q_4$ in $\mathbb{R}[u, v, t]$ such that

$$p = q_0 + g_1 q_1 + g_2 q_2 + g_3 q_3 + g_4 q_4, \quad (33)$$

then p is nonnegative on Δ . Moreover, since p and the g_i are all invariant under $\mathfrak{S}_3$, we may assume without loss of generality that the q_i are also invariant.

Let V_r be the matrix indexed by the monomials on u, v , and t of degree at most $\lfloor r/2 \rfloor$ such that $V_r(m_1, m_2) = m_1 m_2$ for any two such monomials. Every entry of V_r is a polynomial of degree at most r . A polynomial q of degree $2k$ is a sum of squares if and only if there is a positive-semidefinite matrix Q such that $q = \langle Q, V_{2k} \rangle$.

Using this equivalence and restricting the degrees of the polynomials q_i appearing in (33), we can write a sufficient condition for p to be nonnegative on Δ in terms of positive-semidefinite matrices. Namely, if there are positive-semidefinite matrices F_k and Q_i such that

$$\begin{aligned} \sum_{k=0}^{2d} f(k)(1/3)(P_k^n(u) + P_k^n(v) + P_k^n(t)) - \sum_{k=0}^d \langle F_k, \bar{Y}_{k,d}^n(u, v, t) \rangle \\ = \langle Q_0, V_{2d} \rangle + \langle Q_1, g_1 V_{2d-2} \rangle + \langle Q_2, g_2 V_{2d-4} \rangle \\ + \langle Q_3, g_3 V_{2d-6} \rangle + \langle Q_4, g_4 V_{2d-8} \rangle, \end{aligned} \quad (34)$$

then Z given in (31) is in $\mathcal{Q}_1(S^{n-1})$. This leads us to the definition of the following cone for every fixed d :

$$\mathcal{Q}_1^d = \{ (f(0), \dots, f(2d), 0, \dots) \in \mathbb{R}^{\mathbb{N}} : \text{there are positive-semidefinite matrices } F_k \text{ and } Q_i \text{ such that (34) holds} \}$$

We were careful to describe the domain Δ with invariant polynomials so that we could assume that all polynomials q_i are likewise invariant. This can be used to simplify (34), so that we can work with block-diagonal positive-semidefinite matrices Q_i . The original idea was presented by Gatermann and Parrilo [48]; see also Machado and Oliveira [71] and Leijenhorst and De Laat [59, §4] for more recent descriptions of the method and an application to this exact situation. This use of symmetry to reduce the problem's size is essential to reach high degrees.

4. Implementation and verification of the bound

To make our bound on α_n as good as possible, we combine the cone $\mathcal{Q}_1^d$ with constraints from the Boolean quadratic polytope, which for a finite set V

is defined as

$$\text{BQP}(V) = \text{conv}\{xx^\top : x \in \{0, 1\}^V\}.$$

Such constraints were used before by DeCorte, Oliveira, and Vallentin [36].

Given a measurable independent set $I \subseteq S^{n-1}$ of G_n , define the kernel $A = \text{Av}_{\text{O}(n)}(\mathbf{1}_I \otimes \mathbf{1}_I)$. Then,

- (i) A is an $\text{O}(n)$ -invariant continuous kernel (by Lemma 3.4),
- (ii) $A(x, y) = 0$ for all orthogonal $x, y \in S^{n-1}$,
- (iii) A is positive semidefinite and $A \in \mathcal{Q}_r(S^{n-1})^*$ for all $r \geq 1$,
- (iv) $(A(x, y))_{x, y \in U} \in \text{BQP}(U)$ for every finite $U \subseteq S^{n-1}$, and
- (v) $\int_{S^{n-1}} A(x, x) d\omega(x) = \omega(I)$ and $\langle A, J \rangle = \omega(I)^2$.

We use Schoenberg's theorem to express A in terms of Jacobi polynomials as in (27), so

$$A(x, y) = \sum_{k=0}^{\infty} a(k) P_k^n(x^\top y)$$

for some sequence $a \geq 0$. Recall that we normalize the polynomials so $P_k^n(1) = 1$; together with the addition formula, see 3, this gives

$$\int_{S^{n-1}} A(x, x) d\omega(x) = \omega_n \sum_{k=0}^{\infty} a(k) \quad \text{and} \quad \langle A, J \rangle = \omega_n^2 a(0), \quad (35)$$

where $\omega_n = \omega(S^{n-1})$.

Let $U \subseteq S^{n-1}$ be a finite set and let $L \in \mathbb{R}^{U \times U}$ and $\beta \in \mathbb{R}$ such that $\langle L, X \rangle \leq \beta$ for all $X \in \text{BQP}(U)$. Then, defining $r : \mathbb{N} \rightarrow \mathbb{R}$ by

$$r(k) = \sum_{x, y \in U} L(x, y) P_k^n(x^\top y) \quad (36)$$

we have

$$\sum_{k=0}^{\infty} a(k) r(k) \leq \beta. \quad (37)$$

We call (r, β) a $\text{BQP}(S^{n-1})$ -inequality, and we call the points in U the *support points* of the inequality.

Finally, if $(f(0), \dots, f(2d), 0, \dots) \in \mathcal{Q}_1^d$ and $Z(x, y) = \sum_{k=0}^{2d} f(k) P_k^n(x^\top y)$, then $Z \in \mathcal{Q}_1(S^{n-1})$, and from (25) we get

$$\sum_{k=0}^{2d} (a(k)/h_k^n) f(k) = \omega_n^{-2} \langle A, Z \rangle \geq 0,$$

that is, $k \mapsto a(k)/h_k^n$ belongs to $(\mathcal{Q}_1^d)^*$.

Let $(r_1, \beta_1), \dots, (r_N, \beta_N)$ be any $\text{BQP}(S^{n-1})$ -inequalities and fix some integer $d \geq 1$. Put together, our developments lead us to the following optimization problem, whose optimal value gives an upper bound on α_n :

$$\begin{aligned} \sup \quad & \sum_{k=0}^{\infty} a(k) \\ & \sum_{k=0}^{\infty} a(k) P_k^n(0) = 0, \\ & \sum_{k=0}^{\infty} a(k) r_i(k) \leq \beta_i \quad \text{for } 1 \leq i \leq N, \\ & \begin{pmatrix} 1 & \omega_n \sum_{k=0}^{\infty} a(k) \\ \omega_n \sum_{k=0}^{\infty} a(k) & \omega_n^2 a(0) \end{pmatrix} \text{ is positive semidefinite,} \\ & a \geq 0 \text{ and } k \mapsto a(k)/h_k^n \in (\mathcal{Q}_1^d)^*. \end{aligned} \quad (38)$$

The 2×2 matrix comes from (35) and (v) and is used to normalize the problem. The objective function is divided by ω_n , ensuring that we get a bound for $\alpha_n = \alpha(G_n)$ under the measure ω/ω_n . Finally, our problem has infinitely many variables a , but only the first $2d+1$ of them appear in the cone constraint with $(\mathcal{Q}_1^d)^*$. Contrast this with the situation of the block moment hierarchy from §2.

The dual of this problem is

$$\begin{aligned} \inf \quad & z_{11} + \sum_{i=1}^N y_i \beta_i \\ & \lambda + \sum_{i=1}^N y_i r_i(0) - \omega_n z_{12} - \omega_n^2 z_{22} - f(0) \geq 1, \\ & \lambda P_k^n(0) + \sum_{i=1}^N y_i r_i(k) - \omega_n z_{12} - f(k) \geq 1 \quad \text{for all } 1 \leq k \leq 2d, \\ & \lambda P_k^n(0) + \sum_{i=1}^N y_i r_i(k) - \omega_n z_{12} \geq 1 \quad \text{for all } k \geq 2d+1, \\ & \begin{pmatrix} z_{11} & z_{12}/2 \\ z_{12}/2 & z_{22} \end{pmatrix} \text{ is positive semidefinite,} \\ & y \geq 0 \text{ and } k \mapsto h_k^n f(k) \in \mathcal{Q}_1^d. \end{aligned} \quad (39)$$

The objective value of any feasible solution of (39) is greater than or equal to the objective value of any feasible solution of (38). So, any feasible solution of (39) gives an upper bound on α_n . Furthermore, the constraint given by “ $k \mapsto h_k^n f(k) \in \mathcal{Q}_1^d$ ” is expressed in terms of (34), namely we require there to be positive-semidefinite matrices F_k and Q_i such that

$$\begin{aligned} \sum_{k=0}^{2d} f(k) (h_k^n/3) (P_k^n(u) + P_k^n(v) + P_k^n(t)) - \sum_{k=0}^d \langle F_k, \bar{Y}_{k,d}^n(u, v, t) \rangle \\ - \langle Q_0, V_{2d} \rangle - \langle Q_1, g_1 V_{2d-2} \rangle - \langle Q_2, g_2 V_{2d-4} \rangle \\ - \langle Q_3, g_3 V_{2d-6} \rangle - \langle Q_4, g_4 V_{2d-8} \rangle = 0. \end{aligned} \quad (40)$$

So (39) is a semidefinite programming problem with finitely many variables but infinitely many constraints.

To solve (39) we use the package `ClusteredLowRankSolver.jl` of Leijenhorst and De Laat [59]; the input for the solver is generated by a Julia program. The program and all data files used are available in the Harvard Dataverse repository [81].

To find good $\text{BQP}(S^{n-1})$ -inequalities, we use a separation heuristic described by DeCorte, Oliveira, and Vallentin [36]. The inequalities used are also included in the repository and need not be recomputed.

Table 1 contains a detailed account of all the bounds computed from (39). Solving the problem for $d = 14$ and 18 takes time and memory, so files with the corresponding solutions are also available in the repository.

Since (39) has infinitely many linear constraints, to solve it, we select some finite set $S \subseteq \{0, 1, \dots\}$ and consider only the constraints for $k \in S$. After a solution is found it has to be verified, that is, we need to check that all constraints are indeed satisfied.

Let (λ, y, z, f, F, Q) be a candidate solution to (39), where F and Q are as in (40), returned by the solver. The first step is to certify ourselves that f , F , and Q indeed satisfy (40).

This is certainly not true: the solver uses floating-point arithmetic, so (40) will not hold. Rather, the left-hand side of (40) will be a polynomial with coefficients close to 0. Since the `ClusteredLowRankSolver.jl` uses high-precision floating-point arithmetic, the coefficients will be quite small; let η be the largest absolute value of any such coefficient.

It is always possible to perturb the matrices Q_i in order to satisfy the constraint; the order of the perturbation depends on η . We want to do so and keep the Q_i positive semidefinite; as long as the minimum eigenvalues of the matrices Q_i are large enough compared to η , this is always possible. The solutions stored in the repository have large minimum eigenvalues, several orders of magnitude larger than η , so this perturbation of the Q_i can always be carried out. We do not have to actually change the Q_i ; it suffices to know that such a perturbation is possible, since then we know we can get a feasible solution if we want to. This procedure was used before by De Laat, Oliveira, and Vallentin [62].

Checking that the linear constraints for all $k \geq 0$ are satisfied is more difficult; we use the approach outlined in DeCorte, Oliveira, and Vallentin [36].

The idea is as follows. Let $\text{lhs}(k)$ be the left side of the k th linear constraint in (39) and write $\text{lhs}(\infty) = \lim_{k \rightarrow \infty} \text{lhs}(k)$; we will see that this limit exists. We then take the following steps.

n	Lower bound	Old upper bound		d	New upper bound	
		Simple	Best		No BQP	With BQP
3	0.2928...	0.3333...	0.30153	6	0.316925	0.300708
				10	0.309298	0.298998
				14	0.305627	0.298341
				18	0.303294	0.297742
4	0.1816...	0.25	0.21676	6	0.223633	0.207617
				10	0.211825	0.199402
				14	0.205479	0.196162
				18	0.201445	0.194297
5	0.1161...	0.2	0.16765	6	0.167357	0.151541
				10	0.153819	0.141539
				14	0.146612	0.137142
				18	0.142349	0.134588
6	0.0755...	0.1666...	0.13382	6	0.130829	0.116599
				10	0.116509	0.105200
				14	0.109989	0.100374
				18	0.106727	0.098095
7	0.0498...	0.1428...	0.11739	6	0.106059	0.093031
				10	0.091477	0.081221
				14	0.086656	0.077278
				18	0.084787	0.075751
8	0.0331...	0.125	0.09981	6	0.088750	0.076801
				10	0.074309	0.064919
				14	0.071676	0.063287
				18	0.070607	0.061178

TABLE 1. Low and upper bounds for Witsenhausen's parameter α_n . The lower bound is given by the double-cap conjecture. The simple upper bound was given by Witsenhausen [102] and is just $1/n$. The best previous upper bounds are by DeCorte, Oliveira, and Vallentin [36]. The table gives upper bounds obtained from solving (39) with and without BQP(S^{n-1})-inequalities and for several values of the maximum degree d .

- (i) As long as $z_{22} > 0$, we can change z_{11} and z_{12} to get $\text{lhs}(\infty) \geq 1 + \eta$ for some $\eta > 0$. The more we change z_{12} , the more we have to change z_{11} , and the worse the bound gets.
- (ii) Next, for some $\epsilon < \eta$ we find a k_0 such that $|\text{lhs}(k) - \text{lhs}(\infty)| \leq \epsilon$ for all $k \geq k_0$. Then $\text{lhs}(k) \geq \text{lhs}(\infty) - \epsilon \geq 1 + \eta - \epsilon > 1$, and so all constraints are satisfied for $k \geq k_0$.
- (iii) Finally, we check the constraints for $k = 0, \dots, k_0$, and by changing z again we can make all these constraints satisfied.

If we choose our initial sample S well, then all constraints will be almost satisfied, and we will not have to change z too much in order to get a feasible solution. This is the procedure implemented by the `fix_linear_constraints` function in the Julia program in the repository [81].

Let us see the details of the procedure. The asymptotic formula for the Jacobi polynomials [98, Theorem 8.21.8] implies that $P_k^n(t) \rightarrow 0$ as $k \rightarrow \infty$ for all $t \in (-1, 1)$. We make sure that all the $\text{BQP}(S^{n-1})$ -inequalities (37) we use have support points U such that distinct $x, y \in U$ have inner product $x^\top y$ bounded away from ± 1 . So if r is given as in (36), then

$$r(\infty) = \lim_{k \rightarrow \infty} r(k) = \text{Tr} L$$

and

$$|r(k) - r(\infty)| \leq \sum_{\substack{x, y \in U \\ x \neq y}} |L(x, y)| |P_k^n(x^\top y)|. \quad (41)$$

We also have

$$\text{lhs}(\infty) = \sum_{i=1}^N y_i r_i(\infty) - \omega_n z_{12}.$$

Given $\epsilon > 0$ we want to get k_0 as in (ii). Note that

$$|\text{lhs}(k) - \text{lhs}(\infty)| \leq |\lambda| |P_k^n(0)| + \sum_{i=1}^N y_i |r_i(k) - r_i(\infty)|.$$

Fix k_0 . Using (41), we see that to find an upper bound for the left side above for all $k \geq k_0$, it suffices to find for all $k \geq k_0$ an upper bound on $|P_k^n(t)|$ for $t = 0$ and all other $t \in (-1, 1)$ that occur as inner products between distinct support points of the $\text{BQP}(S^{n-1})$ -inequalities we use.

To do so rigorously, we use an integral representation for the ultraspherical polynomials due to Gegenbauer (take $\lambda = (n - 2)/2$ in Theorem 6.7.4 from Andrews, Askey, and Roy [3]):

$$P_k^n(\cos \theta) = R(n)^{-1} \int_0^\pi F(\phi)^k \sin^{n-3} \phi \, d\phi,$$

where

$$F(\phi) = \cos \theta + i \sin \theta \cos \phi \quad \text{and} \quad R(n) = \int_0^\pi \sin^{n-3} \phi \, d\phi.$$

Then, $|F(\phi)|^2 = \cos^2 \theta + \sin^2 \theta \cos^2 \phi$, so

$$|P_k^n(\cos \theta)| \leq R(n)^{-1} \int_0^\pi (\cos^2 \theta + \sin^2 \theta \cos^2 \phi)^{k/2} \sin^{n-3} \phi \, d\phi.$$

The right side is decreasing in k , and we can estimate the integrals rigorously using interval arithmetic.

For (iii) we need to compute $\text{lhs}(k)$ for all $k \leq k_0$. We would like to do this rigorously, using for instance interval arithmetic. The most time-consuming step here is to compute the r_i functions. In practice, this step involves evaluating the polynomials P_k^n for values of k that can exceed 100,000.

The Jacobi polynomials P_k^n are given by a simple recurrence, namely

$$P_k^n(u) = a_k^n(u)P_{k-1}^n(u) + b_k^n P_{k-2}^n(u)$$

for $k \geq 2$ with $P_1^n(u) = u$ and $P_0^n(u) = 1$, where

$$a_k^n(u) = \frac{2k + 2\alpha - 1}{k + 2\alpha} u \quad \text{and} \quad b_k = -\frac{k - 1}{k + 2\alpha}$$

with $\alpha = (n - 3)/2$. This recurrence comes from formula (4.5.1) in Szegő [98], adapted to our normalization of $P_k^n(1) = 1$.

The recurrence is very stable: even using double-precision floating-point arithmetic it is possible to accurately evaluate the polynomial for very high degrees for points in $[-1, 1]$. If we use this recurrence with interval arithmetic though, the error estimation quickly gets out of hand: if $\lim_{k \rightarrow \infty} a_k(u) > 1$, the error bound grows exponentially.

Using interval arithmetic then requires very high precision and is very slow, though not prohibitively so. In any case, we can trust floating-point computations. Using the recurrence amounts to solving a linear system with a triangular matrix whose entries are the numbers $a_k(u)$, b_k , and 1 by backward substitution, and this matrix is well conditioned, so the error we make in solving the system is very small. The error was analyzed for instance by Barrio [9]. The Julia program that performs the verification uses high-precision floating-point arithmetic.

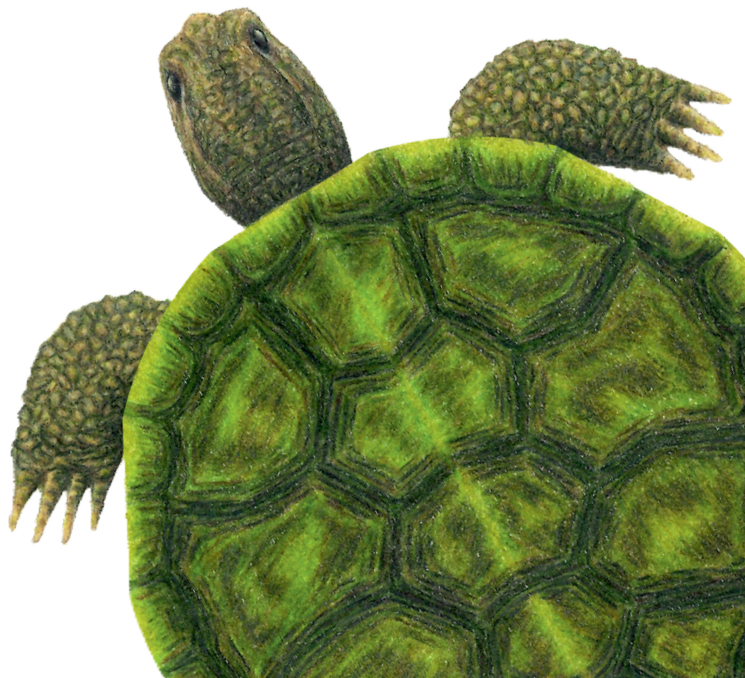
Acknowledgments

We would like to thank David de Laat, Nando Leijenhorst, Fabrício Caluza Machado, and Willem de Muinck Keizer for fruitful discussions. David de Laat and Nando Leijenhorst also gave some much-needed technical support

regarding the `ClusteredLowRankSolver`. The optimization problems were solved in a computational cluster at TU Delft maintained by Joffrey Wallaart.

Part 3

Packings in compact spaces



Copositive programming for compact packing graphs

Part 3 concerns problems in Class II from the introduction. These are counting problems on compact, but infinite, spaces. Prime examples of these are the spherical-codes problems, of which Problem II—the kissing number problem—is a special case.

A spherical-codes problem is as follows. For $\theta \in (0, \pi)$, a *spherical code* with angle θ is a subset $S \subseteq S^{n-1}$ such that the angle between two distinct points $x, y \in S$ is at least θ . The *spherical-codes problem* with angle θ asks what is the largest cardinality $A(n, \theta)$ that a spherical code with angle θ can have. This is the same as asking how many spherical caps of angular radius $\theta/2$ fit on S^{n-1} without overlapping, from which it follows that this maximum cardinality is indeed finite and attained.

We model such problems as independent-set problems on graphs with the following properties. A *packing graph* is a graph $G = (V, E)$ with V a topological space such that every finite clique of G is contained in an open clique. It is called a *compact packing graph* if V is compact. A compact packing graph has a finite independence number, i.e.

$$\alpha(G) = \sup\{|I| : I \subseteq V \text{ is independent}\} < \infty.$$

For a spherical-codes problem, the corresponding packing graph is the graph $G_\theta = (S^{n-1}, E(\theta))$, where $xy \in E(\theta)$ if and only if the angle between x and y is in the interval $(0, \theta)$. A set $S \subseteq S^{n-1}$ is an independent set of G_θ if and only if it is a spherical code with angle θ . Hence, $\alpha(G_\theta)$ is indeed $A(n, \theta)$.

Linear programming methods for spherical codes and other compact packing problems were introduced to the topic early on. Delsarte [37] introduced linear programming methods to the *binary-codes problem* in 1972; this is the spherical-codes problem, but with the sphere replaced by the Hamming cube $(\mathbb{Z}/2\mathbb{Z})^n$, and the angle between two points replaced by the *Hamming distance*: the number of distinct entries between vectors. His method was adapted to the spherical-codes problem in 1977 by Delsarte, Goethals, and Seidel [38].

McEliece, Rodemich, and Rumsey [73], and independently Schrijver [94], observed that these types of linear programming bounds are symmetry-reduced versions of the Lovász ϑ -number, respectively in 1978 and 1979. By today's measures, it is a fairly simple method, but already gave rise to sharp bounds in many cases [69, 79].

Schrijver [95] introduced semidefinite programming to the topic in 2005 by describing a three-point bound that gives upper bounds on the binary codes problem. In their landmark paper, Bachoc and Vallentin [6] extended the three-point bound to spherical codes in 2008.

The moment hierarchy was studied in depth for the independence number of finite graphs in 2003 by Laurent [67], and based on this, Gvozdeni-*no*vić, Laurent, and Vallentin [50] introduced a k -point bound for $k \geq 2$. This enabled the extension of the three-point bound to general compact packing graphs, and for those graphs, a moment hierarchy was introduced by De Laat and Vallentin [63] in 2015 and a block moment hierarchy by De Laat, Machado, Oliveira, and Vallentin [60] in 2022. Low levels of these hierarchies were implemented successfully for several problems on the unit sphere by De Laat, Machado, and De Muinck Keizer [61] in 2023 and Cohn, De Laat, and Leijenhurst [29] in 2024.

Dobre, Dür, Frerick, and Vallentin introduced copositive optimization to the topic in 2016, by describing an exact copositive formulation of the independence number of a compact metrizable packing graph. Based on this result, Kuryatnikova and Vera [57] defined a copositive hierarchy for compact packing graphs with metrizable vertex set in 2017, and they showed it converges to the independence number. These results are analogous to the approach for finite graphs Section 1.3 of this thesis; the hierarchy is based on an extension of Pólya's theorem to continuous kernels on compact spaces. See also the PhD thesis of Kuryatnikova [56].

In this chapter we study the relation between the moment hierarchy, the block moment hierarchy, and the copositive programming hierarchy for compact packing graphs, and show that they converge to the independence number if the vertex set is in addition metrizable. To be precise, the novel contribution in this chapter is a proof that the copositive hierarchy introduced by Kuryatnikova and Vera is weaker than the block moment hierarchy by De Laat, Machado, Oliveira and Vallentin. The convergence of the copositive hierarchy thus implies convergence of the block moment hierarchy. The exposition is based on the preprint [13].

1. Spaces of measures

Let V be a compact Hausdorff space. We say that $\mu \in M(V^k)$ is *symmetric* if for all $E \subseteq V^k$ and all $\pi \in \mathfrak{S}_k$ we have $\mu(\pi E) = \mu(E)$, where

$$\pi E = \{ (v_{\pi 1}, \dots, v_{\pi k}) : (v_1, \dots, v_k) \in E \}.$$

Denote the space of symmetric signed Radon measures on V^k by $M_{\text{sym}}(V, k)$.

The pair $(C(V), M(V))$ is a dual pair with the duality

$$\langle f, \mu \rangle = \mu(f) = \int_V f(v) d\mu(v).$$

This defines a duality between $C_{\text{sym}}(V, k)$ and $M_{\text{sym}}(V, k)$. All asterisks in this chapter refer to the dual with respect to this duality, i.e. if $\mathcal{C} \subset C_{\text{sym}}(V, k)$ a cone, then $\mathcal{C}^* \subset M_{\text{sym}}(V, k)$. The notation $C_{\text{sym}}(V)$ and $M_{\text{sym}}(V)$ denotes $C_{\text{sym}}(V, 2)$ and $M_{\text{sym}}(V, 2)$ respectively.

We define cones of Radon measures on a compact Hausdorff space V through this duality. Thus, the cone of nonnegative Radon measures $M(V)_{\geq 0}$ is the cone of signed Radon measures for which $\langle f, \mu \rangle \geq 0$ for all $f \in C(V)_{\geq 0}$. Likewise, the cone $M_{\text{sym}}(V)_{\geq 0}$ of *positive-semidefinite* measures on V^2 consists of signed Radon measures μ such that $\langle F, \mu \rangle \geq 0$ for each positive-semidefinite kernel $F \in C_{\text{sym}}(V)$.

Let $k \geq 0$ be an integer and $G = (V, E)$ be a graph with V a compact Hausdorff space. Recall from Section 5.1 that $\mathcal{I}_k$ denotes the subset of $\text{Sub}_k(V)$ of independent sets. Let $C_{\text{sym}}(V^2 \times \mathcal{I}_k)$ be the space of functions in $C(V^2 \times \mathcal{I}_k)$ that are invariant under permutation of the first two arguments. Define the space $M_{\text{sym}}(V^2 \times \mathcal{I}_k)$ as the space of all signed Radon measures on $V^2 \times \mathcal{I}_k$ such that $\mu(F^\top) = \mu(F)$ for all measurable $F \subseteq V^2 \times \mathcal{I}_k$, where

$$F^\top = \{ (u, v, w) : (v, u, w) \in F \}.$$

Let $C_{\text{sym}}(V^2 \times \mathcal{I}_k)_{\geq 0}$ be the cone of continuous functions F such that the function $(u, v) \mapsto F(u, v, w)$ is a positive semidefinite kernel for all $w \in \mathcal{I}_k$, and denote its dual by $M_{\text{sym}}(V^2 \times \mathcal{I}_k)_{\geq 0}$.

2. Compact packing graphs

Recall the preliminaries on the standard topology on $\text{Sub}_r(V)$ for a topological space V from Section 1 of Chapter 5, in particular that a basis of the standard topology is given by sets of the form

$$(U_1, \dots, U_t)_r = \{ A \in \text{Sub}_r(V) : A \cap U_i \neq \emptyset \text{ for all } i \text{ and } A \subseteq U_1 \cup \dots \cup U_t \},$$

with t an integer such that $0 \leq t \leq r$ and $U_1, \dots, U_r \subseteq V$ disjoint open sets. Moreover, recall that, for a given graph G , the notation $\mathcal{K}_r$ denotes the set of cliques of size at most r .

The following characterization of packing graphs was not found in the literature yet. It will be of use later.

THEOREM 7.1. *If $H = (V, E)$ is a graph and V is a Hausdorff space, the following are equivalent:*

- (i) G is a packing graph;
- (ii) $\mathcal{K}_r$ is open in $\text{Sub}_r(V)$ for every $r \geq 0$;
- (iii) $\mathcal{K}_2$ is open in $\text{Sub}_2(V)$.

Proof. To see that (i) implies (ii), let $C = \{x_1, \dots, x_s\}$ be a clique of cardinality $s \leq r$. It is contained in an open clique K . Since V is a Hausdorff space, there are disjoint open sets $U_1, \dots, U_s$ such that $x_i \in U_i$ for all i . By taking the intersection of each U_i with K , assume $\bigcup_i U_i$ is a clique. The set $(U_1, \dots, U_s)_r$ is an open set of $\text{Sub}_r(V)$ containing C and consisting only of cliques. This proves that $\mathcal{K}_r$ is open in $\text{Sub}_r(V)$ for all $r \geq 0$.

That (ii) implies (iii) is immediate.

Assume (iii) holds. Let C be a finite clique and $\mathbf{x} = \{x_1, x_2\}$ be a subset of size ≤ 2 . Then there exists a basic open set $(U_{\mathbf{x}}^{x_1}, U_{\mathbf{x}}^{x_2})_2 \subseteq \mathcal{K}_2$, where $U_{\mathbf{x}}^{x_i}$ is an open neighborhood of x_i , because sets of this form produce a basis of the topology. Choose such a basic open set for each $\mathbf{x} \subseteq C$ of size at most 2, and define

$$K = \bigcup_{x \in C} \bigcap_{\substack{\mathbf{x} \in \text{Sub}_2(C) \\ \text{with } x \in \mathbf{x}}} U_{\mathbf{x}}^x.$$

The set K is a union of finite intersections of open sets that contains C , thus it is an open neighborhood of C . Moreover, K is a clique: let $\mathbf{y} \subseteq K$ be a set of cardinality 2, say $\mathbf{y} = \{y_1, y_2\}$. For $i \in \{1, 2\}$ there exists an $x_i \in C$ such that $y_i \in U_{\mathbf{x}}^{x_i}$ for every $\mathbf{x} \subseteq C$ containing x_i . Choose such x_1 and x_2 , and take $\mathbf{x} = \{x_1, x_2\}$. Then, per definition of the $U_{\mathbf{x}}^{x_i}$, $\mathbf{y} \in (U_{\mathbf{x}}^{x_1}, U_{\mathbf{x}}^{x_2})_2 \subseteq \mathcal{K}_2$, so $\mathbf{y}$ is an edge. Thus, we found an open clique containing C , and G is a packing graph. $\square$

The following lemma is essential for much of the analysis in this chapter.

LEMMA 7.2 ([63, Lemma 2]). *If $G = (V, E)$ is a compact packing graph, then $\mathcal{I}_{=r}$ is both open and closed in $\mathcal{I}_k$ for all $r \leq k$. Hence, if Y is a topological space, then $f : \mathcal{I}_k \rightarrow Y$ is continuous if and only if the restriction of f to $\mathcal{I}_{=r}$ is continuous for all $r \leq k$.*

In particular, if $G = (V, E)$ is a compact packing graph, then

$$C(\mathcal{I}_k) \cong \bigoplus_{r=0}^k C(\mathcal{I}_{=r}).$$

3. The moment and block moment hierarchies

The definitions of the moment hierarchy and the block moment hierarchy are similar to those of Chapter 5, but there are notable differences. For compact packing graphs, we may formulate the problem over the space of Radon measures. If we do this, it turns out to be favorable to realize the edge constraint by only considering elements of $M(\mathcal{I}_r)$ rather than $M(\text{Sub}_r(V))$. We use the same notation M_r and $\text{block}M_r$ as in Chapter 5; because the programs there were defined for measurable graphs, this should not lead to confusion.

Recall the operator on $\mathbb{R}^V$ for finite V

$$M_r : \mathbb{R}^{\text{Sub}_{2r}(V)} \rightarrow \text{Sym}(\text{Sub}_r(V)), \quad (M_r \nu)_{S,T} = \nu_{S \cup T}.$$

To define the analogue of this operator on a space of measures, it is easier to first extend its adjoint to function spaces. A direct calculation shows that for all $A \in \text{Sym}(\text{Sub}_r(V))$,

$$(M_r^* A)_S = \sum_{\substack{J, J' \in \text{Sub}_r(V) \\ J \cup J' = S}} A_{J, J'}.$$

To understand this operator, consider that for generic V , the only additional structure given is the union map $\text{Sub}_r(V) \times \text{Sub}_r(V) \rightarrow \text{Sub}_{2r}(V)$. We want a map $\text{Sym}(\text{Sub}_r(V)) \rightarrow \mathbb{R}^{\text{Sub}_{2r}(V)}$ that treats all elements in a fixed fiber of the union map the same. This is achieved by giving all elements the same weight, weight 1, and summing the values together; i.e., we average the function values in a fiber of the union map under a uniform distribution. A similar interpretation holds for the block moment hierarchy, but with the union maps $\text{Sub}_1(V)^2 \times \text{Sub}_{r-2}(V) \rightarrow \text{Sub}_r(V)$, which results in the operators described later.

Let us proceed with defining the hierarchies. Let $G = (V, E)$ be a compact packing graph and let $r \geq 1$ be an integer. Equip the spaces $\mathcal{I}_r$ with the topology induced by the standard topology of $\text{Sub}_r(V)$. Define the operator

$$A_r : C_{\text{sym}}(\mathcal{I}_r) \rightarrow C(\mathcal{I}_{2r}), \quad A_r F(I) = \sum_{\substack{J, J' \in \mathcal{I}_r \\ J \cup J' = I}} F(J, J').$$

It is a priori not clear that the codomain of this operator is correct: $A_r F$ might not be a continuous function. We postpone the proof that the functions $A_r F$

are indeed continuous for now. Theorem 7.3 gives a related result. Assuming this for now, A_r is a bounded linear operator under the supremum norm, thus its continuous adjoint $A_r^* : M(\mathcal{I}_{2r}) \rightarrow M_{\text{sym}}(\mathcal{I}_r)$ exists. Define the r th level of the *moment hierarchy* as the optimization problem

$$\begin{aligned} M_r(G) = \sup \quad & \nu(\mathcal{I}_{=1}) \\ & \nu(\{\emptyset\}) = 1, \\ & A_r^* \nu \in M_{\text{sym}}(\mathcal{I}_r)_{\geq 0}, \\ & \nu \in M(\mathcal{I}_{2r})_{\geq 0}. \end{aligned}$$

Let $r \geq 2$ be an integer, and define the operator

$$\begin{aligned} B_r : C_{\text{sym}}(\mathcal{I}_1^2 \times \mathcal{I}_{r-2}) &\rightarrow C_{\text{sym}}(\mathcal{I}_r), \\ B_r F(I) &= \sum_{\substack{Q \in \text{Sub}_{r-2}(V) \\ Q \cup J \cup J' = I}} \sum_{\substack{J, J' \in \text{Sub}_1(I) \\ J \cup J' = I}} F(Q, J, J'). \end{aligned}$$

By Theorem 7.3, B_r is a bounded linear operator which has a continuous adjoint $B_r^* : M_{\text{sym}}(\mathcal{I}_r) \rightarrow M_{\text{sym}}(\mathcal{I}_1^2 \times \mathcal{I}_{r-1})$. Define the *block moment hierarchy* as the optimization problem

$$\begin{aligned} \text{block}M_r(G) = \sup \quad & \nu(\mathcal{I}_{=1}) \\ & \nu(\{\emptyset\}) = 1, \\ & B_r^* \nu \in M_{\text{sym}}(\mathcal{I}_1^2 \times \mathcal{I}_{r-2})_{\geq 0}, \\ & \nu \in M(\mathcal{I}_r)_{\geq 0}. \end{aligned}$$

The definition of the block moment hierarchy deviates in two ways from the one by De Laat, Machado, Oliveira, and Vallentin [60], resulting in a bound that is at least as strong as theirs. First, the original formulation excludes the empty set from $\mathcal{I}_r$, and second, a different normalization is used. Including the empty set is necessary for our proof of convergence; it seems to give stronger, nonequivalent problems. Changing the normalization does not affect convergence, and in fact, the proof of convergence hinges on this fact.

Like for the measurable independence number, we can restrict a feasible solution of $M_{r+1}(G)$ to one of $M_r(G)$, and the same is true for $\text{block}M_r(G)$, which shows that

$$M_1(G) \geq M_2(G) \geq \cdots \quad \text{and} \quad \text{block}M_1(G) \geq \text{block}M_2(G) \geq \cdots.$$

As was the case in Chapter 5, $M_r(G) \leq \text{block}M_{r+1}(G)$.

Moreover, if $I \subseteq V$ is an independent set of G , then

$$\nu = \sum_{R \in \mathcal{I}_r, R \subseteq I} \delta_R,$$

where δ_R is the Dirac measure at R , is a feasible solution to $\text{blockM}_r(G)$ with objective value $|I|$, so $\text{blockM}_r(G) \geq \alpha(G)$ for all r . Similarly, $\text{M}_r(G) \geq \alpha(G)$ for all r .

We now show that the operator $B_r : C_{\text{sym}}(\mathcal{I}_1^2 \times \mathcal{I}_{r-2}) \rightarrow C(\mathcal{I}_r)$ is well-defined, that is, if $F \in C_{\text{sym}}(\mathcal{I}_1^2 \times \mathcal{I}_{r-2})$, then $B_r F$ is continuous. The proof for A_r is similar. Lemma 7.2 is essential here: it implies we only have to prove continuity on $\mathcal{I}_{=r}$ for all r .

THEOREM 7.3. *If $G = (V, E)$ is a graph with V a compact Hausdorff space, if $r \geq 2$ is an integer, and if $F \in C(\mathcal{I}_1^2 \times \mathcal{I}_{r-2})$, then $B_r F$ is continuous on $\mathcal{I}_{=s}$ for all $s \leq r$.*

Proof. Fix an integer $0 \leq s \leq r$, and let $(I_\alpha)_\alpha$ be a net in $\mathcal{I}_{=s}$ that converges to $I \in \mathcal{I}_{=s}$; the task is to show that $B_r F(I_\alpha)$ converges to $B_r F(I)$.

Say $I = \{x_1, \dots, x_s\}$ and take disjoint open neighborhoods $U_1, \dots, U_s$ of $x_1, \dots, x_s$ respectively. The set $W = (U_1, \dots, U_s)_r$ is an open neighborhood of I , so there exists an α_0 such that $S_\alpha \in W$ for all $\alpha \geq \alpha_0$. This implies that for each i , $S_\alpha \cap U_i \neq \emptyset$. Hence, since the U_i are disjoint, write the set $I_\alpha = \{x_{\alpha,1}, \dots, x_{\alpha,s}\}$, where $x_{\alpha,i} \in U_i$.

This shows that for every $\alpha \geq \alpha_0$, the double sum appearing in $B_r F(I_\alpha)$ can be ordered such that the elements $x_{\alpha,i}$ in a fixed place. The net $(x_{\alpha,i})_\alpha$ converges in V for all i , and so each of the sets being summed over converges in its respective space. Since F is continuous, the theorem follows. $\square$

4. The copositive hierarchy

In the remainder of the chapter we assume that the vertex sets are compact and metrizable. The reason that we assume that they are metrizable is that exactness of the copositive formulation for compact packing graphs was only proved for this case, see [39].

Let V be a compact metrizable space and $k \geq 2$ an integer. Define the *completely positive cone* on $M_{\text{sym}}(V, k)$ by

$$\text{CP}(V, k)_m = \overline{\text{cone}}\{\mu^{\otimes k} : \mu \in M(V)_{\geq 0}\}$$

with closure under the weak* topology on $M(V^k)$. Define the *copositive cone* on $C_{\text{sym}}(V, k)$ as $\text{COP}(V, k)_c = \text{CP}(V, k)_m^*$. So, a continuous k -tensor T

is copositive if and only if for every nonnegative measure $\mu \in M(V)$ we have $\int_{V^k} T(v) d\mu^k(v) \geq 0$.

When V is compact and metrizable, there exists a Radon measure ω with full support on V . With $\text{COP}(V, k) \subseteq L^2(V^k, \omega)$ as in Chapter 2 we have $\text{COP}(V, k)_c = \text{COP}(V, k) \cap C_{\text{sym}}(V, k)$. This is independent of the choice of such ω . This was argued by Dobre, Dür, Frerick, and Vallentin [39] for $k = 2$, but also holds for $k \geq 2$.

It is well known that a continuous tensor T is in $\text{COP}(V, k)_c$ if and only if $T[U] \in \text{COP}(U, k)$ for all finite $U \subseteq V$, where

$$T[U] = (T(v_1, \dots, v_k))_{v_1, \dots, v_k \in U};$$

see [39, Lemma 2.1] for a proof for $k = 2$, which holds for general k . This is equivalent to the identity

$$\text{CP}(V, k) = \overline{\text{cone}}\{\chi^{\otimes k} : \chi \in M(V)_{\geq 0} \text{ discrete and has finite support}\}.$$

An important distinction with the L^2 setting is that $\text{COP}(V, k)_c$ has an algebraic interior.

LEMMA 7.4. *If $k \geq 2$ is an integer and V is a compact metrizable space, then $\mathbf{1}^{\otimes k} \in \text{algint } \text{COP}(V, k)_c$.*

Proof. It suffices to check that for all $F \in C_{\text{sym}}(V, k)$ there exists $\lambda > 0$ such that for every nonzero $\chi \in M(V)_{\geq 0}$ that is discrete and has finite support, $\langle \mathbf{1}^{\otimes k} - \lambda F, \chi^{\otimes k} \rangle \geq 0$. Take $F \in C_{\text{sym}}(V, k)$ nonzero and $\lambda = \|F\|_{\infty}^{-1}$, then, for all $\chi \in M(V)_{\geq 0}$ discrete and with finite support,

$$\langle \mathbf{1}^{\otimes k} - \lambda F, \chi^{\otimes k} \rangle = |\text{supp } \chi|^k - \lambda \langle F, \chi^{\otimes k} \rangle \geq |\text{supp } \chi|^k (1 - \lambda \|F\|_{\infty}) = 0,$$

and the conclusion follows. $\square$

For $r \in \mathbb{N}$, define the operators $\mathcal{T}_r : C(V^k) \rightarrow C_{\text{sym}}(V, k+r)$ similar to the $\mathcal{T}_r$ operators of Chapter 2:

$$\mathcal{T}_r F(v_1, \dots, v_{k+r}) = \frac{1}{(k+r)!} \sum_{\pi \in \mathfrak{S}_{k+r}} (F \otimes \mathbf{1}^{\otimes r})(v_{\pi 1}, \dots, v_{\pi(k+r)}).$$

They are bounded and linear. Define the cones

$$C_r(V, k)_c = \mathcal{T}_r^{-1} C_{\text{sym}}(V, k+r)_{\geq 0},$$

which are weakly closed and convex. Again, if V is metrizable, then these cones are the intersections $C_r(V, k) \cap C_{\text{sym}}(V, k)$ with $C_r(V, k) \subseteq L^2(V^k, \omega)$ with ω a measure with full support.

Kuryatnikova and Vera [56, Theorem 2.9] showed that for $k = 2$

$$C_1(V, k)_c \subseteq C_2(V, k)_c \subseteq \dots \subseteq \text{COP}(V, k)_c,$$

which we confirmed in Chapter 2 for all $k \geq 2$. They also proved, again for $k = 2$, that $\text{algint COP}(V, k)_c \subseteq \bigcup_r C_r(V, k)_c$; their proof goes through for all $k \geq 2$, which was shown in the appendix of [12]. This results in a version of Pólya's theorem for continuous tensors, but under the presence of a Radon measure with full support. Since our definition of $\text{COP}(V, k)_c$ does not depend on this measure, we also do not need it for the conclusion to hold; in fact, the proof goes through with little change. It is really just the second half of that of [12, Theorem A.1]; see also [57, Theorem 2.9].

THEOREM 7.5 (Pólya's theorem for continuous tensors). *Let V be a compact metrizable space. Given $T \in C_{\text{sym}}(V, k)$, let $M = \|T\|_\infty$ and*

$$\lambda = \inf \{ \langle T[U], x^{\otimes k} \rangle : U \subseteq V \text{ finite}, x \in \mathbb{R}_{\geq 0}^U, \text{ and } \mathbf{1}^\top x = 1 \}.$$

If $\lambda > 0$ then, for every $r > k(k-1)M/(2\lambda) - k$, $\text{Av}_{\mathfrak{S}_{k+r}}(T \otimes \mathbf{1}^{\otimes r}) \geq 0$.

Proof. See the proof of Theorem A.1 of [12]. □

This leads to the following approximation result for the copositive cone.

THEOREM 7.6. *If V is a compact metrizable space, then*

$$C_0(V, k)_c \subseteq C_1(V, k)_c \subseteq \cdots \subseteq \text{COP}(V, k)_c \text{ and} \\ \text{algint COP}(V, k)_c \subseteq \bigcup_{r \in \mathbb{N}} C_r(V, k)_c.$$

Proof. Again, see the appendix of [12]. □

Let us now discuss the completely positive and copositive hierarchies for packing graphs. From now on $k = 2$ everywhere, and $G = (V, E)$ is a packing graph with V a compact and metrizable space. Denote the diagonal of V^2 by Δ . Let $C_r(V)_c = C_r(V, 2)_c$, and similarly for $\text{COP}(V, 2)_c$ and $\text{CP}(V, 2)_m$.

For a convex cone $\mathcal{C} \subseteq M_{\text{sym}}(V)$, consider optimization problems of the form

$$\begin{aligned} \vartheta(G, \mathcal{C}) &= \sup \alpha(V^2) \\ \alpha_E &= 0, \\ \alpha(\Delta) &= 1, \\ \alpha &\in \mathcal{C}, \end{aligned} \tag{42}$$

where α_E is the restriction of α to the edge set. For $\mathcal{C}$ a convex cone in $C_{\text{sym}}(V)$,

$$\begin{aligned} \vartheta^*(G, \mathcal{C}) = \inf t \quad & \\ T(v, v) \leq t - 1 \quad & \text{for all } v \in V \\ T(v_1, v_2) \leq -1 \quad & \text{for all } v_1 v_2 \in E \\ T \in \mathcal{C}. \end{aligned}$$

As before, there should be no confusion with ϑ from Part 2, since there we only considered measurable graphs.

Given a closed convex cone $\mathcal{C} \subseteq M_{\text{sym}}(V)_{\geq 0}$, $\vartheta(G, \mathcal{C})$ and $\vartheta^*(G, \mathcal{C}^*)$ are dual in the sense of [10, Ch. IV]. Weak duality $\vartheta(H, \mathcal{C}) \leq \vartheta^*(H, \mathcal{C}^*)$ follows in the usual way, namely by taking feasible solutions α and (t, T) of the respective programs, and evaluating $\langle T, \alpha \rangle$ to obtain $\alpha(V^2) \leq t$. Dobre, Dür, Frerick and Vallentin [39] showed that if $G = (V, E)$ is a packing graph with V compact and metrizable, then we also have strong duality and exactness.

THEOREM 7.7 ([39]). *If $G = (V, E)$ is a packing graph with V compact and metrizable, then*

$$\alpha(G) = \vartheta(G, \text{CP}(V)_{\text{m}}) = \vartheta^*(G, \text{COP}(V)_{\text{c}}).$$

We shall refer to the sequence $(\vartheta(G, C_r(V)_{\text{c}}^*))_r$ as the completely positive hierarchy of G , and to $(\vartheta(G, C_r(V)_{\text{c}}))_r$ as its copositive hierarchy. Kuryatnikova and Vera proved [56, Theorem 2.17] that for certain compact packing graphs, the copositive hierarchy converges to the independence number. Their proof depends on the existence of a point $T_0 \in \text{algint COP}(V)_{\text{c}}$ such that for all $xy \in \overline{E}$, $Z(x, y) \leq -1$. The following theorem extends this result to all packing graphs with a compact and metrizable vertex set, by the fact that such a point exists for every compact Hausdorff space V .

THEOREM 7.8. *If $G = (V, E)$ is a packing graph with V compact and metrizable, then $\alpha(G) = \lim_r \vartheta(G, C_r(V)_{\text{c}}^*) = \lim_r \vartheta^*(G, C_r(V)_{\text{c}})$.*

Proof. It is enough to prove the inequality $\leq$ for each $\vartheta(G, C_r(V)_{\text{c}}^*)$, since weak duality ensures that then $\vartheta^*(G, C_r(V)_{\text{c}})$ for all r follows as well. Since for all r , $\vartheta(G, C_r(V)_{\text{c}}^*)$ is a relaxation of $\vartheta(G, \text{CP}(V)_{\text{m}})$, the inequality indeed holds.

For the other inequality, first prove that there exists $T_0 \in \text{algint COP}(V)_{\text{c}}$ such that $T_0(x, y) \leq -1$ for all $xy \in \overline{E}$. De Laat and Vallentin [63, Lemma 7] showed that there is a positive-semidefinite kernel F such that $F(x, y) \leq -1$ for all $xy \in \overline{E}$. Then $2F$ is positive semidefinite, and $2F(x, y) \leq -2$ for $xy \in \overline{E}$. Now take a suitable convex combination of $2F$ and $\mathbf{1}^{\otimes 2}$ to obtain the required kernel T_0 .

Dobre, Dür, Frerick, and Vallentin [39, Theorem 1.2] showed that for V compact and metrizable, $\vartheta^*(G, \text{COP}(V)_c) = \alpha(G)$. So, take a feasible solution (t, T) of $\vartheta^*(G, \text{COP}(V)_c)$. For every $0 < \epsilon \leq 1$ the kernel $\epsilon T_0 + (1 - \epsilon)T$ is in the algebraic interior of $\text{COP}(V)_c$. So by Theorem 7.6 it belongs to $C_r(V)_c$ for some $r \geq 0$. Thus, by taking $\epsilon \rightarrow 0$, it follows that $\lim_r \vartheta^*(G, C_r(V)_c) \leq t$, and the conclusion follows. $\square$

5. Comparing the hierarchies

Throughout this section, fix $G = (V, E)$ a packing graph with V compact and metrizable. This section contains the main subject of this chapter, namely the proof of the following theorem.

THEOREM 7.9. *If $G = (V, E)$ is a packing graph with V compact and metrizable, then for every $r \geq 1$, $M_{r+1}(G) \leq \text{block}M_{r+2}(G) \leq \vartheta(G, C_r(V)_c^*)$. In particular $\alpha(G) = \text{block}M_r(G) = M_{r-1}(G)$ for all $r \geq \alpha(G) + 2$.*

That $\text{block}M_r(G) = \alpha(G)$ for all $r \geq \alpha(G) + 2$ follows from the first part of the theorem and that if $r \geq \alpha(G) + 2$, then $\mathcal{I}_{r-2}$ is the space of all independent sets, thus the sequence $(\text{block}M_r(G))_r$ stabilizes. We will prove Theorem 7.9 by taking a feasible solution of $\text{block}M_{r+2}(H)$ and making a feasible solution of $\vartheta(H, C_r(V)_c^*)$ with at least the same objective value.

Let $r \geq 0$ be an integer and let $S \subseteq V$. Denote by $N_r(S)$ the number of tuples $v \in V^r$ such that $\llbracket v \rrbracket = S$. Recall the convention $V^0 = \{\emptyset\}$, so the definition includes $r = 0$. We do not specify a domain for N_r ; we will use it as a function on $\mathcal{I}_t$ for any t we need. Note that $N_r(S)$ depends only on the cardinality of S , so the function $N_r : \mathcal{I}_t \rightarrow \mathbb{R}$ is continuous.

Fix integers $r \geq 0$ and $0 \leq t \leq s$. Let $Q_{s,t} : C(V^t) \rightarrow C(\mathcal{I}_r)$ be the map such that

$$Q_{s,t}F(I) = \sum_{\substack{v \in V^s \\ \llbracket v \rrbracket = I}} F(v_1, \dots, v_t).$$

A proof similar to that of Theorem 7.3 shows that $Q_{s,t}F$ is indeed continuous for every continuous F , and each $Q_{s,t}$ is a bounded linear operator. Of course, these maps also depend on r , but we omit it from notation.

LEMMA 7.10. *Let r, s , and t be integers such that $r \geq 0$, $s \geq 1$ and $0 \leq t \leq s$, and let $G = (V, E)$ be a compact packing graph. Then,*

- (i) *for all $\nu \in M(\mathcal{I}_r)$, $Q_{s,t}^* \nu(V^t) = \langle N_s, \nu \rangle$, and*
- (ii) *if $t + 2 \leq s$, then $Q_{s,t+2} \mathcal{T}_t = Q_{s,2}$, hence $\mathcal{T}_t^* Q_{s,t+2}^* = Q_{s,2}^*$.*

Proof. That $Q_{s,t}$ is linear and bounded is clear, so its continuous adjoint is well-defined. If $\nu \in M(\mathcal{I}_r)$, then

$$Q_{s,t}^* \nu(V^t) = \langle Q_{s,t} \mathbf{1}_{V^t}, \nu \rangle = \int_{\mathcal{I}_k} \sum_{\substack{v \in V^s \\ \llbracket v \rrbracket = I}} 1 \, d\nu(I) = \langle N_s, \nu \rangle,$$

proving (i).

To see (ii), let $K \in C(V^2)$. For every $I \in \mathcal{I}_r$,

$$\begin{aligned} (Q_{s,t+2} \mathcal{T}_t F)(I) &= \sum_{\substack{v \in V^s \\ \llbracket v \rrbracket = I}} \frac{1}{(t+2)!} \sum_{\pi \in \mathfrak{S}_{t+2}} F(v_{\pi 1}, v_{\pi 2}) \\ &= \frac{1}{(t+2)!} \sum_{\pi \in \mathfrak{S}_{t+2}} \sum_{\substack{v \in V^s \\ \llbracket v \rrbracket = I}} F(v_{\pi 1}, v_{\pi 2}) \\ &= \sum_{\substack{v \in V^s \\ \llbracket v \rrbracket = I}} F(v_1, v_2) \\ &= Q_{s,2} F(I), \end{aligned}$$

and (ii) follows. $\square$

Proof of Theorem 7.9. Fix an integer $r \geq 1$ and a feasible solution ν of the problem $\text{blockM}_{r+2}(H)$ with positive objective value. Write $Q_t = Q_{r+2,t}$ for short, and take $C(\mathcal{I}_{r+2})$ for its codomain.

Write $\Phi_t = \langle N_t, \nu \rangle$ for $t \in \mathbb{N}$. Assume for now that $\Phi_t > 0$ for all t ; the proof of this will follow later. Write

$$\beta = Q_{r+2}^* \nu \in M(V^{r+2}) \quad \text{and} \quad \alpha = \mathcal{T}_r^* \beta \in M_{\text{sym}}(V).$$

The first objective is to show that $\Phi_{r+1}^{-1} \alpha$ is feasible for $\vartheta(H, C_r(V)_c^*)$.

To begin, if $F \in C(V^{r+2})$ is nonnegative, then so is $Q_{r+2} F$, and it follows that $\langle F, Q_{r+2}^* \nu \rangle = \langle Q_{r+2} F, \nu \rangle \geq 0$, hence β is nonnegative. If $F \in C_{\text{sym}}(V)$ is nonnegative, then so is $\mathcal{T}_r F$, whence $\langle F, \alpha \rangle = \langle F, \mathcal{T}_r^* \beta \rangle = \langle \mathcal{T}_r F, \beta \rangle \geq 0$, and $\alpha \in C_r(V)_c^*$.

Next, if $F \in C_{\text{sym}}(V)$ is a function with support in E , using Lemma 7.10

$$\langle F, \alpha \rangle = \langle F, \mathcal{T}_r^* Q_{r+2}^* \nu \rangle = \langle F, Q_2^* \nu \rangle = \int_{\mathcal{I}_{r+2}} \sum_{\substack{v \in V^{r+2} \\ \llbracket v \rrbracket = I}} F(v_1, v_2) \, d\nu(I) = 0.$$

Since E is open in V^2 , it is itself a locally compact Hausdorff space, and α_E is a signed Radon measure on E . It then follows from the Riesz representation theorem that $\alpha_E = 0$.

Next, calculate $\alpha(\Delta)$. Every vertex is contained in an open clique, so by compactness there are open cliques $C_1, \dots, C_m$ such that their union is V . The set $U = \bigcup_i C_i^2$ is an open set in V^2 whose union contains Δ . Moreover, if $(v_1, v_2) \in U$ such that $v_1 \neq v_2$, then $v_1 v_2 \in E$; thus, $U \setminus \Delta \subseteq E$.

Since Δ is closed, Urysohn's lemma gives a continuous $F : V^2 \rightarrow [0, 1]$ such that $F(v) = 1$ for all $v \in \Delta$ and $F(v) = 0$ for all $v \notin U$. Since $\alpha_E = 0$ and using Lemma 7.10

$$\begin{aligned} \alpha(\Delta) &= \langle F, \alpha \rangle = \langle F, \mathcal{T}_r^* Q_{r+2}^* \nu \rangle = \int_{\mathcal{I}_{r+2}} \sum_{\substack{v \in V^{r+2} \\ \llbracket v \rrbracket = I}} F(v_1, v_2) d\nu(I) \\ &= \int_{\mathcal{I}_{r+2}} \sum_{\substack{v \in V^{r+1} \\ \llbracket v \rrbracket = I}} 1 d\nu(I) = Q_{r+1, r+2}^* \nu(V^{r+2}) = \Phi_{r+1} \end{aligned}$$

To finish, it suffices to show that $\Phi_t > 0$ for all t and $\Phi_{r+2} \Phi_{r+1}^{-1} \geq \nu(\mathcal{I}_{=1})$, as the latter shows that $\alpha(\Delta)^{-1} \alpha$ is a feasible solution with objective at least $\nu(\mathcal{I}_{=1})$. This will follow from the following claim: if $t \leq r$, the matrix

$$\begin{pmatrix} \Phi_t & \Phi_{t+1} \\ \Phi_{t+1} & \Phi_{t+2} \end{pmatrix} \quad (43)$$

is positive semidefinite.

Indeed, assume the claim. Since $\Phi_0 = 1$ and $\Phi_1 = \nu(\mathcal{I}_{=1}) > 0$, $\Phi_2 > 0$ immediately follows. Repeating this argument results in $\Phi_t > 0$ for all t .

For the objective value, if the matrix (43) is positive semidefinite, the inequalities $\Phi_{t+2} \Phi_{t+1}^{-1} \geq \Phi_{t+1} \Phi_t^{-1}$ follow for all t . Repeated application of these inequalities yields

$$\alpha(V^2) \alpha(\Delta)^{-1} = \Phi_{r+2} \Phi_{r+1}^{-1} \geq \Phi_1 \Phi_0^{-1} = \nu(\mathcal{I}_{=1}),$$

as required.

Recall that a measure $\mu \in M_{\text{sym}}(V)$ is positive semidefinite if $\langle F, \mu \rangle \geq 0$ for every positive semidefinite kernel $F \in C_{\text{sym}}(V)$. To prove that (43) is positive semidefinite, we will prove that for every t there is a positive semidefinite measure μ such that $\mu(\mathcal{I}_{=i} \times \mathcal{I}_{=j}) = \Phi_{t+i+j}$ for $i, j \in \{0, 1\}$.

Fix an integer t such that $0 \leq t \leq r$. Employ the Riesz representation theorem to define μ as the measure such that $\langle F, \mu \rangle = \langle F \otimes N_t, B_{r+2}^* \nu \rangle$ for every $F \in C(\mathcal{I}_1^2)$, where $\mathcal{I}_r$ is the domain of N_t , and $(F \otimes N_t)(S, T) = F(S) N_t(T)$ for all $S \in \mathcal{I}_1^2$ and $T \in \mathcal{I}_r$.

To see that μ is positive semidefinite, let $F \in C_{\text{sym}}(\mathcal{I}_1)$ be a positive semidefinite kernel. Since $N_t \geq 0$, the kernel $(S, T) \mapsto F(S, T)N_t(Q)$ is positive semidefinite for every $Q \in \mathcal{I}_r$. Therefore, $F \otimes N_t \in C(\mathcal{I}_1^2 \times \mathcal{I}_r)_{\geq 0}$, and $\langle F, \mu \rangle \geq 0$ since $B_{r+2}^* \nu \in C(\mathcal{I}_1^2 \times \mathcal{I}_r)_{\geq 0}$.

Next, calculate $\mu(\mathcal{I}_{=i} \times \mathcal{I}_{=j})$ for $i, j \in \{0, 1\}$. For every $I \in \mathcal{I}_{r+2}$ and for $i = j = 0$,

$$(B_{r+2}(\mathbf{1}_{\{\emptyset\}}^{\otimes 2} \otimes N_t))(I) = \sum_{Q \in \text{Sub}_r(I)} \sum_{\substack{J, J' \in \text{Sub}_1(I) \\ Q \cup J \cup J' = I}} \mathbf{1}_{\{\emptyset\}}(J) \mathbf{1}_{\emptyset}(J') N_t(Q) = N_t(I).$$

For $i = 0$ and $j = 1$,

$$\begin{aligned} (B_{k+r}(\mathbf{1}_{\{\emptyset\}} \otimes \mathbf{1}_{\mathcal{I}_{=1}} \otimes N_t))(I) &= \sum_{Q \in \text{Sub}_r(I)} \sum_{\substack{x \in I \\ Q \cup \{x\} = I}} N_t(Q) \\ &= \sum_{Q \in \text{Sub}_t(I)} \sum_{\substack{x \in I \\ Q \cup \{x\} = I}} \sum_{\substack{v \in V^t \\ \llbracket v \rrbracket = Q}} 1. \end{aligned}$$

The map $(Q, x, v) \leftrightarrow (x, v)$ is a bijection between the set of triples (Q, x, v) in $\text{Sub}_t(I) \times I \times V^t$ such that $Q \cup \{x\} = I$ and $\llbracket v \rrbracket = Q$ and the set of tuples $v \in V^{t+1}$ such that $\llbracket v \rrbracket = I$. Hence,

$$(B_{k+r}(\mathbf{1}_{\{\emptyset\}} \otimes \mathbf{1}_{\mathcal{I}_{=1}} \otimes N_t))(I) = \sum_{\substack{v \in V^{t+1} \\ \llbracket v \rrbracket = I}} 1 = N_{t+1}(I).$$

Similarly, for $i = j = 1$,

$$\begin{aligned} (B_{k+r}(\mathbf{1}_{\mathcal{I}_{=1}}^{\otimes 2} \otimes N_t))(I) &= \sum_{Q \in \text{Sub}_r(I)} \sum_{\substack{x, y \in I \\ Q \cup \{x, y\} = I}} N_t(Q) \\ &= \sum_{Q \in \text{Sub}_t(I)} \sum_{\substack{x, y \in I \\ Q \cup \{x, y\} = I}} \sum_{\substack{v \in V^t \\ \llbracket v \rrbracket = Q}} 1 \\ &= \sum_{\substack{v \in V^{t+2} \\ \llbracket v \rrbracket = I}} 1 \\ &= N_{t+2}(I). \end{aligned}$$

Putting it all together, (43) is equal to

$$A = \begin{pmatrix} \mu(\mathcal{I}_{=0}^2) & \mu(\mathcal{I}_{=0} \times \mathcal{I}_{=1}) \\ \mu(\mathcal{I}_{=0} \times \mathcal{I}_{=1}) & \mu(\mathcal{I}_{=1}^2) \end{pmatrix}.$$

This matrix is positive semidefinite, since for $x_0, x_1 \in \mathbb{R}$

$$(x_0, x_1)^\top A(x_0, x_1) = \int_{\mathcal{I}_1^2} (x_0 \mathbf{1}_{\mathcal{I}=0} + x_1 \mathbf{1}_{\mathcal{I}=1})^{\otimes 2}(S, T) d\mu(S, T) \geq 0,$$

from which the theorem follows. $\square$

6. What about packing hypergraphs?

Much of the analysis in this chapter works because of Lemma 7.2. On the side of functions, it implies that continuity on $\mathcal{I}_k$ is ensured by separate continuity on each $\mathcal{I}_{=r}$ for all $r \leq k$. On the side of the measures, together with Urysohn's lemma, it gives a degree of control over the measure of lower dimensional sets—like the diagonal in V^2 —that is otherwise out of reach. Not only are these properties important for Theorem 7.9, they are also necessary for the definition of the copositive formulation, the moment hierarchy, and the block moment hierarchy.

For the conclusion of Lemma 7.2 to hold for k -uniform packing hypergraphs, it is sufficient to ask that again every finite clique is contained in an open clique, which is equivalent to the properties of Theorem 7.1 with $\mathcal{K}_2$ replaced by $\mathcal{K}_k$. In this setting, the copositive formulation and all other definitions and results seem to go through unchanged.

However, in the next chapter we will see a 3-uniform hypergraph that reasonably can be called a packing hypergraph, but for which there are pairs—which are vacuously cliques—that are not contained in an open clique; they are not even contained in any clique. Thus, there are independent triples that converge to a pair without every passing through $\mathcal{K}_{=3}$. As a result, it can be shown that there are k for which $\mathcal{I}_k$ is not open in $\mathcal{I}_{k+1}$.

A simple solution to this is to replace the topology of $\mathcal{I}_k$ by the disjoint-union topology $\bigsqcup_{r \leq k} \mathcal{I}_{=r}$. This seems correct for the formulation of a moment and a block moment hierarchy, but it comes with other complications. Most importantly, the space $\bigsqcup_{r \leq k} \mathcal{I}_{=r}$ is not compact, which means that many steps we took might not make sense anymore.

CHAPTER 8

Application: obtuse almost-equiangular sets

This chapter presents a geometric question that can be formulated as the independence number of a 3-uniform hypergraph. This hypergraph behaves like a compact packing graph and can therefore be thought of as an example of a “packing hypergraph”. What follows is a lightly edited version of the preprint by Bachoc, Bekker, Moustrou, and Oliveira [4].

Given $t \in [-1, 1)$, a set $S \subseteq S^{n-1}$ is *t-almost-equiangular* if every 3-subset $\{x, y, z\}$ of S is such that $t \in \{x^\top y, x^\top z, y^\top z\}$. In the literature, the word “almost” is often replaced by “nearly”. An *obtuse almost-equiangular set* is a *t-almost equiangular set* with $t \leq 0$. A 0-almost-equiangular set is also often called *almost-orthogonal*. Similarly, one may define *almost-equidistant* subsets of a metric space, of which almost-equiangular sets are a special case.

Denote the maximum cardinality of a *t-almost-equiangular set* in S^{n-1} by $\alpha(n, t)$. The problem of finding $\alpha(n, t)$ is called the *t-almost-equiangular-set problem*. For $t = 0$, this problem first appears in a paper by Rosenfeld [91], who attributes the question to Erdős. Rosenfeld showed that $\alpha(n, 0) = 2n$; a lower bound is given by the union of two disjoint orthogonal bases and an upper bound is given through an interesting argument involving the spectrum of a matrix associated to an almost-equiangular set. Pudlák [88] and Deaett [34] reproved this result by simpler methods.

Later, Bezdek and Lángi [14] extended Rosenfeld’s spectral bound to all $t \in [-1, \varepsilon]$, where $\varepsilon > 0$ is a number close to 0 that depends on the dimension. In particular, they proved that $\alpha(n, t) \leq 2(n + 1)$ on this interval with equality at $t = -1/n$. An example of an optimal construction at this inner product is the union of two disjoint regular n -simplices. Polyanskii [87] mentioned a simple lifting argument to obtain $\alpha(n, t) \leq 2(n + 1)$ for $t \leq 0$ directly from Rosenfeld’s original result.

The goal of the current work is two-fold. First, to obtain better upper bounds on the number $\alpha(n, t)$, which is done through semidefinite programming and a closer investigation of the spectral bound of Bezdek and Lángi. Both methods reproduce known bounds, and improve many others. Second, to list all t -almost-equiangular subsets of S^{n-1} of size $\alpha(n, t)$ for small n . The spectral bound of Bezdek and Lángi again plays an important role; it is used to derive characterizing properties of those t -almost-equidistant sets in S^{n-1} that are maximum for all $t \in [-1, 0]$.

Upper bounds through semidefinite programming. For $t \in [-1, 1)$, the *equiangular-lines problem* asks for the maximum number of vectors in S^{n-1} such that any two distinct vectors have inner product $\pm t$. This problem can be rephrased as an independence-number problem on a compact packing graph, similar to the spherical-codes problem. The methods discussed in Chapter 7 are thus applicable, and in fact result in the best known bounds on the equiangular-lines problem, see also De Laat, Machado, and De Muinck Keizer [61].

The t -almost-equiangular-set problem can be rephrased as a question on independent sets of a hypergraph. Given $n \geq 2$ and $t \in [-1, 1)$, let $H(n, t)$ be the 3-uniform hypergraph whose vertex set is S^{n-1} and in which a 3-set $\{x, y, z\}$ of points is an edge if $t \notin \{x^\top y, x^\top z, y^\top z\}$. Then independent sets of H correspond to t -almost-equiangular sets and vice versa. It follows that $\alpha(n, t) = \alpha(H(n, t))$.

This connection again opens the door to the development of optimization upper bounds for $\alpha(n, t)$. Castro-Silva, Oliveira, Slot and Vallentin [24] proposed an extension of the Lovász theta number to finite hypergraphs. A further extension to infinite hypergraphs by the same authors [23] has applications in Euclidean Ramsey theory. The underlying hypergraphs are unlike packing graphs, and the setting is more in the spirit of what we studied in Part 2 of this thesis. The current chapter proposes an alternative extension of the theta number to infinite hypergraphs like $H(n, t)$ based on the moment hierarchy and the block moment hierarchy [63, 60]. This bound is strongly related to the semidefinite programming methods developed in [15, 16], where similar techniques were used to reprove Rosenfeld's original bound, and further apply them to energy minimization questions on hypergraphs.

This allows for the computation of upper bounds for $\alpha(n, t)$ through the use of sums of squares and semidefinite programming. Analytic bounds can be obtained by interpolating solutions of the resulting semidefinite programming problems, leading to the following theorem proved in Section 3.

THEOREM 8.1. *If $t \in [-1, 0]$ and $n \geq 3$, and if*

$$f(n, t) = p^2 n(1 - t)^2 / (2(nt^2 + 1)),$$

where

$$p = \frac{8n^2 t^4 (2n - 1) - 9n^2 t^3 (n - 1) + (2nt^2 - 3t + 4)(7n + 1)}{2(1 - t)(1 + 7n - 2n^2 t^3 (2n - 1))},$$

then $\alpha(n, t) \leq \lfloor f(n, t) \rfloor \leq \lfloor (16t - 9)^2 / (128t^2) \rfloor$.

The bound $(16t - 9)^2 / (128t^2)$ in this theorem is an asymptotic bound; it is the limit of $f(n, t)$ as n goes to infinity. That there exists an upper bound that does not depend on the dimension n is consistent with the existence of the constructions considered in this chapter not explicitly depending on the embedding dimension n if t is far enough removed from $-1/n$.

Lower bounds through constructions. If S is a t -almost-equiangular set, then its *distance- t graph*, namely the graph with vertex set S in which x and y are adjacent if $x^\top y = t$, is *anti-triangle free*, that is, its complement does not contain triangles.

Necessary and sufficient conditions for some anti-triangle-free graphs to be the distance graph of an almost-equiangular set are given in Section 4. Together with the optimization bound of Theorem 8.1 and the results of Section 5, this leads to constraints for the existence of t -almost-equiangular sets of certain sizes, making it possible to list all optimal such sets for dimensions $n = 2$ and 3 . This search leads to the optimal constructions listed in Section 6 and summarized in Figure 3.

Maximum obtuse almost-equiangular sets. Both the semidefinite programming bound of Theorem 8.1 and the spectral bound of Rosenfeld [91] and Bezdek and Lángi [14] show that $\alpha(n, t) \leq 2(n + 1)$ for all n and $t \leq 0$. In Section 5 the spectral bound is investigated further to show that equality for nonpositive t is only attained at $t = -1/n$. In light of this, call the maximum $(-1/n)$ -almost-equidistant sets on S^{n-1} *maximum obtuse almost-equiangular sets*. Inspection of the matrices that are associated to the maximum obtuse almost-equiangular sets in the proof of the spectral bound reveals several interesting properties of these sets, like the following result.

THEOREM 8.2. *A maximum obtuse almost-equiangular set on S^{n-1} is a spherical 2-design.*

Deaett proved [34] that there is a bijection between the maximum almost-orthogonal sets in S^{n-1} and certain $2n \times 2n$ symmetric orthogonal matrices.

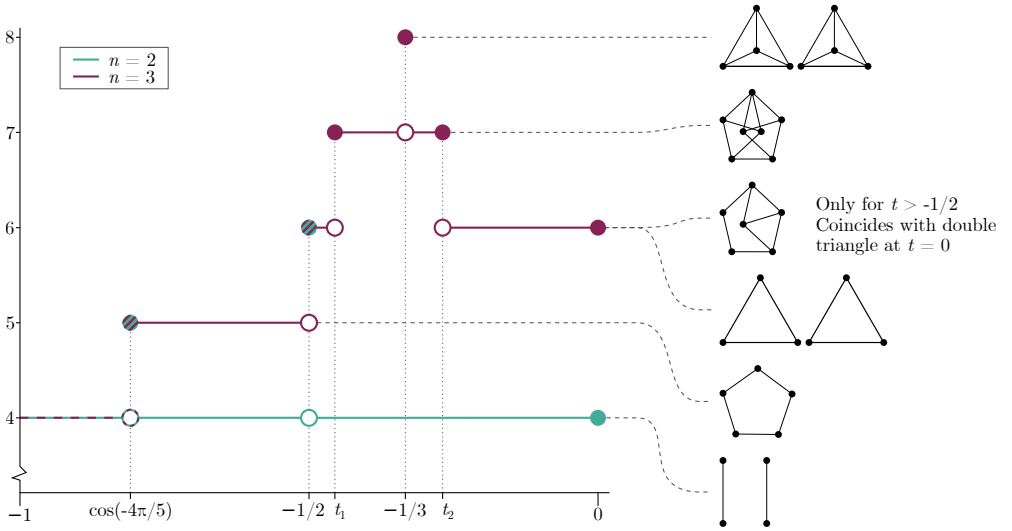


FIGURE 3. The classification of maximum-cardinality t -almost-equiangular sets in S^1 and S^2 for $t \in [-1, 0]$. The vertical axis is the cardinality of the set, the horizontal axis is the inner product t . Open bullets indicate that a point is excluded from the interval while closed bullets indicate that the point is included. Stripes indicate that in both dimensions the same maximum cardinality is attained. The numbers t_1 and t_2 are the first two roots of (54) with $k = 2$.

Any t -almost-equiangular set in S^{n-1} with $t \leq 0$ can be lifted to an almost-orthogonal set on S^n [87], and so it is expected that there is a version of this bijection for maximum nonpositive almost-equiangular sets as well. The bijection is made precise in the following theorem, which is Deaett's correspondence with the addition of the eigenvector condition (i). Here, e is the all-ones vector.

THEOREM 8.3. *There exists a bijection between the maximum obtuse almost-equiangular subsets of S^{n-1} up to orthogonal transformations and symmetric, orthogonal matrices $O \in \text{Sym}(2(n+1))$ such that*

- (i) $Oe = e$;
- (ii) $O_{ii} = 0$ for all i ;
- (iii) $O_{ij}O_{jk}O_{ki} = 0$ for all i, j , and k .

The union of two disjoint regular n -simplices is called a *double-regular n -simplex*. It remains an open question whether a maximum obtuse almost-equiangular set is always a double-regular n -simplex. However, with the help of Theorem 8.3 the question is settled for $2 \leq n \leq 5$.

THEOREM 8.4. *If $2 \leq n \leq 5$, then any maximum obtuse almost-equiangular set in S^{n-1} is a double-regular n -simplex.*

1. Preliminaries

Hypergraphs. Given a hypergraph H , we will denote its vertex set by $V(H)$ and its edge set by $E(H)$. Given $S \subseteq V(H)$, the subgraph of H induced by S , denoted by $H[S]$, is the hypergraph with vertex set S whose edges are all edges of H contained in S . For all $S \subseteq V(H)$, write $H - S = H[V \setminus S]$.

Geometry. Given $V \subseteq S^{n-1}$ and $t \in [-1, 1)$, the *distance- t graph* of V is the graph whose vertex set is V and in which x and y are adjacent if $x^\top y = t$. A graph $G = (V, E)$ is (n, t) -*realizable* if there is an injection $f: V \rightarrow S^{n-1}$ such that $f(x)^\top f(y) = t$ for every $xy \in E$. If $xy \notin E$, then there is no constraint on $f(x)^\top f(y)$.

An $(n-1)$ -*sphere* is a translated and scaled copy of S^{n-1} . Let S be an $(n-1)$ -sphere S with radius r , and let $k \leq n-1$. A *great k -sphere of S* is a k -sphere with radius r contained in S . A great k -sphere of S^{n-1} is then the intersection of S^{n-1} with a $(k+1)$ -dimensional linear subspace of $\mathbb{R}^n$. A great 1-sphere is called a *great circle*.

An n -*simplex* is the convex hull of $n+1$ affinely independent points in Euclidean space. An n -simplex is often identified with its set of $n+1$ vertices. A *regular n -simplex* with *inner product t* is a regular simplex whose vertices all lie on a unit sphere and have pairwise inner product t . The t -distance graph of a regular n -simplex with inner product t is isomorphic to K_{n+1} , the complete graph on $n+1$ vertices. Conversely, for all $n \geq k$ and $d > 0$, the graph K_{k+1} is (n, t) -realizable, and its realization is a regular k -simplex with inner product t .

The *circumsphere* of an n -simplex in $\mathbb{R}^n$ is the unique sphere that goes through all the vertices of the simplex [44, Section 1.4]. For $S \subseteq \mathbb{R}^n$, let $\text{Aff } S$ denote the affine span of S . In general, if S is a k -simplex contained in $\mathbb{R}^n$, define its circumsphere as the circumsphere of S in $\text{Aff } S$. With this definition, the circumsphere of a k -simplex in $\mathbb{R}^n$ is unique.

2. The block moment hierarchy for almost-equiangular sets

For integer $n \geq 2$ and $t \in [-1, 1)$, let $H = H(n, t)$ be the 3-uniform hypergraph whose vertex set is S^{n-1} and in which three distinct points $x, y,$

and z form an edge if $t \notin \{x^\top y, x^\top z, y^\top z\}$. Then the t -almost-equiangular sets are exactly the independent sets of H , and so the goal is to compute the independence number $\alpha(H) = \alpha(n, t)$ of H .

Recall the three-point bound from the block moment hierarchy for graphs from Chapter 7. In Section 7.6 we saw that the standard topology puts unnecessary restrictions on continuous functions on $\mathcal{I}_3$. The solution suggested there is to replace the standard topology by the disjoint union topology $\bigsqcup_{i=0}^3 \mathcal{I}_{=i}$ where each $\mathcal{I}_{=i}$ has the standard topology. This suffices for the definition of a block moment hierarchy. Thus, instead of optimizing over $C(\mathcal{I}_3)$, we optimize over $\bigoplus_{i=0}^3 C(\mathcal{I}_{=i})$.

Let

$$B_3 : C(\mathcal{I}_1^3) \rightarrow \bigoplus_{k=0}^3 C(\mathcal{I}_{=k}),$$

be defined by

$$(B_3 A)(I) = \sum_{\substack{Q, J, J' \in \mathcal{I}_1 \\ J \cup J' \cup Q = I}} A(J, J', Q)$$

for all $A \in C(\mathcal{I}_3)$ and $I \in \mathcal{I}_3$. Then, $B_r A$ is continuous on each $\mathcal{I}_{=i}$ by exactly the same argument with which we proved Theorem 7.3. Define the optimization problem $\text{blockM}_3(H(n, t))$ as

$$\begin{aligned} \text{blockM}_3(H(n, t)) &= \sup \nu(\mathcal{I}_{=1}) \\ \nu(\{\emptyset\}) &= 1, \\ B_3^* \nu &\in M_{\text{sym}}(\mathcal{I}_1^3)_{\succeq 0}, \\ \nu &\in M(\mathcal{I}_3)_{\succeq 0}. \end{aligned}$$

we will in fact implement its dual, which is

$$\begin{aligned} \min (B_3 A)(\emptyset) \\ (B_3 A)(\{x\}) &\leq -1 \quad \text{for all } x \in S^{n-1}, \\ (B_3 A)(S) &\leq 0 \quad \text{for all } S \in \mathcal{I}_3 \text{ with } |S| \geq 2, \\ A \in C(\mathcal{I}_1^3) &\text{ is slice positive.} \end{aligned} \tag{44}$$

Recall that A being slice positive means that for every $Q \in \mathcal{I}_1$ the kernel $(S, T) \mapsto A(S, T, Q)$ is positive semidefinite.

One benefit of using the minimization formulation (44) is that every feasible solution gives an upper bound on $\alpha(n, t)$. Contrast this with the situation for Witsenhausen's problem in Chapter 6, where we had to go to great lengths to obtain and verify an upper bound.

THEOREM 8.5. *If A is a feasible solution of (44), then $\alpha(H(n, t)) \leq (B_3 A)(\emptyset)$.*

Proof. Let $I \subseteq S^{n-1}$ be an independent set of $H(n, t)$. On the one hand,

$$\begin{aligned} \sum_{\substack{J \subseteq I \\ |J| \leq 3}} (B_3 A)(J) &= \sum_{\substack{J \subseteq I \\ |J| \leq 3}} \sum_{\substack{S, T, Q \in \mathcal{I}_1 \\ S \cup T \cup Q = J}} A(S, T, Q) \\ &= \sum_{\substack{S, T, Q \subseteq I \\ |S|, |T|, |Q| \leq 1}} A(S, T, Q) \\ &\geq 0, \end{aligned}$$

where the last inequality follows from A being positive semidefinite.

On the other hand,

$$\sum_{\substack{J \subseteq I \\ |J| \leq 3}} (B_3 A)(J) = (B_3 A)(\emptyset) + \sum_{x \in I} (B_3 A)(\{x\}) + \sum_{\substack{J \subseteq I \\ |J| \geq 2}} (B_3 A)(J) \leq (B_3 A)(\emptyset) - |I|,$$

whence $|I| \leq (B_3 A)(\emptyset)$. $\square$

We will now see how to use a semidefinite programming solver to obtain feasible solutions to (44). As in Chapter 6, we exploit the action of $O(n)$ to greatly reduce the size of the program. Denote the Haar measure on $O(n)$ by μ , and assume $\mu(O(n)) = 1$. The orthogonal group $O(n)$ acts on S^{n-1} by rotation. Extend this action to $\mathcal{I}_1$ by acting trivially on $\emptyset$. The induced action on tensors is the diagonal action. See Section 3.1 and the appendix for more details on group actions. Simplify notation by identifying $\mathcal{I}_1$ with $\{\emptyset\} \cup S^{n-1}$, so below $x \in \mathcal{I}_1$ is either $\emptyset$ or an element of S^{n-1} .

Any feasible solution of (44), and in particular any $O(n)$ -invariant feasible solution, gives an upper bound for $\alpha(H)$, where $H = H(n, t)$. Moreover, nothing is lost by restricting (44) to invariant solutions. Indeed, every rotation in $O(n)$ is an automorphism of H , and the objective of (44) is preserved under this action. It follows that, if A is a feasible solution of (44), then

$$\text{Av}_{O(n)}(x, y, z) = \int_{O(n)} A(Tx, Ty, Tz) d\mu(T),$$

is an $O(n)$ -invariant feasible solution providing the same bound as A .

In Chapter 6 we saw how to parametrize continuous slice-positive 3-tensors on S^{n-1} by spherical harmonics. Since we are working on $\mathcal{I}_1$, we need to slightly extend this parametrization to account for the empty set. For a full discussion of the parametrization of such tensors on $\mathcal{I}_k$, see the thesis [77].

Consider an $O(n)$ -invariant slice-positive function $A: \mathcal{I}_1^3 \rightarrow \mathbb{R}$. The kernel $K_\emptyset: \mathcal{I}_1^2 \rightarrow \mathbb{R}$ defined by

$$K_\emptyset(x, y) = A(x, y, \emptyset)$$

for $x, y \in \mathcal{I}_1$ is positive semidefinite and $O(n)$ -invariant.

Fix $e \in S^{n-1}$, then for all $z \in S^{n-1}$ there is a $T \in O(n)$ such that $Tz = e$, so $A(x, y, z) = A(Tx, Ty, e)$. Let $K_e: \mathcal{I}_1^2 \rightarrow \mathbb{R}$ be the kernel such that

$$K_e(x, y) = A(x, y, e).$$

This kernel is positive semidefinite and invariant under the *stabilizer subgroup* of e , namely the subgroup $\text{Stab}(e)$ of $O(n)$ that fixes e .

It follows that an $O(n)$ -invariant slice-positive function $A \in C(\mathcal{I}_1^3)$ can be represented by two positive-semidefinite kernels in $C(\mathcal{I}_1^2)$, namely $K_\emptyset$ and K_e , the kernel $K_\emptyset$ being $O(n)$ -invariant and the kernel K_e being $\text{Stab}(e)$ -invariant. The correspondence is simply

$$\begin{aligned} A(x, y, \emptyset) &= K_\emptyset(x, y) \quad \text{and} \\ A(x, y, z) &= K_e(Tx, Ty), \end{aligned}$$

where T is any element of $O(n)$ such that $Tz = e$. It follows from the invariance of K_e that A is well-defined, since if $T_1z = T_2z = e$, then $T_2T_1^{-1} \in \text{Stab}(e)$ and $K_e(T_1x, T_1y) = K_e(T_2x, T_2y)$.

In Section 1 of Chapter 6, we saw that Schoenberg's theorem [93] characterizes $O(n)$ -invariant positive-semidefinite kernels on S^{n-1} in terms of Gegenbauer polynomials, and that a theorem of Bachoc and Vallentin [6] characterizes $\text{Stab}(e)$ -invariant positive-semidefinite kernels on S^{n-1} using multivariate Gegenbauer polynomials. Both characterizations can be easily adapted to kernels on $\mathcal{I}_1$. For this the following lemma is useful.

LEMMA 8.6. *Let V be a topological space, $f_1, \dots, f_N: V \rightarrow \mathbb{R}$ be continuous functions, and for $x, y \in V$ consider the $N \times N$ matrix such that*

$$Z(x, y)_{ij} = f_i(x)f_j(y).$$

If $A \in \mathbb{R}^{N \times N}$ is positive semidefinite, then the kernel $K: V^2 \rightarrow \mathbb{R}$ such that

$$K(x, y) = \langle A, Z(x, y) \rangle$$

is positive semidefinite.

Proof. Let $x_1, \dots, x_k \in V$ and take $u \in \mathbb{R}^k$. Since A is positive semidefinite, the matrix $A \otimes uu^T$ is also positive semidefinite; its rows and columns are indexed by $I = \{1, \dots, N\} \times \{1, \dots, k\}$. Setting $g(i, k) = f_i(x_k)$, it follows

that

$$\begin{aligned} \sum_{k,l=1}^k K(x_k, x_l) u_k u_l &= \sum_{k,l=1}^k u_k u_l \sum_{i,j=1}^N A_{ij} f_i(x_k) f_j(x_l) \\ &= \sum_{(i,k),(j,l) \in I} (A \otimes uu^\top)_{(i,k),(j,l)} g(i,k) g(j,l) \\ &\geq 0, \end{aligned}$$

as wanted. $\square$

Start with $K_\emptyset$. Let P_k^n denote the Jacobi polynomial of degree k with parameters $\alpha = \beta = (n-3)/2$ normalized so $P_k^n(1) = 1$. For $k \geq 1$, let $Z_k^\emptyset: \mathcal{I}_1^2 \rightarrow \mathbb{R}$ be such that

$$Z_k^\emptyset(x, y) = \begin{cases} P_k^n(x^\top y) & \text{if } x, y \in S^{n-1}; \\ 0 & \text{otherwise.} \end{cases}$$

Let $Z_0^\emptyset: \mathcal{I}_1^2 \rightarrow \mathbb{R}^{2 \times 2}$ be such that, for $x, y \in S^{n-1}$,

$$\begin{aligned} Z_0^\emptyset(\emptyset, \emptyset) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, & Z_0^\emptyset(x, \emptyset) &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\ Z_0^\emptyset(\emptyset, x) &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, & Z_0^\emptyset(x, y) &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

It follows from the addition formula for Gegenbauer polynomials [3, §9.6] that for every $k > 0$ the kernel $(x, y) \mapsto Z_k^\emptyset(x, y)$ is positive semidefinite. From Lemma 8.6 it follows that if $A \in \mathbb{R}^{2 \times 2}$ is positive semidefinite, then the kernel $(x, y) \mapsto \langle A, Z_0^\emptyset(x, y) \rangle$ is positive semidefinite. So, for every $d \geq 0$, any kernel of the form

$$(x, y) \mapsto \langle A_0^\emptyset, Z_0^\emptyset(x, y) \rangle + \sum_{k=1}^d a_k Z_k^\emptyset(x, y) \quad (45)$$

for positive-semidefinite $A_0^\emptyset \in \mathbb{R}^{2 \times 2}$ and nonnegative numbers a_k is $O(n)$ -invariant and positive semidefinite. The only difference with Schoenberg's theorem is that $Z_0^\emptyset$ is then a single number.

Next consider K_e . Recall, for $k \geq 0$, the polynomials

$$Q_k^n(u, v, t) = (1 - u^2)^{k/2} (1 - v^2)^{k/2} P_k^{n-1} \left(\frac{t - uv}{(1 - u^2)^{1/2} (1 - v^2)^{1/2}} \right);$$

this is a polynomial on u, v , and t of degree $2k$, as defined in §6.1. Also recall the matrices

$$(Y_{k,d}^n)_{i,j}(u, v, t) = u^i v^j Q_k^{n-1}(u, v, t);$$

here, we allow $d = \infty$, so that the above is defined for any integers i and j , and we denote the corresponding infinite matrix by Y_k^n .

For $k > 0$ and $x, y \in S^{n-1}$, let $Z_k^e(x, y)$ be the infinite matrix indexed by integers $i, j \geq 0$ such that

$$Z_k^e(x, y)_{i,j} = (Y_k^n)_{i,j} (e^\top x, e^\top y, x^\top y).$$

Note that this is a polynomial on $e^\top x$, $e^\top y$, and $x^\top y$ of degree $i + j + 2k$. If $x = \emptyset$ or $y = \emptyset$, set $Z_k^e(x, y)_{ij} = 0$.

For integer $i \geq 0$, let $f_i: \mathcal{I}_1 \rightarrow \mathbb{R}$ be such that

$$f_i(x) = \begin{cases} 0 & \text{if } x = \emptyset; \\ (e^\top x)^i & \text{otherwise.} \end{cases}$$

Let $f_\emptyset: \mathcal{I}_1 \rightarrow \mathbb{R}$ be such that $f_\emptyset(\emptyset) = 1$ and $f_\emptyset(x) = 0$ if $x \in S^{n-1}$. Define the infinite matrix $Z_0^e(x, y)$, indexed by $U = \{\emptyset\} \cup \{i \in \mathbb{Z} : i \geq 0\}$, by setting

$$Z_0^e(x, y)_{\alpha\beta} = f_\alpha(x) f_\beta(y)$$

for $\alpha, \beta \in U$.

Let A be a positive-semidefinite matrix indexed by a finite set of nonnegative integers. For $k > 0$, Bachoc and Vallentin [6] showed that the kernel

$$(x, y) \mapsto \langle A, Z_k^e(x, y) \rangle \quad (46)$$

on S^{n-1} is positive semidefinite. In the trace inner product in (46), the matrix Z_k^e is truncated, that is, only the finite submatrix corresponding to the rows and columns of A is considered. From this it immediately follows that the kernel (46) is positive semidefinite as a kernel over $\mathcal{I}_1$ as well.

As for $k = 0$, if A is a positive-semidefinite matrix indexed by a finite subset of the index set U , then the kernel $(x, y) \mapsto \langle A, Z_k^e(x, y) \rangle$ is positive semidefinite, as follows directly from Lemma 8.6. So, if $A_0^e, \dots, A_d^e$ are positive-semidefinite matrices, with A_0^e indexed by a subset of U and A_k^e indexed by a subset of the nonnegative integers for $k > 0$, then

$$K_e(x, y) = \sum_{k=0}^d \langle A_k^e, Z_k^e(x, y) \rangle \quad (47)$$

is positive semidefinite and, by construction, $\text{Stab}(e)$ -invariant. Every $\text{Stab}(e)$ -invariant positive-semidefinite continuous kernel K_e can be uniformly approximated by kernels with the above expression, see for example the appendix of [58].

With this, it is possible to express the function $A \in C(\mathcal{I}_1^3)$ of (44) in terms of polynomials. Here, d in (45) and (47) is fixed and the matrices A_k in (47) are truncated appropriately to bound the total degree of the polynomials used. The constraints of (44) are modeled as polynomial constraints using sums of squares. In this way, (44) can be solved numerically with the computer, and

solutions can even be found analytically. Both approaches are discussed in Section 3.

3. Upper bounds from the three-point bound

As shown in Section 2, the bound (44) can be expressed in terms of a polynomial optimization problem once d is fixed in (45) and (47) and the $Z_k^\emptyset$ and Z_k^e matrices are truncated to finite matrices.

So implemented, the three-point bound (44) gives particularly good results for $t \leq 0$. Figure 4 shows a plot of this bound as a function of $t \in [-1, 0]$; it was computed by a Julia program using the package `ClusteredLowRankSolver.jl` [59]. These are numerical results of very high precision that can be turned into rigorous results with some effort. The Julia package `AlmostEquiangular.jl`, contained in the arXiv supplement to [4], includes a function to compute the three-point bound.

Using `ClusteredLowRankSolver.jl` [59] and its rounding routine [29], it is possible to obtain a rational analytic solution for fixed dimension $n \geq 3$ and for inner products 0 and $-1/n$. At these points the bound is exactly equal to the maximum size of an almost-equiangular set. These solutions can then be interpolated to obtain a rational function in n and t that gives an upper bound for the size of a t -almost-equiangular set in S^{n-1} for $t \in [-1, 0]$, leading to Theorem 8.1.

A union of two disjoint regular $(n-1)$ -simplices in S^{n-1} gives a 0-almost-equiangular set with $2n$ points; Rosenfeld [91] showed that this construction is optimal. A union of two disjoint regular n -simplices in S^{n-1} gives a $(-1/n)$ -almost-equiangular set with $2(n+1)$ points; Bezdek and Lángi [14] showed that this construction is optimal. The bound of Theorem 8.1 is sharp in both cases, providing a new proof of the optimality of these constructions.

Proof of Theorem 8.1. The proof of the theorem is by exhibiting a solution to the three-point bound that has the objective value in the statement. To keep the solution as simple as possible, use a degree-0 kernel $K_\emptyset$ and a degree-4 kernel K_e . Thus, the set of positive-semidefinite variables is $A_\emptyset^\emptyset$ and A_k^e with $0 \leq k \leq 2$.

Let

$$p = \frac{8n^2t^4(2n-1) - 9n^2t^3(n-1) + (2nt^2 - 3t + 4)(7n+1)}{2(1-t)(1+7n-2n^2t^3(2n-1))}$$

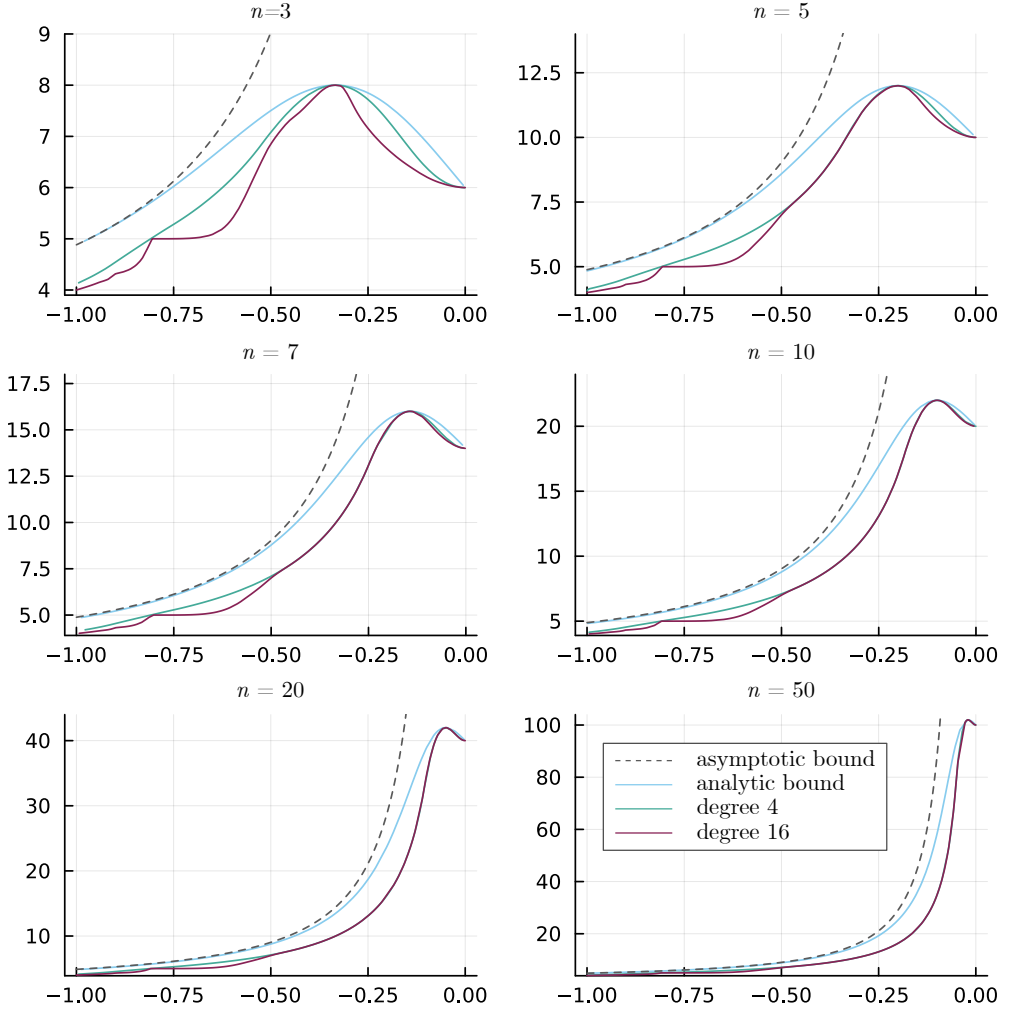


FIGURE 4. The numeric solutions of degree 4 and 16, and the analytic (Theorem 8.1) and asymptotic solutions to the bound (44). The asymptotic bound is the limit as $n \rightarrow \infty$ of the analytic bound.

and

$$\begin{aligned}
A_0^\emptyset &= \begin{pmatrix} \frac{n(1-t)^2}{2(nt^2+1)}p^2 & -\frac{1}{2}p \\ * & \frac{nt^2+1}{2n(1-t)^2} \end{pmatrix}, \\
A_0^e &= \begin{pmatrix} p & -\frac{1}{4n(1-t)^2} - \frac{t^2}{(1-t)^2} & \frac{3t}{2(1-t)^2} & -\frac{3}{4(1-t)^2} \\ * & \frac{1}{4n(n-1)(1-t)^3} - \frac{t^3}{2(1-t)^3} & \frac{3t^2}{4(1-t)^3} & -\frac{n+1}{8n(n-1)(1-t)^3} \\ * & * & -\frac{3t}{2(1-t)^3} & 0 \\ * & * & * & \frac{2n-1}{4(n-1)(1-t)^3} \end{pmatrix}, \\
A_1^e &= \begin{pmatrix} 0 & 0 \\ * & \frac{n+1}{2n(1-t)^3} \end{pmatrix}, \\
A_2^e &= \begin{pmatrix} \frac{n-2}{4n(n-1)(1-t)^3} \end{pmatrix}.
\end{aligned}$$

The *s indicate that the entries are determined by the symmetry of the matrices.

All matrices above, except for A_0^e , can be checked by hand to be positive semidefinite in the domain given by $n \geq 3$ and $t \in [-1, 0]$. To check that A_0^e is positive semidefinite in the required domain, first decompose it as $A_0^e = LDL^\top$, where L and D are matrices of rational functions on n and t and D is diagonal, and then check that the diagonal entries of D are nonnegative in the domain.

These diagonal entries are rational functions, which can be rigorously checked to be nonnegative by a sum-of-squares approach. The arXiv supplement to [4] contains the Julia package `AlmostEquiangular.jl`, which provides sum-of-squares certificates for the nonnegativity of the diagonal entries of D . The same package also provides a sum-of-squares certificate for the inequality $f(n, t) \leq (16t - 9)^2/(128t^2)$.

The Julia package also checks if, for the corresponding function $A \in C(\mathcal{I}_1^3)$,

$$\begin{aligned}
B_3A(\{x\}) &= -1, \\
B_3A(\{x, y\}) &= 0 && \text{for all } x \neq y, \text{ and} \\
B_3A(\{x, y, z\}) &= \frac{3(t - x^\top z)(t - y^\top z)(t - x^\top y)}{(t - 1)^3} && \text{for all } x, y, \text{ and } z \text{ distinct.}
\end{aligned}$$

In particular, if $t \in \{x^\top z, y^\top z, x^\top y\}$, then $B_3A(\{x, y, z\}) = 0$. $\square$

The solution constructed in the proof above can in principle be improved; the issue is to get a good compromise between simplicity and quality. For instance, by forcing some matrix entries to be zero as done above, it becomes possible to find a simple rational expression as given in the theorem.

4. Realizability of anti-triangle-free graphs

A graph is *anti-triangle free* if its complement is triangle free. This is equivalent to saying that every triple of vertices contains an edge. The distance graphs of almost-equiangular sets are anti-triangle free and, conversely, realizable anti-triangle-free graphs give almost-equiangular sets. Hence, to construct good almost-equiangular sets, one has to show that given anti-triangle-free graphs are realizable.

Recall the definition of realizability from Section 1. The goal of this section is to determine whether certain anti-triangle-free graphs are (n, t) -realizable. A construction of interest is the (k, l) -spindle, denoted by $S(k, l)$ with $k, l \geq 1$, defined later in this section, of which the Moser spindle is a special case. In order to bound the inner products at which $S(k, l)$ is realizable, and to offer some tools for other calculations, it is useful to derive realizability of some commonly appearing subgraphs of the spindle, namely the simplex and the rhombus.

The simplex. A nice reference for simplex geometry is Fiedler [44]; see in particular Theorem 4.5.1 of this book for the following facts. The inner products of distinct vertices of a regular n -simplex inscribed in S^{n-1} is $-1/n$. So K_{n+1} is (n, t) -realizable if and only if $t = -1/n$.

If $k < n$, then K_{k+1} is (n, t) -realizable if and only if $t \geq -1/k$. Indeed, the circumradius of a regular k -simplex with inner product t is

$$r_k(t) = \sqrt{\frac{(1-t)k}{k+1}}$$

and the circumsphere of a k -simplex is a $(k-1)$ -sphere. For $k < n$, the sphere S^{n-1} contains a $(k-1)$ -sphere of every radius less than or equal to 1, so K_{k+1} is (n, t) -realizable if and only if $r_k(t) \leq 1$. This happens if and only if $t \geq -1/k$.

The $k+1$ vertices of a regular k -simplex on S^{n-1} are by definition affinely independent, and so a regular k -simplex contains at least k linearly independent points. If $t = -1/k$, then $r_k(t) = 1$, and the circumsphere is a great sphere, which lies on a linear subspace of dimension k . However, if $k < n$ and $t > -1/k$, then $r_k(t) < 1$, and so the linear span of the k -simplex has dimension $k+1$. In this case, the vertices of the k -simplex are linearly independent.

The rhombus. A useful subgraph of a spindle is the union of two complete graphs on $k+1$ vertices that have exactly k vertices in common. This is the distance graph of a pair of regular k -simplices that share exactly one

facet. Alternatively, it is the complete graph K_{k+2} with one edge removed. Call this graph a k -rhombus. By the previous paragraph, necessary conditions for realizability are $k \leq n$, $t = -1/k$ if $k = n$, and $t \geq -1/k$ otherwise.

In what follows, let R be a k -rhombus that is the union of two instances of K_{k+1} , denoted by Σ_1 and Σ_2 , let e be the unique vertex of $\Sigma_1 - V(\Sigma_2)$, let p be the unique vertex in $\Sigma_2 - V(\Sigma_1)$, and let $B = V(\Sigma_1) \cap V(\Sigma_2)$. Refer to $R[B]$ as the *base* of the rhombus. It is an instance of K_k . Up to orthogonal transformations, an (n, t) -realization of $R[B]$ is uniquely determined, so assume its vectors are known and denote the realization by B as well. The following lemma is comparable to [7, Lemma 7].

LEMMA 8.7. *With e, p as above, the k -rhombus is (n, t) -realizable if and only if $k \leq n - 1$ and $t > -1/k$. If these conditions hold, then e and p lie on an $(n - k - 1)$ -sphere of radius $\sqrt{1 - 2kt^2/((k - 1)t + 1)}$. In particular, let k be an integer such that $k \leq n - 1$ and $t > -1/k$, and*

$$\tau = \frac{2kt^2}{(k - 1)t + 1} - 1.$$

If $k < n - 1$, then $e^\top p \geq \tau$, and if $k = n - 1$, then $e^\top p = \tau$.

Conversely, if $k < n - 1$ and $t' \in [\tau, 1)$, then there exists an (n, t) -realization of the k -rhombus in which $e^\top p = t'$. If $k = n - 1$, then the points e and p are uniquely determined.

Proof. The k -rhombus with base B has a subgraph isomorphic to K_{k+1} , and so necessary conditions for realizability are $t \geq -1/k$ and $k \leq n$. Assume that these hold. If $k = n$, the vectors in B already determine a full rank system, so then p will coincide with e . Consequently, another necessary condition is $k \leq n - 1$.

Since $t \geq -1/k$, the k -rhombus is realizable if and only if the affine space

$$A = \{x \in \mathbb{R}^n : x^\top b = t \text{ for all } b \in B\}$$

intersects S^{n-1} in more than one point, that is, if and only if

$$\inf \{ \|a\| : a \in A \} < 1;$$

this infimum is attained in A .

Let U be the linear span of B and let W be its orthogonal complement. Then the shortest vector a_0 in A is in U . Indeed, if $a_0 = \sum_{b \in B} \lambda_b b + w \in A$ with $w \in W$, then by orthogonality $\|a_0\|^2 = \|\sum_{b \in B} \lambda_b b\|^2 + \|w\|^2$. Translating by a vector orthogonal to U does not change the inner product with any of the elements in B . So, if $\|a_0\|$ is minimal, then $w = 0$.

All that is left is to calculate the coefficients λ_b . Since A is convex, a_0 is the unique shortest vector. Because $b^\top b' = t$ for all distinct $b, b' \in B$, applying a permutation to the coefficients gives another vector in A with the same norm. By uniqueness, this forces all λ_b to have the same value λ . For every $b \in B$,

$$t = a_0^\top b = \lambda \sum_{b' \in B} b'^\top b = \lambda((k-1)t + 1),$$

so $\lambda = t/((k-1)t + 1)$ and

$$\|a_0\|^2 = \frac{kt^2}{(k-1)t + 1}.$$

Since $t \geq -1/k$, it follows that $\|a_0\| < 1$ if and only if $kt^2 - (k-1)t - 1 < 0$. As a polynomial in t it has roots 1 and $-1/k$, so the k -rhombus is realizable if and only if $-1/k < t < 1$.

The intersection of A with S^{n-1} gives an $(n-k-1)$ -sphere S whose radius r is $\sqrt{1 - \|a_0\|^2} = \sqrt{1 - kt^2/((k-1)t + 1)}$. Any two distinct points on S are valid realizations of p and e . If τ is the minimum possible inner product between points on S , then $2r = \sqrt{2(1-\tau)}$, so

$$\tau = 1 - 2r^2 = \frac{2kt^2}{(k-1)t + 1} - 1. \quad \square$$

In Lemma 8.7, the inner product t does not depend on the embedding dimension, something that often happens for these types of constructions.

The spindle. The (k, l) -*spindle*, notation $S(k, l)$, is described as follows: let R_1 be a k -rhombus; say e and p_1 are the vertices of its unique nonedge. Attach at e an l -rhombus R_2 with nonedge ep_2 , so $V(R_1) \cap V(R_2) = \{e\}$. Finally, add the edge p_1p_2 . Figure 5 shows several spindles. For $k \geq 1$, the spindle $S(k, k)$ is called the k -*Moser spindle*, denoted by $MS(k)$. If $k \leq l$, then $S(k, l) \subseteq MS(l)$.

The (k, l) -spindle is an anti-triangle-free graph of order $k+l+3$. The Moser spindle in particular is well studied. For example, the spindle $MS(n-1)$ was already pointed out by Bezdek and Lánci [14] as a t -almost-equiangular set for t close to 1. However, they did not attempt to calculate for which t the graph is realizable, and did not consider the case of negative t or $k \neq l$. This is done in the following theorem.

THEOREM 8.8. *If $k \geq 1$ and $i \in \{1, 2, 3\}$, then all roots $t_{k,i}$ of the polynomial*

$$8k^2t^3 - (k^2 - 10k + 1)t^2 - 2(k-1)t - 1$$

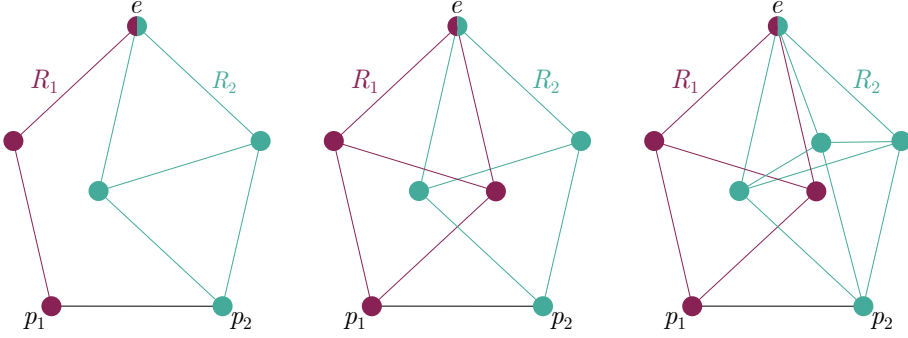


FIGURE 5. From left to right: S(1,2), MS(2), and S(2,3).
The R_i are indicated with their respective colors.

with respect to t are real and can be ordered such that $t_{k,1} \leq t_{k,2} < 0 < t_{k,3}$. The (k,l) -spindle is (n,t) -realizable if and only if $k, l \leq n-1$ and $t \in [-1,1)$ satisfy

$$t = (-1/4)(1 \pm \sqrt{5}) \quad \text{if } n = 2 \text{ and } k = l = 1, \quad (48)$$

$$t_{k,1} \leq t \leq t_{k,2} \quad \text{or} \quad t_{k,3} \leq t \quad \text{if } n > 2 \text{ and } k = l = n-1, \quad (49)$$

$$t_{k,1} \leq t \quad \text{if } n > 2 \text{ and } k = l < n-1, \quad \text{or} \quad (50)$$

$$-1/l < t \quad \text{if } n > 2 \text{ and } k < l \leq n-1. \quad (51)$$

The following simple lemma does a lot of the work in the proof of Theorem 8.8.

LEMMA 8.9. *If S_1 and S_2 are subsets of S^{n-1} that are invariant under the subgroup of $O(n)$ that stabilizes a point e and if $\inf \{x^\top y : x \in S_1, y \in S_2\}$ is attained by points $p_1 \in S_1$ and $p_2 \in S_2$, then e , p_1 , and p_2 lie on a great circle C . Moreover, if $f \in C$ is orthogonal to e and if $p_1, p_2 \neq \pm e$, then $f^\top p_1$ and $f^\top p_2$ have opposite signs.*

Proof. If p_1 or p_2 is $\pm e$, then the result is clear. So assume $p_1, p_2 \neq \pm e$.

Let $U = \text{span}\{e, p_1\}$ and let f be a unit vector in U orthogonal to e such that $f^\top p_1 > 0$. Write $p_1 = \alpha e + \beta f$ and $p_2 = \lambda e + \kappa f + w$ with $w \in U^\perp$, so $|\kappa| \leq \sqrt{1 - \lambda^2}$. By invariance under the stabilizer of e , any point p'_2 on the sphere with $p_2^\top e = p'^\top_2 e$ is also in S_2 . Let $p'_2 = \lambda e - \sqrt{1 - \lambda^2} f \in S_2$. Then

$$p_1^\top p'_2 = \alpha\lambda - \beta\sqrt{1 - \lambda^2} \leq \alpha\lambda + \beta\kappa = p_1^\top p_2.$$

It follows that $w = 0$ and that $f^\top p_2 = \kappa < 0$, as wanted. $\square$

Proof of Theorem 8.8. Let $n \geq 2$ and $1 \leq k \leq l \leq n-1$ be integers and let t be in $(-1/l, 1)$. A (k, l) -spindle contains the union of a k - and an l -rhombus that intersect in a single point. Let R_1 be the k -rhombus with unique nonedge ep_1 and R_2 the l -rhombus with unique nonedge ep_2 , so $V(R_1) \cap V(R_2) = \{e\}$. A necessary and sufficient condition for realizability is that there are realizations of R_1 and R_2 such that $p_1^\top p_2 = t$.

Let S_i be the set of all possible images of p_i under (n, t) -realizations of R_i that map e to the north pole $(1, 0, \dots, 0)$, that is,

$$S_i = \{ f(p_i) : f \text{ is an } (n, t)\text{-realization of } R_i \text{ such that } f(e) = (1, 0, \dots, 0) \}.$$

Let

$$\tau_1 = \frac{2kt^2}{(k-1)t+1} - 1 \quad \text{and} \quad \tau_2 = \frac{2lt^2}{(l-1)t+1} - 1. \quad (52)$$

If $n > 2$, then Lemma 8.7 guarantees the existence of an (n, t) -realization of R_1 with $e^\top p_1 = \tau_1$. By rotating the realization, e can be placed at the north pole. If $k < n-1$, the lemma similarly guarantees the existence of an (n, t) -realization of R_1 with $e^\top p_1 = t'$ for all $t' \in [\tau_1, 1]$ with e at the north pole. This goes through analogously for R_2 . Since the action of the stabilizer of e in $O(n)$ is transitive on the set of points p that have inner product t' with e for all $t' \in [-1, 1]$, this shows that if $k = n-1$ or $l = n-1$, the corresponding S_i is

$$S_i = \{ p \in S^{n-1} : e^\top p = \tau_i \}$$

and if $k < n-1$ or $l < n-1$, the corresponding S_i is

$$S_i = \{ p \in S^{n-1} : e^\top p \in [\tau_i, 1] \}.$$

In particular, they are invariant under the stabilizer of e in $O(n)$.

Furthermore, if $k \leq l \leq n-1$ and $t > -1/l$, then $\tau_1 \leq \tau_2$ for fixed t , so that $S_2 \subseteq S_1$. It follows that there is ξ such that

$$\{ p^\top q : p \in S_1, q \in S_2 \} = [\xi, 1].$$

Note that ξ is a function of k, l , and t . Given $n > 2$ and k and l , it is then enough to find the values of t for which $\xi \leq t$. Let $q_1 \in S_1$ and $q_2 \in S_2$ be such that $\xi = q_1^\top q_2$. The goal is then to have $q_1^\top q_2 \leq t$. The following simple fact will be useful:

If S_1 and S_2 are arcs of the unit circle S^1 such that the infimum $\inf \{ x^\top y : x \in S_1, y \in S_2 \}$ is attained, then the infimum is attained by an antipodal pair or by endpoints of the arcs. (53)

By Lemma 8.9 it can be assumed that q_1, q_2 , and e all lie on the same great circle C . By (53), either the q_i are endpoints of $S_i \cap C$ or they are antipodal.

If the q_i are endpoints, then $e^\top q_i = \tau_i$. Using Lemma 8.9 again gives

$$E = q_1^\top q_2 = \tau_1 \tau_2 - \sqrt{1 - \tau_1^2} \sqrt{1 - \tau_2^2}.$$

Hence, in this case the spindle is (n, t) -realizable if and only if $E \leq t$.

The q_i are antipodal only if $\tau_1 \leq -\tau_2$. In this case, the spindle is (n, t) -realizable. This gives necessary and sufficient conditions for realizability in the $n > 2$ case.

If $n = 2$, then $k = l = 1$. The sets S_i then each contain only two choices for q_i such that $e^\top q_i = \tau_i$. A necessary and sufficient condition for realizability is then that $q_1^\top q_2 = t$.

To summarize, necessary and sufficient conditions for (n, t) -realizability of the (k, l) -spindle are:

- (i) $E = t$ if $n = 2$;
- (ii) $E \leq t$ if $n > 2$ and $k = l = n - 1$;
- (iii) $E \leq t$ or $\tau_1 \leq -\tau_2$ otherwise.

Recall from (52) that the τ_i are functions of k, l , and t , and hence so is E . The goal is now to determine, for each case above, the values of t for which the conditions hold.

In most of the cases below, one has $k = l$. Then $\tau_1 = \tau_2 =: \tau$, and so

$$E = 2\tau^2 - 1.$$

Plug (52) into the right-hand side above to see that $E \leq t$ if and only if

$$8k^2t^3 - (k^2 - 10k + 1)t^2 - 2(k - 1)t - 1 \quad (54)$$

is nonnegative, with equality when t is a root of the polynomial. In what follows, this and other polynomials considered are seen as polynomials on t only, that is, k is fixed.

CASE (i). If $n = 2$, then $k = l = 1$, and there are only two values of t for which MS(1) is realizable. To see this, factor the polynomial (54) as

$$8t^3 + 8t^2 - 1 = (2t + 1)(4t^2 + 2t - 1).$$

For the root $t = -1/2$, the points p_1 and p_2 coincide with other points in the spindle. The other roots are $t = -(1/4)(1 \pm \sqrt{5})$. These inner products correspond to the pentagon and pentagram. This gives (48).

CASE (ii). If $n > 2$ and $k = l = n - 1$, then (ii) is satisfied if and only if the polynomial (54) has a nonnegative value at t . Its discriminant is positive, so it only has real roots. Denote them by $t_{k,1} \leq t_{k,2} \leq t_{k,3}$. The constant and linear terms are negative, so $t_{k,1} \leq t_{k,2} < 0 < t_{k,3}$. At $t = 0$ the polynomial is negative, thus the polynomial must be nonnegative for $t_{k,1} \leq t \leq t_{k,2}$

and $t \geq t_{k,3}$. So $\text{MS}(n-1)$ is realizable if and only if $t_{k,1} \leq t \leq t_{k,2}$ or $t \geq t_{k,3}$. This establishes (49).

CASE (iii). It remains to consider $n > 2$ and $l < n-1$. The discussion splits into two cases: (a). $k = l$ and (b). $k < l$.

Case (a). If $k = l < n-1$, either one of the conditions in (iii) has to be satisfied. The first one is again equivalent to finding t such that the polynomial (54) is nonnegative, and so a sufficient condition for realizability is $t_{k,1} \leq t \leq t_{k,2}$ or $t \geq t_{k,3}$.

The second condition is $\tau_1 \leq -\tau_2$. Since $\tau_1 = \tau_2 =: \tau$ one has $\tau \leq 0$. From (52), this happens if and only if $g = 2kt^2 - (k-1)t - 1 \leq 0$. This polynomial has a positive and a negative root and is negative at 0. At both roots, (54) is positive. This can be seen by taking the remainder of (54) after division by g , and testing it at a convenient value smaller than the smallest root of g (for example $t = -1/k$), since the remainder is linear and increasing in t . So $\text{MS}(k)$ with $k < n-1$ is realizable if and only if $t_{k,1} \leq t$. This establishes (50).

Case (b). The final case is $n > 2$ and $k < l \leq n-1$. We will see later that it suffices to consider the case $l = k+1$.

So assume $l = k+1$. Let

$$f = 8k^2(k-1)t^4 - (k^3 - 19k^2 + 8k + 4)t^3 - k(3k-14)t^2 - 3(k-1)t - 1.$$

The inequality $E \leq t$ is satisfied if and only if $f \geq 0$.

If $k = 1$, then f is of degree 3. Computing its roots, one gets conditions for the inequality above to be satisfied, obtaining a set of values of t for which the spindle is realizable. Similarly, the condition $\tau_1 \leq -\tau_2$ is satisfied if and only if $t^3 + 3t^2 - t - 1 \leq 0$. This gives another set of values of t for which the spindle is realizable. Taking the union of both sets, one gets the condition $t > -1/2 = -1/l$ for realizability.

If $k > 1$, then f has degree 4 and its discriminant is negative, so it has exactly two real roots $f_1 \leq f_2$. At $t = 0$ it is negative and at $t = 1$ and at $t = -1/(k+1)$ it is positive, hence $-1/(k+1) < f_1 < 0 < f_2 < 1$ and $f \geq 0$ for $-1/(k+1) < t \leq f_1$ and $f_2 \leq t < 1$.

The condition $\tau_1 \leq -\tau_2$ is equivalent to the condition

$$g = (2k^2 - 1)t^3 - (k^2 - 3k - 1)t^2 - (2k - 1)t - 1 \leq 0.$$

By an analysis similar as before, this polynomial has three real roots given as $g_1 \leq g_2 < 0 < g_3$. It is negative at 0, so it is nonpositive for all t such that $t \leq g_1$ or $g_2 \leq t \leq g_3$. The next objective is to show $g_2 \leq f_1 \leq f_2 \leq g_3$, so that the result follows; see Figure 6.

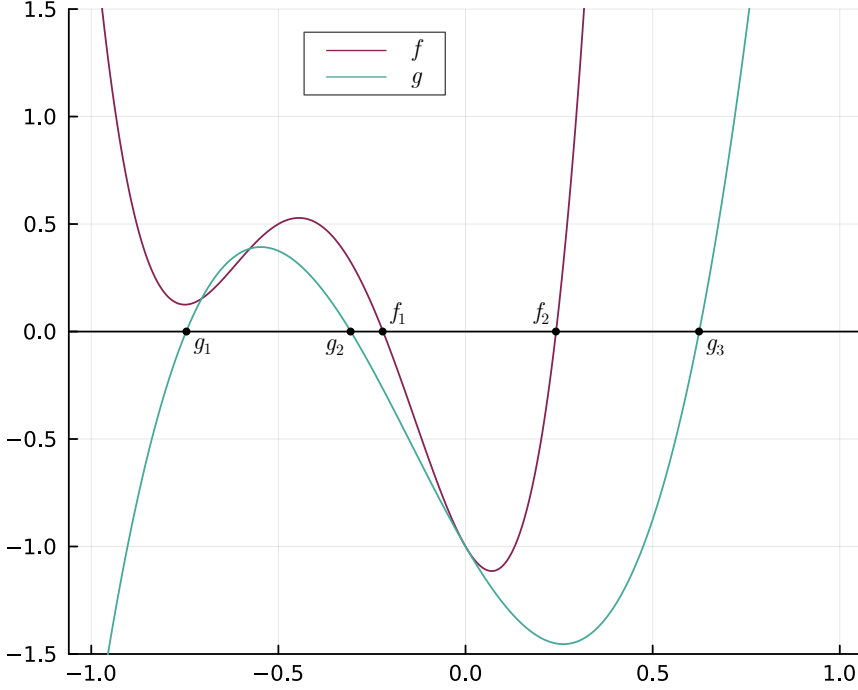


FIGURE 6. A plot of f and g for $k = 2$. The horizontal axis is the inner product t . Clearly $-1/2 < g_2 \leq f_1 < 0 < f_2 \leq g_3$. If $n > 2$, then $S(1, 2)$ is (n, t) -realizable if and only if $t > -1/2$, and $f \geq 0$ or $g \leq 0$. This plot shows that it is (n, t) -realizable if and only if $t > -1/2$.

To determine the order of the roots f_1, f_2, g_1, g_2 , and g_3 , take the remainder r of f after division by g . The remainder has degree 2 and has two real roots; denote the roots of r by $r_1 \leq r_2$. Then f is nonnegative at a g_i if and only if r is. Both roots of r are negative for any $k \geq 2$. Moreover, g is positive at r_1 and r_2 , so they lie between g_1 and g_2 . The coefficient of the quadratic term of r is positive, so it has a global minimum, meaning it is positive for all $t > r_2 > g_1$, so f is positive at g_2 and g_3 . This determines the order of the roots $g_2 \leq f_1 < 0 < f_2 \leq g_3$. The spindle is realizable if $-1/(k+1) < t \leq f_1, f_2 \leq t \leq 1$ and $g_2 \leq t \leq g_3$, so putting all of this together, $S(k, k+1)$ is realizable if and only if $-1/(k+1) < t < 1$.

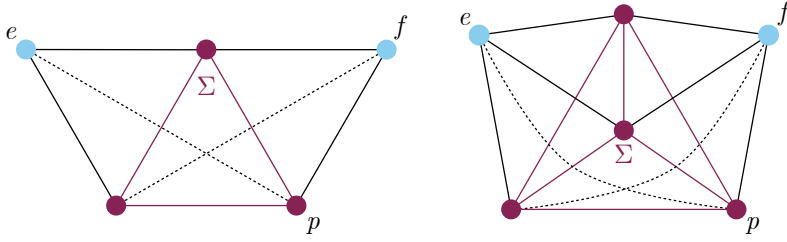


FIGURE 7. The extended 2-rhombus on the left and the extended 3-rhombus on the right. The dotted lines are edges that follow from Lemma 8.7, forcing f to coincide with e .

From $l = k + 1$ all other cases follow. Indeed, $S(k, l)$ with $k < l - 1$ is a subgraph of $S(l - 1, l)$, and so a sufficient condition for realizability is $t > -1/l$, which was already seen to be necessary. This settles (51). $\square$

Some results on non-realizability. To classify almost-equiangular sets in low dimension, it is necessary to show that given anti-triangle-free graphs are not (n, t) -realizable for certain n and t .

The extended rhombus. Let $t = -1/n$ and take two $(n - 1)$ -rhombi, R_1 and R_2 , that intersect in an induced subgraph Σ isomorphic to K_n (see Figure 7). Call this graph an *extended $(n - 1)$ -rhombus*. Let e and p be the endpoints of the unique nonedge of R_1 with $p \in \Sigma$. Let f be the endpoint of the nonedge of R_2 not contained in Σ .

If $t = -1/n$ and $k = n - 1$, then by Lemma 8.7, $e^\top p = -1/n$ in any realization of R_1 . So a realization of R_1 actually forms an n -simplex, and analogously the same holds for R_2 . But then e and f are uniquely determined by Σ , and must coincide, hence the extended $(n - 1)$ -rhombus is not $(n, -1/n)$ -realizable.

The complement of the split k -cycle. Let $k \geq 4$. The *split k -cycle* is the graph W_k on vertices $p_0, \dots, p_{k-1}, q_0, \dots, q_{k-1}$ in which the neighborhood of both p_i and q_i is $\{p_{i-1}, q_{i-1}, p_{i+1}, q_{i+1}\}$ with all indices modulo k (see Figure 8). It is obtained from a k -cycle by splitting each vertex. Deaett proved [34, Theorem 4.11] that the graph $\overline{W_n}$, the complement of W_n , is $(n, 0)$ -realizable.

For even k , the graph W_k is bipartite with parts of size k , since the set of all even-indexed points is independent and so is its complement. This means that $\overline{W_k}$ is a union of two $(k - 1)$ -simplices with some extra edges and therefore does not give a new construction.

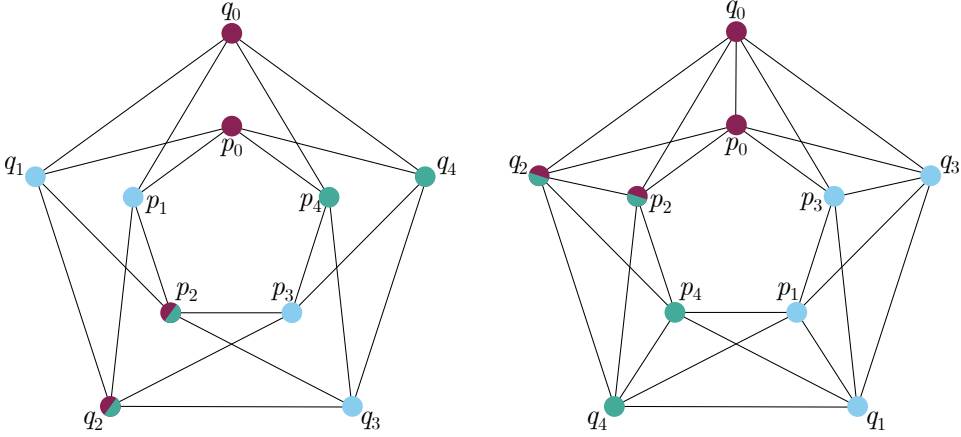


FIGURE 8. On the left the graph W_5 with Σ_0 and Σ_2 indicated by color. In this case, $T_2 = \{p_2, q_2\}$. On the right $\overline{W}_5$ after rearranging the vertices, with Σ_0 and Σ_2 colored as well. The similarity between the two graphs is incidental for $k = 5$.

For $k = 5$, the graph $\overline{W}_5$ is $(5, 0)$ -realizable (see Figure 8). It is the smallest example of an optimal $(n, 0)$ -realizable anti-triangle-free graph that is not a union of two $(n - 1)$ -simplices [34]. Balko, Pór, Scheucher, Swanepoel, and Valtr showed [7, Theorem 2] that $\overline{W}_5$ cannot be embedded in $\mathbb{R}^3$ such that adjacent vertices are at distance 1. Since there are $(4, -1/4)$ -realizable graphs of order 10, a priori $\overline{W}_5$ could be $(4, -1/4)$ -realizable. It turns out, however, that $\overline{W}_k$ with odd $k \geq 5$ is not $(k - 1, t)$ -realizable for any negative t .

Indeed, take W_k with odd $k \geq 5$. The optimization bound (Theorem 8.1) shows that if $n < k$ and $t \in [-1, 0]$, then the maximum cardinality of a t -almost-equiangular set on S^{n-1} is $\leq 2(n + 1)$, with equality only at $t = -1/n$. Since W_k has order $2k \geq 2(n + 1)$, it can only be (n, t) -realizable for $n < k$ when $n = k - 1$ and $t = -1/n$.

Hence, the goal is to show $\overline{W}_k$ is not $(n, -1/n)$ -realizable with $n = k - 1$. So assume that $\overline{W}_k$ is realizable.

In what follows, indices are taken modulo k . Let Σ_i be the set of all vertices p_{i+2j} and q_{i+2j} for $0 \leq j \leq (k - 3)/2$ and set $T_i = \Sigma_{i-2} \cap \Sigma_i$ (see Figure 8).

The Σ_i are independent in W_k and so form $(k - 2)$ -simplices in a realization of $\overline{W}_k$. Take the sets Σ_0 and Σ_2 . Then $\Sigma_0 \setminus T_2$ and $\Sigma_2 \setminus T_2$ both consist of two points that lie in the intersection of hyperplanes defined by the

equations $l^T x = t$ for all $l \in T_2$. The realization of T_2 is a $(k-4)$ -simplex, so by Section 4, T_2 consists of $k-3$ linearly independent vectors and the dimension of the intersection of these hyperplanes is 2. Therefore, p_0, q_0, p_{k-1} , and q_{k-1} are coplanar and lie on a circle C_1 . Repeat this for Σ_1 and Σ_3 to see that p_0, q_0, p_1 , and q_1 are also coplanar and lie on a circle C_2 .

Since $K = \{p_1, q_1, p_{k-1}, q_{k-1}\}$ is a clique in $\overline{W_k}$, it defines a regular tetrahedron, hence its affine span is 3-dimensional, and the circles C_1 and C_2 are distinct. Denote the circumsphere of K by S , which is a 2-sphere. The affine span of $\{p_0, q_0, p_1, q_1, p_{k-1}, q_{k-1}\}$ is also 3-dimensional, since these points lie on two distinct planes intersecting on a line. Then $p_0, q_0 \in \text{Aff } K$. By uniqueness of the circumsphere of a simplex this means p_0 and q_0 also lie on S .

Since Σ_0 can be completed to a regular $(k-1)$ -simplex for $t = -1/(k-1)$ by adding a point on $z \in C_1$, it follows that C_1 is a circumcircle of a regular triangle on S whose vertices are z, p_{k-1} , and q_{k-1} . However, there are only two such regular triangles on S , namely $\{p_{k-1}, q_{k-1}, p_1\}$ and $\{p_{k-1}, q_{k-1}, q_1\}$. So C_1 contains $p_0, q_0, p_{k-1}, q_{k-1}$ and either p_1 or q_1 . By a similar argument, C_2 contains p_0, q_0, p_1, q_1 and either p_{k-1} or q_{k-1} . Then C_1 intersects C_2 in at least four points, a contradiction.

5. Maximum obtuse almost-equiangular sets

Theorem 8.10 below establishes that $\alpha(n, t) \leq 2(n+1)$ for all $t \leq 0$, with equality only for $t = -1/n$. This motivates calling a $(-1/n)$ -almost-equiangular set with $2(n+1)$ points a *maximum obtuse almost-equiangular set*.

The proof of Theorem 8.10 follows a spectral analysis of matrices associated to the Gram matrix of such a set, done by Rosenfeld [91] and Bezdek and Lánci [14]. Further analysis of these matrices gives useful properties of maximum obtuse almost-equiangular sets; they turn out to be spherical 2-designs, and are in bijection with certain symmetric orthogonal matrices.

Finally, this leads to a proof that the only maximum obtuse almost-equiangular set is the double-regular n -simplex for $n = 2, \dots, 5$.

The spectral analysis. Bezdek and Lánci prove in [14] that a t -almost-equiangular subset of S^{n-1} with $t \leq 0$ cannot have more than $2(n+1)$ points by analyzing the eigenvalues of a certain matrix related to the Gram matrix of the set. Their method is revisited here to strengthen their result as follows.

THEOREM 8.10. *If $t \in [-1, 0]$, then $\alpha(n, t) \leq 2(n+1)$, with equality only at $t = -1/n$. The Gram matrix of a maximum obtuse almost-equiangular set has rank n , its only nonzero eigenvalue is $2(1 + 1/n)$, and the all-ones*

vector e is in its kernel. In particular, the barycenter of a maximum obtuse almost-equianangular set is 0.

Proof. Following [14], let U be the Gram matrix of a t -almost-equianangular subset of S^{n-1} of cardinality N , let $C = U - tJ$, and $B = U - tJ - (1-t)I$, where J is the all-ones matrix and I is the identity matrix. The diagonal coefficients of B are 0, hence $\text{Tr} B = 0$. The coefficients of B corresponding to pairs of points with inner product t are equal to 0, hence the set being almost equianangular translates to $B_{ij}B_{jk}B_{ki} = 0$ for all $1 \leq i, j, k \leq N$, whence $\text{Tr}(B^3) = 0$.

These two properties give rise to equations for the eigenvalues of B . Because $\text{rank } C \leq n+1$, the matrix B has at least $N - (n+1)$ eigenvalues equal to $-(1-t)$. If $\lambda_1, \dots, \lambda_{n+1}$ denote the remaining ones, then

$$\sum_{i=1}^{n+1} \lambda_i = (N - n - 1)(1 - t) \quad \text{and} \quad \sum_{i=1}^{n+1} \lambda_i^3 = (N - n - 1)(1 - t)^3.$$

Since $t < 0$, the matrix C is positive semidefinite, and so the smallest eigenvalue of B is $-(1-t)$. Hence, if $y_i = \lambda_i/(1-t)$, then $y_i \geq -1 > -\sqrt{3}$ and the problem

$$\begin{aligned} z^* = \max \quad & \sum_{i=1}^{n+1} y_i \\ & \sum_{i=1}^{n+1} y_i - y_i^3 = 0, \\ & y_i \geq -\sqrt{3} \quad \text{for } i = 1, \dots, n+1 \end{aligned} \tag{55}$$

gives an upper bound for $N - n - 1$.

Let

$$L(y) = \sum_{i=1}^{n+1} y_i + (1/2) \left(\sum_{i=1}^{n+1} y_i - y_i^3 \right) = \sum_{i=1}^{n+1} (3/2)y_i - (1/2)y_i^3$$

and

$$d^* = \max \{ L(y) : y_i \geq -\sqrt{3} \text{ for all } i \}. \tag{56}$$

If (56) has an optimal solution y^* that is feasible for (55), then it is also optimal for (55). Conversely, if $z^* = d^*$ and y^* is optimal for (55) then it is optimal for (56).

In an optimal solution of (56) all the y_i have the same value, namely

$$\max \{ p(y) : y \geq -\sqrt{3} \},$$

where $p(y) = (3/2)y - (1/2)y^3$. A boundary and critical point analysis on p shows it has a unique maximum for $y \geq -\sqrt{3}$ given by $p(1) = 1$.

Therefore, the problem (56) has a unique optimal solution y^* with $y_i^* = 1$ for all i , and its optimal value is $n+1$. Since y^* is also feasible for (55), it is its unique optimal solution with optimal value $n+1$. So, $N \leq 2(n+1)$,

and equality holds if and only if the matrix B has exactly $n + 1$ eigenvalues equal to $1 - t$ and $n + 1$ eigenvalues equal to $-(1 - t)$. It then follows that if $N = 2(n + 1)$, then C has exactly one nonzero eigenvalue, namely $2(1 - t)$ with multiplicity $n + 1$.

Assume that $N = 2(n + 1)$, so the set attains the maximum cardinality. Then, the all-ones vector e is in the kernel of U , and U has rank n . Indeed, $\text{rank } U \leq n < n + 1 = \text{rank } C$, and since $C = U - tJ$ it follows that e is not in the column space of U , so e is in the column space of C . The column space E of C is the eigenspace of C with eigenvalue $2(1 - t)$. Let $S \subseteq E$ be the orthogonal complement to the span of e in E . If $x \in S$, then

$$Ux = Cx + tJx = Cx = 2(1 - t)x,$$

hence x is an eigenvector of U . Since $\text{rank } U < \text{rank } C$ it follows that S is the only eigenspace of U with nonzero eigenvalue. Hence, $Ue = 0$ and U has rank n .

The equation $Ue = 0$ means that the barycenter of the set is 0. Moreover, from $0 = Ue = (C + tJ)e = (2(1 - t) + 2(n + 1)t)e$ it follows that $t = -1/n$. $\square$

By the continuous dependence of eigenvalues on the coordinates of a matrix, the bound $\alpha(n, t) \leq 2(n + 1)$ can be extended to $[-1, \varepsilon(n))$, where $\varepsilon(n)$ is some (small) positive number depending on n , something Bezdek and Lángi already showed. However, it is not true that this bound is global on $t \in [-1, 1)$, as a construction of Larman and Rogers [64] shows. Namely, let $n = 5$ and S be the set of vertices of the cube $[-1, 1]^5$ that have an odd number of positive signs. Then $|S| = 16$ with vectors of norm $\sqrt{5}$. Rescaling by $\sqrt{5}$ gives a $(1/5)$ -almost-equiangular set on S^4 of cardinality 16.

The proof of Theorem 8.10 moreover links the maximum obtuse almost-equiangular sets to the theory of spherical designs; see the survey by Bannai and Bannai [8] for more on spherical designs.

Proof of Theorem 8.2. According to Theorem 8.10, $\sum_{i=1}^{2(n+1)} x_i = 0$, the Gram matrix U of S satisfies $U^2 = 2(1 + 1/n)U$, and U has rank n . Moreover, the identity $U^2 = 2(1 + 1/n)U$ translates to

$$\sum_{k=1}^{2(n+1)} (x_i^\top x_k)(x_k^\top x_j) = 2(1 + 1/n)(x_i^\top x_j) \quad \text{for all } 1 \leq i, j \leq 2(n + 1).$$

By linearity, x_i and x_j can be replaced by any vector of $\mathbb{R}^n$. In particular, for all $u \in S^{n-1}$,

$$\sum_{k=1}^{2(n+1)} (u^\top x_k)^2 = 2(1 + 1/n).$$

This identity, together with $\sum_{i=1}^{2(n+1)} x_i = 0$, characterizes the spherical designs of strength 2. For a proof of the latter, see [8, Theorem 2.2], but note that in property (6) of this theorem the first appearance of the exponent k is wrong and should be $2k$. $\square$

Relation to orthogonal matrices. The union of two vertex-disjoint regular n -simplices, called a double regular n -simplex, is an example of a maximum obtuse almost-equiangular set. A natural question is whether this construction is unique. The affirmative answer for $n \leq 5$ is established in Theorem 8.4. Theorem 8.3 works towards this proof, and is interesting by itself.

Proof of Theorem 8.3. With similar notation as in the proof of Theorem 8.10, let B denote the matrix associated to a maximum obtuse almost-equiangular set of unit vectors. The matrix B has only two eigenvalues, namely $\pm(1+1/n)$, and hence satisfies $B^2 = (1+1/n)^2 I$. Moreover, $Be = (1+1/n)e$. Let $O = (1+1/n)^{-1}B$; it is clear from the properties of B that O is symmetric and orthogonal and that it satisfies the conditions (i)–(iii).

Conversely, given a symmetric and orthogonal matrix O satisfying (i)–(iii), let

$$U = (1+1/n)O - (1/n)J + (1+1/n)I$$

and let $E_{\pm 1}$ be the eigenspaces of O associated with the two eigenvalues ± 1 . Both of them have dimension $n+1$ because $\text{Tr}(O) = 0$ due to (ii). The kernel of U is the subspace $E_{-1} \oplus \mathbb{R}e$ of dimension $n+2$; its orthogonal complement is the eigenspace of U associated to the eigenvalue $2(1+1/n)$. So U is the Gram matrix of a set of $2(n+1)$ unit vectors in $\mathbb{R}^n$. Condition (iii) ensures that this set is $(-1/n)$ -almost equiangular. $\square$

Any t -almost-equidistant set in S^{n-1} with $t \leq 0$ can be lifted to an almost-orthogonal set on S^n [87]. Since $\alpha(n+1, 0) = 2(n+1)$, every maximum obtuse almost-equidistant set gives a maximum almost-orthogonal set in this way. Since $\overline{W_5}$ is $(5, 0)$ -realizable but not $(4, -1/4)$ -realizable (see the last subsection of Section 4), the converse is not the case. Deaett characterized the maximum almost-orthogonal sets by a statement similar to Theorem 8.3; it differs only by the eigenvector condition (i). Hence, the eigenvector condition distinguishes between those maximum almost-orthogonal sets that show this form of descent, and those that do not.

The distance graph of a maximum obtuse almost-equiangular set. A graph is *quadrangular* if no two vertices have exactly one neighbor in common.

LEMMA 8.11. *The following properties hold for the distance graph G of a maximum obtuse almost-equiangular subset S of S^{n-1} .*

- (i) *If G contains a K_{n+1} , then S is a double-regular n -simplex.*
- (ii) *The graph $\overline{G}$ is quadrangular.*
- (iii) *The degree of a vertex in $\overline{G}$ lies between 1 and $n+1$. If there is a vertex of degree $n+1$ in $\overline{G}$, then S is a double-regular n -simplex. If there is a vertex x with exactly one neighbor y in $\overline{G}$, then $G[S \setminus \{x, y\}]$ is the distance graph of a maximum obtuse almost-equiangular subset of S^{n-2} .*

Proof. Let $O = (n/(n+1))U + (1/(n+1))J - I$, where U is the Gram matrix of S , be the matrix of Theorem 8.3. The entries of O are equal to 0 on the diagonal and at pairs of vectors with inner product $-1/n$, so the adjacency matrix A of $\overline{G}$ is such that $A_{ij} = 0$ if $O_{ij} = 0$ and $A_{ij} = 1$ if $O_{ij} \neq 0$.

If G contains a K_{n+1} , then O is of the form $O = \begin{pmatrix} 0 & B \\ B^\top & D \end{pmatrix}$ where B and D are $(n+1) \times (n+1)$ matrices. The condition $O^2 = I$ leads to $BB^\top = I$ and $BD = 0$. But then B is invertible and so $D = 0$, which proves (i).

Property (ii) follows from the columns of O being pairwise orthogonal: if two vertices x_i, x_j share a single neighbor x_k in $\overline{G}$, then $O_{ki}O_{kj} \neq 0$, while $O_{li}O_{lj} = 0$ for $l \neq k$. But then the columns i and j of O would not be orthogonal.

To prove (iii), note that $\overline{G}$ is triangle free. Let x be a vertex and let $\overline{N}_x$ denote its set of neighbors in $\overline{G}$. Two vertices in $\overline{N}_x$ cannot be adjacent in $\overline{G}$, otherwise they would form a triangle with x . So $\overline{N}_x$ is a clique in G , that is, it is a regular simplex, which proves that the degree of x in $\overline{G}$ is at most $n+1$. Moreover, if x has degree $n+1$, then it follows from (i) that G contains a K_{n+1} , and hence that S is a double-regular n -simplex.

Next, given a vertex x , let N_x be its neighborhood in G . All vertices in N_x have inner product $-1/n$ with x , and so lie in an affine hyperplane, and hence belong to an $(n-2)$ -sphere C . By scaling and translating C via an affine transformation, it can be mapped to S^{n-2} , and then N_x is mapped to a t -almost-equidistant set for some $t \leq 0$. It then follows from Theorem 8.10 that $|N_x| \leq 2n$, and so the degree of x in $\overline{G}$ is at least 1. Moreover, if $|N_x| = 2n$, then $t = -1/(n-1)$. $\square$

Uniqueness of the double-regular simplex. The goal in this section is to prove Theorem 8.4. For a given dimension n , the theorem is false if there is an $(n, -1/n)$ -realizable anti-triangle-free graph of order $2(n+1)$ whose complement is not bipartite. It turns out that, to prove the theorem, it is

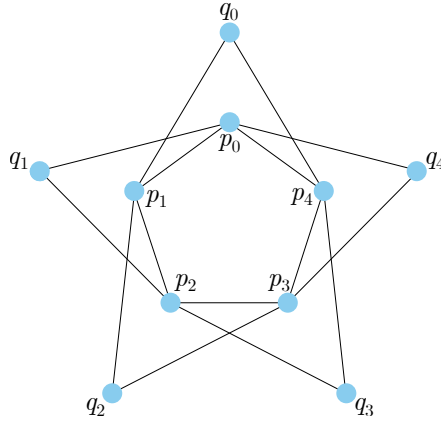


FIGURE 9. If $\overline{G}$ contains a 5-cycle, it contains this subgraph.

enough to show that such a graph whose complement contains a 5-cycle is not realizable.

Proof of Theorem 8.4. The distance graph of any maximum obtuse almost-equiangular set is anti-triangle free and, by Lemma 8.11, has a quadrangular complement. Moreover, if the set is not a double-regular n -simplex, then the complement is not bipartite. The goal of the proof is then to show that, if G is an anti-triangle-free graph of order $2(n+1)$ whose complement is quadrangular and nonbipartite, then G is not $(n, -1/n)$ -realizable. For $2 \leq n \leq 5$, this is done below.

Let G be an anti-triangle-free graph of order $2(n+1)$ whose complement is quadrangular. Say that $\overline{G}$ does not contain odd cycles of length $3, 5, \dots, 2k-1$, but contains an odd cycle of length $2k+1$ with vertices $p_0, \dots, p_{2k}$. Since $\overline{G}$ is quadrangular, every pair of vertices p_i, p_{i+2} , with indices taken modulo $2k+1$, has at least two common neighbors. One of the neighbors is p_{i+1} ; denote the other by q_{i+1} . Since $\overline{G}$ does not contain odd cycles of length less than $2k+1$, the vertices p_i and q_i must all be distinct, and so the order of G is at least $2(2k+1)$, whence $n \geq 2k$. This settles the case $n = 3$.

DIMENSIONS 4 AND 5. It follows that, for $n \leq 5$, if G is an anti-triangle-free graph of order $2(n+1)$ whose complement is quadrangular and nonbipartite, then $\overline{G}$ has an odd cycle of length 5, and since $n \geq 2k$ as shown above, it is necessary that $n \geq 4$. So it suffices to show that such a graph G for $n = 4$ and 5 is not $(n, -1/n)$ -realizable.

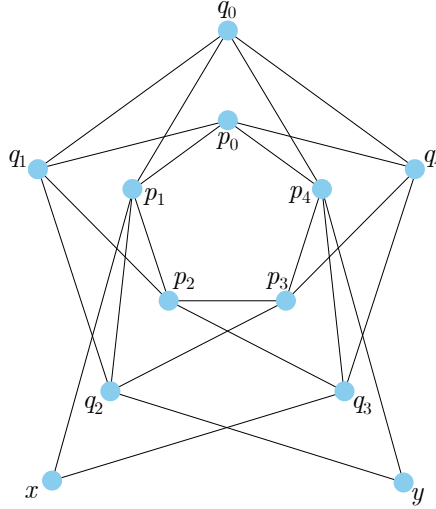


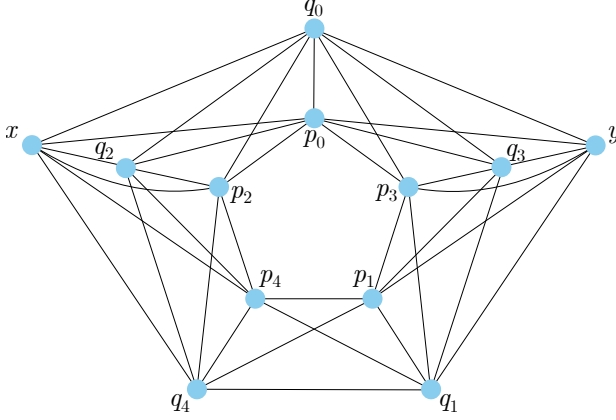
FIGURE 10. When $n = 5$, the graph $\overline{G}$ may contain this graph as a subgraph.

To this end, note that if $p_0, \dots, p_4$ is a 5-cycle in $\overline{G}$ and if $q_0, \dots, q_4$ are the common neighbors defined above, then $\overline{G}$ has the graph in Figure 9 as a subgraph. Again since $\overline{G}$ is quadrangular, the pairs p_i, q_{i+2} must have another common neighbor besides p_{i+1} . If $n = 4$, there are no other vertices available, so the only possibility is that $q_0, \dots, q_4$ is a cycle, that is, $\overline{G}$ is isomorphic to W_5 (see Figure 8). The graph $\overline{W}_5$ is not $(4, -1/4)$ -realizable (see the end of Section 4), so the proof is finished for $n = 4$.

The remaining case is $n = 5$, for which G has order 12. Call x, y the two vertices of G other than the p_i and q_i . By an argument similar to the one above, $\overline{G}$ contains as a subgraph either W_5 , as was the case for $n = 4$, or, without loss of generality, the graph in Figure 10.

Dimension 5 and $\overline{G}$ contains the graph of Figure 10. If $\overline{G}$ contains the graph of Figure 10, then since $\overline{G}$ is triangle free and x is adjacent to p_1 and q_3 in $\overline{G}$, it must be that x is adjacent to p_0, q_0, p_2, q_2, p_4 , and q_4 in G . The same reasoning for y shows that G contains as a subgraph the graph H_{12} from Figure 11. It will turn out that H_{12} is not $(5, -1/5)$ -realizable.

Dimension 5 and $\overline{G}$ contains W_5 . If $\overline{G}$ contains W_5 , then the graph G contains a subgraph isomorphic to H_{12} as well. Indeed, in this case, the vertices x and y must be adjacent in $\overline{G}$ to the subgraph W_5 , otherwise by (iii) of Lemma 8.11

FIGURE 11. The graph H_{12} is not $(5, -1/5)$ -realizable.

the graph $\overline{W_5}$ would be the distance graph of a $(-1/4)$ -almost-equiangular set in S^3 with 10 points that is not a regular double simplex, a contradiction.

If v is a vertex of $\overline{W_5}$ in $\overline{G}$, then the neighborhood of v in $\overline{W_5}$ is an independent set, since $\overline{G}$ is triangle free. The neighborhood forms a clique in G ; call it C_v . If x is adjacent to v in $\overline{G}$, again since $\overline{G}$ is triangle free, x is adjacent to all vertices of C_v in G .

Since x is adjacent in $\overline{G}$ to at least one vertex v of $\overline{W_5}$, without loss of generality say x is adjacent to p_1 . Then, x is adjacent in G to $C_{p_1} = \{p_0, q_0, p_2, q_2\}$. But then without loss of generality x is adjacent in G to $\{p_0, q_0, p_2, q_2, p_4, q_4\}$. Namely, if x is not adjacent to any of p_3, q_3, p_4 , and q_4 in $\overline{G}$, the statement follows immediately. Otherwise, if x is adjacent, say, to p_3 in $\overline{G}$, then x is adjacent in G to $C_{p_1} \cup C_{p_3} = \{p_0, q_0, p_2, q_2, p_4, q_4\}$.

It remains to show that y is adjacent in G to all vertices in $\{p_1, q_1, p_3, q_3\}$; applying the previous reasoning to y shows that if this is the case, y is adjacent to all vertices in either $\{p_4, q_4, p_1, q_1, p_3, q_3\}$ or $\{p_1, q_1, p_3, q_3, p_0, q_0\}$, meaning that a subgraph isomorphic to H_{12} occurs in G .

To prove that y is adjacent to all vertices in $\{p_1, q_1, p_3, q_3\}$, consider the following. In order to arrive at a contradiction, assume y is adjacent to p_1 in $\overline{G}$, again without loss of generality. Then y is connected in G to $C_{p_1} = \{p_0, q_0, p_2, q_2\}$. But x is also adjacent in G to these vertices, and if G contains a K_6 , then it is not $(5, -1/5)$ -realizable ((i) of Lemma 8.11), so $\{x, y\}$ is independent in G . Now the contradiction comes from the quadrangularity of $\overline{G}$; indeed, if x and y are not adjacent in G , then y is a common

neighbor of x and p_1 in $\overline{G}$. But it is not possible that x and p_1 have a second common neighbor because x is not connected to any neighbor of p_1 in $\overline{G}$ other than y .

To complete the proof, it remains to show that the graph H_{12} is not $(5, -1/5)$ -realizable. This is a specialization of a part of the proof of the nonrealizability of W_k from the end of Section 4. In fact, the graph H_{12} is a subgraph of $\overline{W}_7$, with two vertices and some edges removed. The removed edges play no role in the proof, and the two vertices only play a role for nonrealizability for $n = 6$, but for $n = 5$ they are superfluous. $\square$

6. Classification in dimensions 2 and 3

Section 4 gives exact conditions on the dimension n and inner product t for which simplices, rhombi, and spindles are (n, t) -realizable. For every integer $m \geq 1$, dimension n , and inner product t , this gives sufficient conditions for the existence of (n, t) -realizable anti-triangle-free graphs of order m . These realizable graphs then give t -almost-equiangular sets of cardinality m in dimension n . In this section a converse result is obtained in low dimension: list all maximum-cardinality, almost-equiangular sets in S^{n-1} , with $n \leq 3$.

Say that an anti-triangle-free graph is *minimal* if the removal of any edge results in a graph that is not anti-triangle free. Given n and t , say that an anti-triangle-free graph is (n, t) -*optimal* if it is (n, t) -realizable and if it has order $\alpha(n, t)$. If a graph is the unique minimal (n, t) -optimal graph up to isomorphism, then it is called a *unique optimal construction*.

Finding all minimal (n, t) -optimal graphs for low dimension n is done by performing a graph search. The results from Section 4 provide the conditions for this search. They also give lower bounds on $\alpha(n, t)$. Theorem 8.1 and Theorem 8.10 provide an upper bound of $\alpha(n, t) \leq 2(n + 1)$ for $t \in [-1, 0]$, which is only attained at $t = -1/n$. There exists a global lower bound of $\alpha(n, t) \geq 4$, given by the disjoint union of two edges.

Perform the graph search as follows. Let $t_{k,i}$ be the i th root of the polynomial (54) for fixed k . Given t and $2 \leq k \leq n$, list all graphs G of a given order that do not contain a subgraph isomorphic to:

- an anti-triangle;
- K_{n+2} ;
- an n -rhombus;
- a k -rhombus if $t \leq -1/k$;
- K_{k+1} if $t < -1/k$;
- K_{n+1} if $t > -1/n$;
- an extended $(n - 1)$ -rhombus if $t = -1/n$;

- $MS(k)$ if $t < t_{k,1}$;
- $MS(n-1)$ if $t > t_{n-1,2}$.

The search is implemented in SageMath in a script in the supplement of [4]. The code is a modified version of the code used in [7]. Given a dimension n , all anti-triangle-free graphs of cardinality at most $2(n+1)$ not containing a K_{n+2} are generated. A second script reduces the size of these sets greatly by only taking the minimal anti-triangle-free graphs. Finally, each graph is searched for the above list of subgraphs. The results below are summarized in Figure 3.

DIMENSION 2. The three-point bound $\text{blockM}_3(H(n, t))$ for $t \leq 0$ proves a global upper bound $\alpha(2, t) \leq 6$, which is only achieved at $t = -1/2$ by the double triangle. A graph search on order 6 graphs that are anti-triangle-free and do not contain K_4 or a 2-rhombus shows that this is the only minimal order 6 construction on the circle.

The Moser Spindle $MS(1)$ is realizable for $t = -(1/4)(1 \pm \sqrt{5})$ and has order 5. Its graph is a 5-cycle, which is the unique anti-triangle-free graph of order 5 containing no K_3 . The inner product $t = -(1/4)(1 + \sqrt{5}) = \cos(-4\pi/5)$ corresponds to the pentagram and $t = -(1/4)(1 - \sqrt{5}) = \cos(-2\pi/5)$ corresponds to the regular pentagon.

Every other anti-triangle-free graph of order 5 satisfying the constraints above contains the disjoint union of a triangle and an edge, which is only realizable at $t = -1/2$. This shows that the 5-cycle is the unique optimal construction at $t = -(1/4)(1 \pm \sqrt{5})$.

At every other inner product the maximum cardinality is 4, attained by two disjoint edges, which is the unique optimal construction of this order.

DIMENSION 3. The three-point bound $\text{blockM}_3(H(n, t))$ for $t \leq 0$ proves a global upper bound $\alpha(3, t) \leq 8$. This is achieved by the double tetrahedron for $t = -1/3$. The graph search shows that the double tetrahedron is the only minimal construction of order 8.

In the region $t_{2,1} \leq t \leq t_{2,2}$, the Moser spindle $MS(2)$ is $(3, t)$ -realizable and of order 7. Excluding this subgraph from the graph search shows that it is the unique optimal construction for $t_{2,1} \leq t < -1/3$ and $-1/3 < t \leq t_{2,2}$.

For $t \geq -1/2$, the double triangle is realizable. For $t = -1/2$, it is the unique optimal construction. For $t > -1/2$, the spindle $S(1, 2)$ is realizable and of order 6. Excluding these subgraphs from the graph search shows there are no other $(3, t)$ -realizable graphs of order 6 with $t > -1/2$. So for $-1/2 < t < t_{2,1}$ and $t_{2,2} < t \leq 0$, there are two optimal constructions of order 6.

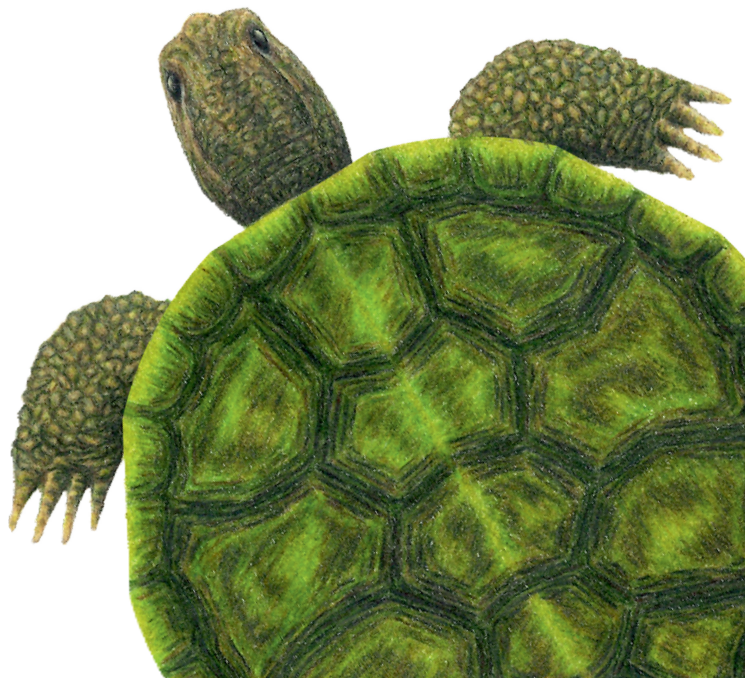
For all $-(1/4)(1 - \sqrt{5}) \leq t < -1/2$, the Moser spindle $MS(1)$ is the unique optimal construction and has order 5, as described above.

Acknowledgments

We thank Nando Leijenhorst for help with `ClusteredLowRankSolver.jl` and Willem de Muinck Keizer for helpful discussions. Alexey Glazyrin has pointed us to references [16, 15].

Part 4

Euclidean space



Distance-avoiding sets

This chapter is part of ongoing work together with Fernando Mário de Oliveira Filho.

The compilation *Problems, problems, problems* by Moser [75] contains the following question—first posed by Moser’s brother, Leo Moser:

LM 25 (1966) Estimate the “size” of the largest measurable point set in a large square, which does not determine unit distance.

This is precisely Problem III from the introduction of this thesis.

Denote the Lebesgue measure on $\mathbb{R}^n$ by λ . In integrals, the notation dx always means $d\lambda(x)$. For a Lebesgue-measurable set $S \subseteq \mathbb{R}^n$, define the *upper density* of S by

$$\bar{\delta}(S) = \limsup_{T \rightarrow \infty} \frac{\lambda(S \cap [-T/2, T/2]^n)}{T^n}.$$

Problem III can then be formulated as: what is the largest upper density a Lebesgue-measurable subset of $\mathbb{R}^n$ not containing pairs at distance 1 can have. We denote this number by $m_1(\mathbb{R}^n)$.

Erdős [42] conjectured that $m_1(\mathbb{R}^n) < 2^{-n}$. This was recently confirmed for $n = 2$ by Ambrus, Csiszárík, Matolcsi, Varga, and Zsámboki [2], who gave the upper bound $m_1(\mathbb{R}^2) \leq 0.2470$. The bound was partly based on convex optimization techniques introduced by Oliveira and Vallentin [82] and DeCorte, Oliveira, and Vallentin [36], using semidefinite programming. The gap between lower and upper bounds on $m_1(\mathbb{R}^n)$ is currently still quite large—the lower bound being $m_1(\mathbb{R}^2) \geq 0.22936$, as the tortoise construction by Croft shows [33]. The paper [2] contains a more detailed account of the history of $m_1(\mathbb{R}^n)$.

Recall that a choice of set $D \subseteq (0, \infty)$ defines the distance graph $G(D)$ with vertex set $\mathbb{R}^n$ and edge set $E(D)$ with $xy \in E(D)$ if and only if $\|x - y\| \in D$. In this chapter, we study the problem of determining

$$m_D(\mathbb{R}^n) = \sup\{\bar{\delta}(S) : S \subseteq \mathbb{R}^n \text{ is measurable and avoids } D\}.$$

A completely positive formulation for the general distance-avoiding-set problem on $\mathbb{R}^n$ was introduced by DeCorte, Oliveira, and Vallentin [36]. They prove this bound is exact, and compute new upper bounds on $m_1(\mathbb{R}^n)$ in low dimensions by including well-chosen constraints coming from the completely positive cone. They also use the full completely positive formulation to reprove a theorem by Bukh [21] about sets avoiding many distances. The theorem says that if $n, m \geq 2$ are integers, and $d_1, \dots, d_m$ are positive numbers, then $m_{\{d_1, \dots, d_m\}}(\mathbb{R}^n)$ approaches $m_1(\mathbb{R}^n)^m$ as the ratios d_i/d_{i+1} go to infinity.

In the next section, we describe a completely positive hierarchy for these problems, based on the outer approximation of $\text{CP}(\mathbb{R}^n)_{\text{inv}}$ described in Chapter 3, and show it converges to the exact optimal upper density. We end with some notes on how we can use the results from Chapter 5 to say something about other hierarchies for distance-avoiding-set problems on $\mathbb{R}^n$.

1. Lattices, tori, and densities

Preliminaries on harmonic analysis can be found in Section 3.1 and the appendix.

Fix an integer $n \geq 2$. Analysis on $\mathbb{R}^n$ can often be factored through its tori. It suffices to consider tori of the form $\mathbb{R}^n/L\mathbb{Z}^n$. The necessary preliminaries on lattices are taken from [11, Appendix B.2].

To be precise, let $L > 0$ be a real number, then $L\mathbb{Z}^n$ is a lattice in $\mathbb{R}^n$. The associated torus is $\mathbb{T}_L = \mathbb{R}^n/L\mathbb{Z}^n$, where we omit n from the notation, as it is always fixed and clear from context. Let $p_L : \mathbb{R}^n \rightarrow \mathbb{T}_L$ be the quotient map. The torus $\mathbb{T}_L$ is itself a compact group under the quotient topology. Each torus $\mathbb{T}_L$ is a metric space when equipped with the metric

$$d_L(p_L(x), p_L(y)) = \inf_{v \in L\mathbb{Z}^n} \|x - y + v\|,$$

which is a right-invariant density metric.

A *fundamental domain* of a lattice $L\mathbb{Z}^n$ is a Borel set F with respect to which $\mathbb{R}^n = \bigcup_{v \in L\mathbb{Z}^n} F + v$ and $(F + v) \cap (F + v') = \emptyset$ for all v and $v' \in L\mathbb{Z}^n$ such that $v \neq v'$. We choose for $L > 0$ the fundamental domain $F_L = [-L/2, L/2]^n$. Thus, $\mathbb{R}^n$ is a discrete union of translates of F_L for each L . By Equation (12),

there is a Haar measure μ_L on $\mathbb{T}_L$ such that

$$\lambda(F_L) = \int_{\mathbb{R}^n} \mathbf{1}_{F_L}(x) dx = \int_{\mathbb{T}_L} \sum_{v \in L\mathbb{Z}^n} \mathbf{1}_{F_L}(x+v) d\mu_L(p(x)) = \mu_L(\mathbb{T}_L).$$

It is common to work with this Haar measure, so we do so as well; thus, in all that follows, $\mu_L(\mathbb{T}_L) = L^n$.

A set $S \subseteq \mathbb{R}^n$ is called *periodic with period L* if for all $v \in L\mathbb{Z}^n$, $S+v = S$. A *periodic set* is a set that is periodic with period L for some L . Sets periodic with period L define a subset of $\mathbb{T}_L$ by taking the quotient. On the other hand, a subset $S \subseteq \mathbb{T}_L$ defines a subset of $\mathbb{R}^n$ by the section $s : \mathbb{T}_L \rightarrow F_L$ of the quotient map, and taking $\bigcup_{v \in L\mathbb{Z}^n} s(S) + v$. This correspondence is bijective.

Likewise, a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called *periodic with period L* if it is invariant under translation by elements of $L\mathbb{Z}^n$, that is, $f(x+v) = f(x)$ for all $x \in \mathbb{R}^n$ and $v \in L\mathbb{Z}^n$. Call f *periodic* if it is a periodic function with period L for some L . A function on $\mathbb{T}_L$ defines a periodic function on $\mathbb{R}^n$ by composition with the quotient map. Vice versa, a periodic function on $\mathbb{R}^n$ defines a function on $\mathbb{T}_L$ by composition with the section $s : \mathbb{T}_L \rightarrow F_L$.

Let for $D \subseteq (0, \infty)$ and $L > 0$ the graph $G_L(D) = (\mathbb{T}_L, E(D))$ be the distance graph given by $p_L(x)p_L(y) \in E(D)$ if and only if there is a $v \in L\mathbb{Z}^n$ such that $\|x-y+v\| \in D$. Then, G_L is a homogeneous graph on $\mathbb{T}_L$ as defined in Chapter 4. To avoid confusion, we denote the measurable independence number of $G_L(D)$ under μ_L by $\alpha_{\mu_L}(G_L(D))$.

The key insight for dealing with distance-avoiding sets on $\mathbb{R}^n$ is that they are approximated by periodic distance-avoiding sets. For a measurable periodic set S with period L the upper density is $\bar{\delta}(S) = L^{-n}\mu_L(S)$, which relates $m_D(\mathbb{R}^n)$ to the optimal densities $\alpha_{\mu_L}(G_L(D))$.

A theorem by Furstenberg, Katznelson, and Weiss [47, Theorem A] shows that if D is unbounded, $m_D(\mathbb{R}^n) = 0$. The following lemma says that if D is bounded, we may approximate a measurable D -avoiding set by periodic measurable D -avoiding sets, that is, independent sets of $G(D)$ are approximated arbitrarily well by independent sets of $G_L(D)$. A proof was given by DeCorte, Oliveira, and Vallentin [36, Lemma 6.2].

LEMMA 9.1. *If $n \geq 2$ and $D \subseteq (0, \infty)$ is bounded, then*

$$m_D(\mathbb{R}^n) = \limsup_{L \rightarrow \infty} \frac{\alpha_{\mu_L}(G_L(D))}{L^n}.$$

This behavior is reflected by the optimization upper bounds on $m_D(\mathbb{R}^n)$ we consider. Define the operator

$$Mf = \limsup_{T \rightarrow \infty} T^{-n} \int_{[-T/2, T/2]^n} f(x) dx.$$

For $\mathcal{C} \subseteq L^\infty(\mathbb{R}^n)$ a convex cone, let

$$\begin{aligned} \vartheta(G(D), \mathcal{C}) = \sup \quad & Mf \\ & f(0) = 1 \\ & f(x) = 0 \quad \text{if } x \in D \\ & f \in \mathcal{C} \text{ is of positive type,} \end{aligned}$$

where we use that a function of positive type is in particular continuous.

As usual, denote $C_r(V) = C_r(V, 2)$ and $\text{CP}(V) = \text{CP}(V, 2)$. For a homogeneous graph $G = (V, E)$ with V compact and a convex cone $\mathcal{C} \subseteq C_{\text{sym}}(V)$, we have $\vartheta_{\text{small}}(G, \mathcal{C}) = \vartheta_{\text{big}}(G, \mathcal{C})$; denote this program by $\vartheta(G, \mathcal{C})$. DeCorte, Oliveira, and Vallentin showed that, when D is closed, $m_D(\mathbb{R}^n)$ has the completely positive formulation $\vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}})$ by proving

$$\vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}}) = \limsup_{L \rightarrow \infty} \frac{\vartheta(G_L(D), \text{CP}(\mathbb{T}_L)_c)}{L^n}, \quad (57)$$

for all closed $D \subseteq (0, \infty)$ [36, Theorem 6.3]. Together with Lemma 9.1, Equation (57) implies $m_D(\mathbb{R}^n) = \vartheta(G(D), \text{CP}(V)_{\text{inv}})$.

The requirement that D is closed is unnecessary for us: it is only there to ensure that the graphs $G_L(D)$ are locally independent. However, we managed to prove Theorem 4.12 without this assumption, and an inspection of the proof of [36, Theorem 6.3] shows that this implies the following theorem.

THEOREM 9.2. *If $n \geq 2$, $D \subseteq (0, \infty)$ is bounded, and $m_D(\mathbb{R}^n) > 0$, then*

$$m_D(\mathbb{R}^n) = \vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}}).$$

Proof. See the discussion above. □

It may seem that Theorem 9.2 misses an assumption: it is not clear that $\vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}})$ is feasible. For example, if $D = (0, 1)$, then the conditions $f(x) = 0$ if $\|x\| \in D$ and $f(0) = 1$ are in contradiction with continuity of f . On the other hand, this is the only thing that can prevent feasibility. Likewise, if $m_D(\mathbb{R}^n) > 0$, the set D must be bounded away from 0. Thus,

positivity of the density is equivalent to feasibility of $\vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}})$. We already saw this in the finite-measure setting of Chapter 4.

2. Convergence of the completely positive hierarchy

Given Lemma 9.1, we can hope that to prove convergence of a completely positive hierarchy for $m_D(\mathbb{R}^n)$, we can somehow factor the hierarchy through a converging completely positive hierarchy for each $\alpha_{\mu_L}(G_L(D))$. However, for nonthick edge sets, we only managed to prove convergence for the unit sphere and similar spaces, a result which dependent highly on the representation theory of these spaces. Particularly, in Section 5.4 we saw that our method does not work for nonthick distance graphs on tori. Thus, this approach does not work.

Luckily, like for distance-avoiding sets on S^{n-1} , the representation theory of $\mathbb{R}^n$ saves us. We will show it is enough to consider feasible solutions that are *radial*: these are the functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $f(Tx) = f(x)$ for all $T \in \text{O}(n)$. Thus, radial functions only depend on the norm of their argument. Radial functions of positive type are given by the integral of a function in $C_0(\mathbb{R})$ under a nonnegative Borel measure, so that a weak*-converging sequence of such measures gives a sequence of radial functions of positive type that converges pointwise, except at 0.

As we will see, the programs $\vartheta(G(D), \mathcal{C})$ with $\mathcal{C} = C_r(\mathbb{R}^n)^*_{\text{inv}}$ or $\text{CP}(\mathbb{R}^n)_{\text{inv}}$ are invariant under the action of $\text{O}(n)$. We can thus assume that a feasible solution f is radial and positive type. Schoenberg [92, Theorem 1] showed that for any complex-valued radial function of positive type, there exists a Borel measure $\alpha \in M(\mathbb{R})_{\geq 0}$ such that

$$f(x) = \int_0^\infty \Omega_n(t\|x\|) d\alpha(t), \text{ with}$$

$$\Omega_n(\|x\|) = 1/\omega_n \int_{S^{n-1}} e^{ix^\top \xi} d\omega(\xi), \quad (58)$$

with ω the standard surface measure on the sphere and $\omega_n = \omega(S^{n-1})$. The proof goes as follows: by Bochner's theorem [45, Theorem 4.19], if $f \in \text{P}(\mathbb{R}^n)$, then there exists a measure $\nu \in M(\mathbb{R}^n)_{\geq 0}$ such that for all $x \in \mathbb{R}^n$ we have $f(x) = \int_{\mathbb{R}^n} e^{ix^\top \xi} d\nu(\xi)$. Then, by interchanging integrals,

$$\omega_n^{-1} \int_{S^{n-1}} \int_{\mathbb{R}^n} e^{i\|x\|y^\top \xi} d\nu(\xi) d\omega(x) = \int_{\mathbb{R}^n} \Omega_n(\|x\|\|\xi\|) d\nu(\xi).$$

Hence, define α by $\alpha(g) = \int_{\mathbb{R}^n} g(\|x\|) d\nu(x)$.

The functions Ω_n lie in $C_0(\mathbb{R})$ for $n \geq 2$. This follows, for example, from the expansion $\Omega_n(t) = \Gamma(n/2)(2/t)^{(n-2)/2} J_{(n-2)/2}(t)$ for $t > 0$, where $J_{(n-2)/2}$ is a Bessel function of the first kind—Equation (1.8) in [92]—together with an asymptotic formula for the Bessel functions—equation (1) in §7.21 of [101]—which says that $J_{(n-2)/2}(t) \rightarrow 0$ as $t \rightarrow \infty$.

THEOREM 9.3. *If $n \geq 2$, $D \subseteq (0, \infty)$, and $m_D(\mathbb{R}^n) > 0$, then*

$$m_D(\mathbb{R}^n) = \lim_{r \rightarrow \infty} \vartheta(G(D), \check{C}_r(\mathbb{R}^n)_{\text{inv}}^*).$$

Proof. The inequality $\lim_r \vartheta(G(D), \check{C}_r(\mathbb{R}^n)_{\text{inv}}) \geq \vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}})$ follows from the inclusions $\text{CP}(\mathbb{R}^n)_{\text{inv}} \subseteq \check{C}_r(\mathbb{R}^n)_{\text{inv}}$.

The first step to show the other inequality, is to prove that the image of each of the cones $\check{C}_r(\mathbb{R}^n)_{\text{inv}}^* \cap \text{P}(\Gamma)$ and $\text{CP}(\mathbb{R}^n)_{\text{inv}}$ under $\text{Av}_{\text{O}(n)}$ is contained in the original cone. Then show that the objective and constraints are also preserved under $\text{Av}_{\text{O}(n)}$, which implies that it suffices to consider radial functions.

Denote the Borel sets with finite Borel measure by $\mathcal{B}_{\text{fin}}$, and recall that for all $A \in \mathcal{B}_{\text{fin}}$, $\text{K}_A f(x, y) = f(x - y)$ for all x and $y \in A$. By Lemma 3.6

$$\check{C}_r(\mathbb{R}^n)_{\text{inv}}^* \cap \text{P}(\mathbb{R}^n) = \bigcap_{A \in \mathcal{B}_{\text{fin}}} \text{K}_A^{-1} C_r(A)^* \cap \text{P}(\mathbb{R}^n).$$

Let ν be the Haar measure on $\text{O}(n)$, and let $\text{Av}_{\text{O}(n)} f(x) = \int_{\text{O}(n)} f(Tx) d\nu(T)$ for all $f \in L^\infty(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$. Take $f \in \check{C}_r(\mathbb{R}^n)_{\text{inv}}^*$, take $A \in \mathcal{B}_{\text{fin}}$, and take $K \in C_r(A)$, then by an application of Fubini-Tonelli,

$$\begin{aligned} \langle \text{K}_A \text{Av}_{\text{O}(n)} f, K \rangle &= \int_A \int_A \int_{\text{O}(n)} f(Tx - Ty) K(x, y) d\nu(T) dx dy \\ &= \int_{\text{O}(n)} \int_{TA} \int_{TA} f(x - y) K(T^{-1}x, T^{-1}y) dx dy d\nu(T) \\ &= \int_{\text{O}(n)} \langle \text{K}_{TA} f, TK \rangle d\nu(T) \geq 0, \end{aligned}$$

where $TK(x, y) = K(T^{-1}x, T^{-1}y)$, so that $TK \in C_r(TA)$ for all $T \in \text{O}(n)$. This uses that the Lebesgue measure is also invariant under the orthogonal group.

With a similar calculation it follows that

$$\langle \text{Av}_{\text{O}(n)} f, \rho \star \rho \rangle = \int_{\text{O}(n)} \langle f, \rho_T \star \rho_T \rangle d\nu(T) \geq 0$$

for all $\rho \in L^1(\mathbb{R}^n)$, where $\rho_T(x) = \rho(Tx)$. Thus, $\text{Av}_{\text{O}(n)} f \in \text{P}(\mathbb{R}^n)$. It is similarly clear that $\text{Av}_{\text{O}(n)} \text{CP}(\mathbb{R}^n)_{\text{inv}} \subseteq \text{CP}(\mathbb{R}^n)$.

Taking $\omega_n^{-1} \text{Av}_{O(n)} f$ preserves the objective. Indeed, let $f \in P(\mathbb{R}^n)$, by Bochner's theorem there exists $\alpha \in M(\mathbb{R}^n)_{\geq 0}$ such that $f(x) = \int_{\mathbb{R}^n} e^{ix^\top \xi} d\alpha(\xi)$. Then, $Mf = \alpha(0)$; to see this, let $\phi_T = T^{-n} \int_{[-T/2, T/2]^n} e^{iu^\top x} dx$. Then, by pointwise convergence $\phi_T \rightarrow \mathbf{1}_{\{0\}}$ and the dominated convergence theorem, $\alpha(0) = \lim_{T \rightarrow \infty} \alpha(\phi_T)$. The measure from Bochner's theorem corresponding to $\omega_n^{-1} \text{Av}_{O(n)} f$ is given by (58), so $Mf = \Omega_n(0)\alpha(0) = \alpha(0)$, since $\Omega_n(0) = 1$. The normalization and edge are constraints also preserved under this operation. Thus, in the remainder assume that all solutions are radial.

From here on, the proof follows steps similar to that of Theorem 5.5. That is, take for all $r \in \mathbb{N}$ a radial feasible solution f_r of $\vartheta(G(D), \check{C}_r(\mathbb{R}^n)_{\text{inv}})$ such that $Mf_r \geq m_D(\mathbb{R}^n)/2$. Then, show that there exists a feasible solution of $\vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}})$ such that for all $r_0 \in \mathbb{N}$ there is an $r \geq r_0$ such that $f(0)^{-1}Mf \geq Mf_r$. This and $\lim_r \vartheta(G(D), \check{C}_r(\mathbb{R}^n)_{\text{inv}}^*) \geq m_D(\mathbb{R}^n)$ concludes the proof.

Since each f_r is radial, there exists for all r an $\alpha_r \in M(\mathbb{R})_{\geq 0}$ such that $f_r(x) = \int_{\mathbb{R}} \Omega_n(t\|x\|) d\alpha_r(t)$. Then, $1 = f_r(0) = \alpha_r(\mathbb{R}) = \|\alpha_r\|$, thus the sequence $(\alpha_r)_r$ lies in the unit ball of $M(\mathbb{R})$, which is compact by Banach-Alaoglu [32, Theorem V.3.1]. Since $C_0(\mathbb{R})$ is separable, the weak* topology on the unit ball is metrizable [32, Theorem V.5.1], thus it is sequentially compact, and $(\alpha_r)_r$ has a converging subsequence; assume the sequence itself converges to $\alpha \in M(\mathbb{R})$. Then, α is nonzero, since for all r and continuous functions ϕ with $0 \in \text{supp } \phi$, $\alpha(\phi) = \lim_r \alpha_r(\phi) \geq \lim_r \alpha_r(0) \geq m_D(\mathbb{R}^n)/2$.

Define $f \in L^\infty(\mathbb{R}^n)$ by $f(x) = \int_{\mathbb{R}} \Omega_n(t\|x\|) d\alpha(t)$, which by Bochner's theorem is positive type, thus continuous, and $f(0) = \|\alpha\| \leq 1$. Moreover, note that $f_r \rightarrow f$ under the weak* topology in $L^\infty(\mathbb{R}^n)$.

Since $\check{C}_{r+1}(\mathbb{R}^n)_{\text{inv}}^* \subseteq \check{C}_r(\mathbb{R}^n)_{\text{inv}}^*$ for all $r \in \mathbb{N}$, and each cone is weak* closed, it follows that $f \in \bigcap_r \check{C}_r(\mathbb{R}^n)_{\text{inv}} = \text{CP}(\mathbb{R}^n)$ by Theorem 3.7. Moreover, since $\Omega_n \in C_0(\mathbb{R})$, it follows that for all $x \in \mathbb{R}^n \setminus \{0\}$ the map $t \mapsto \Omega_n(t\|x\|)$ is in $C_0(\mathbb{R})$, so $f(x) = \int_{\mathbb{R}} \Omega_n(t\|x\|) d\alpha(t) = \lim_r \int_{\mathbb{R}} \Omega_n(t\|x\|) d\alpha_r(t) = \lim_r f_r(x)$. This shows that f_r converges to f pointwise on $\mathbb{R}^n$, except perhaps at 0. In particular, if $x \in D$, $f(x) = 0$.

It follows that $f(0)^{-1}f$ is feasible for $\vartheta(G(D), \text{CP}(\mathbb{R}^n)_{\text{inv}})$. It is left to show that $f(0)^{-1}Mf \leq \lim_r Mf_r$. Moreover, with $\phi_T = T^{-n} \int_{[-T/2, T/2]^n} e^{ix^\top \xi} d\alpha(\xi)$ as before, since $\phi_T(0) = 1$ for all T and $\alpha_r \geq 0$ for all r , it follows that,

$$\alpha(\phi_T) = \lim_{r \in \mathbb{N}} \alpha_r(\phi_T) \geq \lim_{r \in \mathbb{N}} \alpha_r(0).$$

Therefore, $\alpha(0) = \lim_T \alpha(\phi_T) \geq \lim_T \lim_r \alpha_r(0) = \lim_r \alpha_r(0)$. Since $\|\alpha\| \leq 1$, indeed $M(f(0)^{-1}f) \geq \lim_r Mf_r$, and the result follows. $\square$

3. Some discussion of the implementation

By analogy with Witsenhausen's problem (Chapter 6) we might hope that an implementation of $\vartheta(G(\{1\}), \check{C}_1(\mathbb{R}^2)_{\text{inv}}^*)$ leads to better bounds on $m_1(\mathbb{R}^2)$. Although there indeed is again a closely related optimization problem that gives provable upper bounds on $m_1(\mathbb{R}^2)$ which are at least as strong as the best known bounds, we have not yet found strictly better bounds. It is unclear whether this is a shortcoming in the implementation or in the bound itself.

Sphere packing

This chapter is part of ongoing work with David de Laat and Fernando Mário de Oliveira Filho.

A *sphere packing* is a collection of congruent balls in $\mathbb{R}^n$ having pairwise-disjoint interiors. Problem IV from the introduction asks for the largest fraction of Euclidean space that can be covered by a sphere packing; this is called the *maximal sphere-packing density*. The problem of determining this number is the *sphere-packing problem*. Linear programming bounds were introduced to the sphere-packing problem by Cohn and Elkies [27] and proved highly successful. In dimensions 8 and 24 they are even sharp [100, 28].

Attention has recently moved to improving the upper bound by restricting the linear programming bound. Cohn, De Laat, and Salmon [30] introduced a three-point bound inspired by the block moment hierarchy for compact packing graphs. Cohn and Salmon [31] made an in-depth study of the relation between bounds on compact packing graphs and bounds on the sphere-packing problem, leading to a definition of a moment hierarchy for sphere packing, including a proof of its convergence.

In this chapter, we complement this work by introducing a copositive formulation of the maximal sphere-packing density and proving convergence of a corresponding hierarchy. The proof that the copositive formulation is an upper bound is a straight-forward modification of the proof that Cohn and Elkies gave that their linear programming bound upper bounds the sphere packing density. Exactness then follows by a reduction to compact packing graphs.

We will optimize over functions in a vector space X that lies between the space of compactly supported smooth functions $C_c^\infty(\mathbb{R}^n)$ and $C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$.

There is a good reason to treat the space X as a parameter. From the perspective of the theory it seems sufficient to only ask that the functions are integrable and continuous. On the other hand, the space of Schwartz functions has proven especially effective for calculating bounds [27, 26, 30]. The *Schwartz space*, denoted by $\mathcal{S}(\mathbb{R}^n)$, is defined as the space of all smooth functions $f(x)$ such that all partial derivatives decay faster than any negative power of $\|x\|$. Viazovska showed [100] that for the sphere-packing problem in $\mathbb{R}^8$, the Cohn-Elkies programming bound has an optimizer f_{opt} such that $f_{\text{opt}} \in \mathcal{S}(\mathbb{R}^n)$, and whose objective value is exactly the maximal sphere packing density $\mathbb{R}^8$. This result was later also obtained for $\mathbb{R}^{24}$ [28].

It is an open question whether there exists an optimizer that is Schwartz for every dimension. We will see, however, that for any space X between $C_0(\mathbb{R}^n)$ and $C_c^\infty(\mathbb{R}^n)$, the maximal sphere-packing density has a copositive formulation over X . Hence, there is a copositive formulation over $\mathcal{S}(\mathbb{R}^n)$.

The Fourier transform plays an important role in the optimization problems in this chapter; for $f \in L^1(\mathbb{R}^n)$, let

$$\widehat{f}(x) = \int_{\mathbb{R}^n} f(y) e^{-2\pi i x^\top y} dy$$

be the Fourier transform of f . In this chapter, the optimization bounds we will investigate are of the form

$$\begin{aligned} \vartheta_{\text{SP}}^X(\mathcal{C}) &= \inf f(0) + g(0) \\ &\quad \widehat{f}(0) = 1, \\ &\quad f(x) + g(x) \leq 0 \quad \text{for all } \|x\| \geq 1, \\ &\quad f \in \mathcal{P}(\mathbb{R}^n) \cap X, \quad g \in \mathcal{C} \cap X, \end{aligned} \tag{59}$$

where $\mathcal{C} \subseteq L^1(\mathbb{R}^n)$ is a convex cone and $C_c^\infty(\mathbb{R}^n) \subseteq X \subseteq C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ a linear subspace. The Cohn-Elkies bound is of this form, with X a suitable space and $\mathcal{C} = \{0\}$.

1. Sphere-packing densities

The *maximal sphere-packing density* of $\mathbb{R}^n$ is the quantity

$$\sup\{\bar{\delta}(S) : S \text{ is a sphere packing}\}.$$

This number is scaling invariant, in the sense that the radius of the spheres making up the packing does not matter, as long as they are all the same. In the rest of this chapter, the radius is $1/2$, which matches much of the literature. Denote a ball of radius r and center p by $B_r(p)$. The volume of the ball $B_{1/2}(0) \subseteq \mathbb{R}^n$ is denoted $V_{1/2}^n = \lambda(B_{1/2}(0))$.

It is more practical to work with the *maximal center density* Δ_n . Write the cube with side length T centered at 0 as $C_T = [-T/2, T/2]^n$. The maximal center density is the average number of centers of spheres in a packing per unit volume:

$$\Delta_n = \sup \left\{ \limsup_{T \rightarrow \infty} \frac{|P \cap C_T|}{T^n} : P \subseteq \mathbb{R}^n \text{ such that } \|x - y\| \geq 1 \right. \\ \left. \text{for all } x, y \in P \right\}. \quad (60)$$

Then, $V_{1/2}^n \Delta_n$ is exactly the maximal sphere-packing density of $\mathbb{R}^n$.

In Chapter 9 we saw that the density of a D -avoiding set can be approximated by the density of a periodic set, which is given by a single integral over the fundamental domain. Something similar holds for the center density, as explained by Cohn and Elkies [27, Appendix A].

Recall that for $n \in \mathbb{N}$ and $L > 0$ a real number, $\mathbb{T}_L = \mathbb{R}^n / L\mathbb{Z}^n$. Define the graph $G_L = (\mathbb{T}_L, E_L)$ by saying that $xy \in E_L$ if and only if $d_L(x, y) < 1$; then, G_L is the graph $G_L(D)$ of Chapter 9 with $D = (0, 1)$. The graphs G_L are compact packing graphs. It can be shown that

$$\Delta_n = \lim_{L \rightarrow \infty} \frac{\alpha(G_L)}{L^n}. \quad (61)$$

It is sometimes easier to work with the closure of the fundamental domain C_L instead of with the torus itself. Let $G_{C_L} = (C_L, E_L)$ be the graph with vertex set C_L and $xy \in E_L$ if and only if $\|x - y\| < 1$. The graphs G_{C_L} are also compact packing graphs. From (61), one can obtain

$$\Delta_n = \lim_{L \rightarrow \infty} \frac{\alpha(G_{C_L})}{L^n}. \quad (62)$$

2. Some more harmonic analysis

The following lemmas describe mechanics by which we can approximate feasible solutions of $\vartheta_{\text{SP}}^X(\mathcal{C})$ by solutions with compact support, having additional properties. In Sections 3 and 4 we will use these as reduction steps to proof that there are copositive formulations for Δ_n , and that there is a converging copositive hierarchy. See Section 3.1 for background and important inequalities.

LEMMA 10.1. *Let $g \in L^1(\mathbb{R}^n)$. If $\rho \in P(\mathbb{R}^n)$, then $(g \star g) \star \rho \in P(\mathbb{R}^n)$. If $\rho \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$ and $g \geq 0$, then $(g \star g) \star \rho \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$.*

Proof. The correlation of a function $\rho \in L^1(\mathbb{R}^n)$ and $f \in L^p(\mathbb{R}^n)$ is in $L^p(\mathbb{R}^n)$ for all $1 \leq p \leq \infty$, see Section 3.1. So, $(g \star g) \star \rho \in L^1(\mathbb{R}^n)$, and similar for the convolution. In particular, if $f \in L^2(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n)$, then $f \star g \in L^2(\mathbb{R}^n)$.

Let ρ and $g \in L^1(\mathbb{R}^n)$ and $f \in L^2(\mathbb{R}^n)$, then by a change of variables and Fubini-Tonelli,

$$\langle g \star \rho, f \rangle = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(y) \rho(x) f(x - y) dy dx = \langle \rho, g \star f \rangle.$$

Thus, for all ρ and $g \in L^1(\mathbb{R}^n)$ and $f \in L^2(\mathbb{R}^n)$,

$$\langle (g \star g) \star \rho, f \star f \rangle = \langle \rho, (g \star g) \star (f \star f) \rangle.$$

Furthermore, for all $x \in \mathbb{R}^n$, by Fubini-Tonelli and the two changes of variables $y' = y + z$ and $w' = w + z$,

$$\begin{aligned} (g \star g) \star (f \star f)(x) &= \int_{\mathbb{R}^n} (g \star g)(y) (f \star f)(x - y) dy \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(z) g(y + z) f(w) f(x - y + w) dw dz dy \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(z) g(y) f(w - z) f(x + w - y) dw dy dz \\ &= \int_{\mathbb{R}^n} (g \star f)(w) (g \star f)(x + w) dw \\ &= (g \star f) \star (g \star f)(x). \end{aligned}$$

Hence, if $\rho \in P(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ and $g \in L^1(\mathbb{R}^n)_{\geq 0}$, then for every $f \in L^2(\mathbb{R}^n)$ we have $f \star g \in L^2(\mathbb{R}^n)$, and

$$\langle (g \star g) \star \rho, f \star f \rangle = \langle \rho, (g \star f) \star (g \star f) \rangle \geq 0,$$

thus $(g \star g) \star \rho \in P(\mathbb{R}^n)$. If $\rho \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$ and $f \geq 0$, since $g \geq 0$ it follows in the same way that $(g \star g) \star \rho \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$. $\square$

The next lemma is a well-known method of bounding the support of a positive-type function, while preserving the constraints that figure in the definition of $\vartheta_{\text{SP}}^X(\mathcal{C})$, see for example the proof of [62, Theorem 3.1]. This approximation is used to prove that $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\text{COP}_{\text{inv}}(\mathbb{R}^n))$ is an upper bound on Δ_n . For a bounded set $C \subseteq \mathbb{R}^n$, let $\psi_C = \lambda(C)^{-1} \mathbf{1}_C \star \mathbf{1}_C$.

LEMMA 10.2. *Let $\rho \in C_0(\mathbb{R}^n)$, then*

- (i) $\psi_{C_T} \rho \in C_c(\mathbb{R}^n)$ for all $T > 0$,
- (ii) if $\rho \in P(\mathbb{R}^n)$, then $\psi_{C_T} \rho \in P(\mathbb{R}^n)$ for all $T > 0$,
- (iii) if $\rho \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$, then $\psi_{C_T} \rho \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$ for all $T > 0$, and
- (iv) $\lim_T \|\psi_{C_T} \rho - \rho\|_{\infty} = 0$.

Proof. Let $\rho \in C_0(\mathbb{R}^n)$ and $T > 0$ a real number. The support of ψ_{C_T} is contained in $C_T + C_T$, and the correlation of a bounded function with an integrable function is continuous, so the first statement follows immediately.

The second and third statement can be understood as a consequence of the fact that the Hadamard product of two positive-semidefinite matrices is positive semidefinite, and likewise, the Hadamard product of a positive-semidefinite matrix and a copositive matrix is copositive. Indeed, a continuous function is positive type if and only if for all $U \subseteq \mathbb{R}^n$ the matrices $f[U] = (f(u_i - u_j))_{u_i, u_j \in U}$ are positive semidefinite [45, Proposition 3.35]. Take x and $y \in \mathbb{R}^n$, then by a change of variables

$$\begin{aligned} T^n(\psi_{C_T}\rho)(x-y) &= \int_{\mathbb{R}^n} \mathbf{1}_{C_T}(z) \mathbf{1}_{C_T}(x-y+z) dz \rho(x-y) \\ &= \langle \mathbf{1}_{C_T+x}, \mathbf{1}_{C_T+y} \rangle \rho(x-y), \end{aligned}$$

and the inner product is given by a positive definite form, which proves the second statement. The third follows by a similar argument.

Finally, for convergence, note that

$$\|\psi_{C_T}\rho - \rho\|_\infty = \sup_{x \in \mathbb{R}^n} \left| \frac{\lambda(C_T \cap (C_T - x))}{T^n} - 1 \right| |\rho(x)|.$$

Since $\rho \in C_0(\mathbb{R}^n)$, for every $\epsilon > 0$ there is a compact K_ϵ such that $|\rho(x)| \leq \epsilon$ on $\mathbb{R}^n \setminus K_\epsilon$. Surely, since $\lambda(C_T \cap (C_T - x)) \leq T^n$, this implies

$$\sup_{x \in \mathbb{R}^n \setminus K_\epsilon} \left| \frac{\lambda(C_T \cap (C_T - x))}{T^n} - 1 \right| |\rho(x)| \leq \sup_{x \in \mathbb{R}^n \setminus K_\epsilon} |\rho(x)| \leq \epsilon.$$

On the other hand, the sequence $T^{-n}\lambda(C_T \cap (C_T - x))$ can be shown to converge to 1, uniformly for x in a compact set. Take T large enough such that $|T^{-n}\lambda(C_T \cap (C_T - x)) - 1| \leq \epsilon/\|\rho\|_\infty$ for all $x \in K_\epsilon$, then

$$\sup_{x \in K_\epsilon} \left| \frac{\lambda(C_T \cap (C_T - x))}{T^n} - 1 \right| |\rho(x)| \leq \epsilon \|\rho\|_\infty / \|\rho\|_\infty = \epsilon,$$

and the result follows. $\square$

For $r \in \mathbb{R}$, let $\rho_r(x) = \rho(rx)$. We will use the next lemma to smoothen functions, while again preserving the constraints from $\vartheta_{\text{SP}}^X(\mathcal{C})$.

LEMMA 10.3. *If $\rho \in C_c(\mathbb{R}^n)$ such that $\rho(x) \leq 0$ for all x with $\|x\| \geq 1$, if $r_i = 1/(1 - 1/i)$ for all $i \in \mathbb{N}_{\geq 2}$, and if $(\phi_i)_{i \in \mathbb{N}_{\geq 2}}$ is an approximate identity such that $\text{supp } \phi_i \subseteq B_{1/i}(0)$ for all i , then*

- (i) $r_i^n \phi_i \star \rho_{r_i}$ is in $C_c(\mathbb{R}^n)$ for all i ,
- (ii) $(r_i^n \phi_i \star \rho_{r_i})(x) \leq 0$ for all x with $\|x\| \geq 1$ and all i ,

- (iii) $\int_{\mathbb{R}^n} (r_i^n \phi_i \star \rho_{r_i})(x) dx = \int_{\mathbb{R}^n} \rho(x) dx$, and
 (iv) $\lim_i |(r_i^n \phi_i \star \rho_{r_i})(x) - \rho(x)| = 0$ uniformly on bounded sets.

Proof. The first statement follows from the fact that the cross-correlation of two bounded functions with compact support is continuous, for instance by Lemma 3.4, and the support of the correlation is contained in the sum of the support, which follows from a direct calculation.

The second statement follows from the definition of the correlation and the reverse triangle inequality. Indeed, let $\|x\| \geq 1$ and $\|y\| \leq 1/i$, then

$$\|r_i(x - y)\| \geq \frac{\|x\| - \|y\|}{1 - 1/i} \geq 1.$$

Then, for all x with $\|x\| \geq 1$,

$$\phi_i \star \rho_{r_i}(x) = \int_{B_{1/i}(0)} \phi_i(-y) \rho(r_i(x - y)) dy \leq 0.$$

For the third statement, use $\int \phi_i(x) dx = 1$ for all i and Fubini's theorem, invariance of the Haar measure, and the change of variables $x' = r_i(x - y)$, which gives

$$\begin{aligned} \int_{\mathbb{R}^n} r_i^n \phi_i \star \rho_{r_i}(x) dx &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} r_i^n \phi_i(-y) \rho(r_i(x - y)) dy dx \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \phi_i(y) \rho(x) dx dy = \int_{\mathbb{R}^n} \rho(x) dx. \end{aligned}$$

For the fourth statement, by Hölder's inequality,

$$\begin{aligned} r_i^n \phi_i \star \rho_{r_i}(x) - \rho(x) &= \int_{\mathbb{R}^n} \phi_i(-y) (r_i^n \rho(r_i(x - y)) - \rho(x)) dy \\ &\leq \sup_{y \in B_{1/i}(0)} |r_i^n \rho(r_i(x - y)) - \rho(x)|. \end{aligned}$$

Whence, it suffices to find for all $\epsilon > 0$, an i_ϵ such that for all $i \geq i_\epsilon$

$$\sup_{y \in B_{1/i}(0)} |r_i^n \rho(r_i(x - y)) - \rho(x)| \leq \epsilon,$$

which only depends on $\|x\|$.

Fix a real number $M \geq 0$, a vector $x \in \mathbb{R}^n$ such that $\|x\| \leq M$, and a vector $y \in B_{1/i}(0)$. Then

$$\|x - r_i(x - y)\| = \left\| \frac{x/i - y}{1 - 1/i} \right\| \leq \frac{\|x\| + i\|y\|}{i - 1} \leq \frac{M + 1}{i - 1},$$

Hence, for all $\delta > 0$ there exists an i_δ such that for all $i \geq i_\delta$, $\|x - r_i(x - y)\| \leq \delta$ for all $y \in B_{1/i}(0)$.

The function ρ is uniformly continuous, so for every $\epsilon > 0$ there exists a $\delta(\epsilon)$ such that, if $\|x - y\| \leq \delta(\epsilon)$, then $|\rho(x) - \rho(y)| \leq \epsilon$. In particular, for all $x \in B_M(0)$ and $y \in B_{1/i}(0)$ with $i \geq i_{\delta(\epsilon)}$, $|\rho(r_i(x - y)) - \rho(x)| \leq \epsilon$.

Finally, let $\epsilon > 0$, $\bar{\epsilon} = \epsilon/\|\rho\|_\infty + 1$, and $i_0 = \max\{i_{\delta(\epsilon)}, \lceil \bar{\epsilon}^{1/n}/(\bar{\epsilon}^{1/n} - 1) \rceil\}$. Then, for all $i \geq i_0$, since $i \geq 2$ and $i/(i - 1)$ is decreasing in i ,

$$\begin{aligned} |r_i^n \rho(r_i(x - y)) - \rho(x)| &= \left| \frac{\rho(r_i(x - y)) - \rho(x)(1 - 1/i)^n}{(1 - 1/i)^n} \right| \\ &\leq \frac{|\rho(r_i(x - y)) - \rho(x)|}{(1 - 1/i)^n} + \frac{|((1 - 1/i)^n - 1)\rho(x)|}{(1 - 1/i)^n} \\ &\leq 2^n |\rho(r_i(x - y)) - \rho(x)| + \left(\left(\frac{i}{i - 1} \right)^n - 1 \right) \|\rho\|_\infty \\ &\leq 2^n |\rho(r_i(x - y)) - \rho(x)| + (\bar{\epsilon} - 1) \|\rho\|_\infty \\ &\leq \epsilon(2^n + 1), \end{aligned}$$

which goes to 0 as ϵ goes to 0. $\square$

3. Upper bounds on the maximal sphere-packing density

Because ϑ_{SP} is a minimization problem, for any $X \subseteq Y \subseteq L^1(\mathbb{R}^n) \cap C_0(\mathbb{R}^n)$ and any convex cone $\mathcal{C} \subseteq L^1(\mathbb{R}^n)$,

$$\vartheta_{\text{SP}}^Y(\mathcal{C}_Y) \leq \vartheta_{\text{SP}}^X(\mathcal{C}_X). \quad (63)$$

We will use this principle to show that any subspace X of $C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ that contains $C_c^\infty(\mathbb{R}^n)$ satisfies $\vartheta_{\text{SP}}^X(\text{COP}(\mathbb{R}^n)_{\text{inv}}) = \Delta_n$. We will first show the inequality $\Delta_n \leq \vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$.

Recall that for a compact packing graph $G = (V, E)$ and a convex cone of continuous symmetric kernels $\mathcal{C} \subseteq C_{\text{sym}}(V)$,

$$\begin{aligned} \vartheta^*(G, \mathcal{C}) &= \inf t \\ &\quad \begin{aligned} T(x, x) &\leq t - 1 && \text{for all } x \in V, \\ T(x, y) &\leq -1 && \text{for all } xy \in E, \\ T &\in \mathcal{C}. \end{aligned} \end{aligned}$$

Dobre, Dür, Frerick, and Vallentin [39] showed that $\alpha(G) = \vartheta^*(G, \text{COP}(V)_c)$. Together with identity (61), this implies

$$\Delta_n = \lim_{L \rightarrow \infty} \frac{\vartheta^*(G_L, \text{COP}(\mathbb{T}_L)_c)}{L^n}. \quad (64)$$

To see how the problems $\vartheta^*(G_L, \text{COP}(\mathbb{T}_L)_c)$ correspond to the problem $\vartheta_{\text{SP}}^{C_0(X) \cap L^1(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$, think of the functions f and g as the image under $\text{Av}_{\mathbb{R}^n}$ of kernels on $\mathbb{R}^n$. The main difference between the ϑ^* for compact

packing graphs as above and ϑ_{SP} is the presence of the auxiliary function f in the latter.

Fix a compact packing graph $G = (V, E)$. Let us first bring $\vartheta^*(G, \mathcal{C})$ into a similar form. Of course, $\vartheta^*(G, \mathcal{C})$ is equivalent to

$$\begin{aligned} \inf t \\ T(x, x) \leq t \quad \text{for all } x \in V, \\ T(x, y) \leq 0 \quad \text{for all } xy \in E, \\ T - J \in \mathcal{C}, \end{aligned}$$

where $J = \mathbf{1}^{\otimes 2}$. If $\mathcal{C}$ is a convex cone that contains $\text{PSD}(V)$, then $T - J \in \mathcal{C}$ if and only if $T = F + G$ where $F - J \in \text{PSD}(V)$ and $G \in \mathcal{C}$; indeed, if $T - J \in \mathcal{C}$, then take $F = J$ and $G = T - J$. On the other hand, if $F - J \in \text{PSD}(V)$ and $G \in \mathcal{C}$, then $G + F - J \in \mathcal{C}$ because $\mathcal{C}$ is convex. It follows that for every convex cone $\mathcal{C}$ containing the positive-semidefinite cone,

$$\begin{aligned} \vartheta^*(G, \mathcal{C}) = \inf t \\ F(x, x) + G(x, x) \leq t \quad \text{for all } x \in V, \\ F(x, y) + G(x, y) \leq 0 \quad \text{for all } xy \in E, \\ F - J \in \text{PSD}(V) \text{ and } G \in \mathcal{C}. \end{aligned}$$

The split of T into F and G is important. The constraint $F - J \in \text{PSD}(V)$ comes down to a simple eigenvalue condition, whereas, a priori, the constraint $T - J \in \text{COP}(V)_c$ is a difficult one.

The identity (64) suggests an approach for the proof of the Theorem 10.4: show that for every feasible solution (f, g) of $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$, for large enough L , there is a feasible solution (t, F, G) of $\vartheta^*(G_L, \text{COP}(\mathbb{T}_L)_c)$ such that $f(0) + g(0) \geq t/L^n$.

THEOREM 10.4. *For all $n \geq 2$, $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}}) \geq \Delta_n$.*

Proof. Let (f, g) be a feasible solution of $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$. First, they may be assumed to have compact support: let $\psi_T = T^{-n} \mathbf{1}_{C_T} \star \mathbf{1}_{C_T}$. By Lemma 10.2, the functions $\psi_T f$ and $\psi_T g$ converge uniformly to f and g respectively, have compact support, and $\psi_T f \in \text{P}(\mathbb{R}^n)$ and $\psi_T g \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$. Moreover, $\psi_T f(x) + \psi_T g(x) \leq 0$ if $\|x\| \geq 1$.

If $\tau_T = \widehat{\psi_T f}(0) = \int_{\mathbb{R}^n} \psi_T(x) f(x) dx$, by the dominated convergence theorem, $\lim_T \tau_T = \int_{\mathbb{R}^n} f(x) dx = \widehat{f}(0)$. So, the pair $(\tau_T^{-1} \psi_T f, \psi_T g)$ is feasible for $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$ and converges uniformly to (f, g) . Whence, f and g can be assumed to have compact support.

Let (f, g) be a feasible solution to $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$ with compact support, and let L be large enough such that $\text{supp } f$ and $\text{supp } g$ lie entirely

in $[-L/2, L/2]^n$. Let $p: \mathbb{R}^n \rightarrow \mathbb{T}_L$ be the quotient map. Define the functions

$$\phi(x) = \sum_{v \in L\mathbb{Z}^n} f(x+v) \quad \text{and} \quad \gamma(x) = \sum_{v \in L\mathbb{Z}^n} g(x+v);$$

the sums are finite, since f and g have compact support. Since these functions are invariant under the action of $\mathbb{T}_L$, the kernels $F(p(x), p(y)) = \phi(x - y)$ and $G(p(x), p(y)) = \gamma(x - y)$ are well-defined on $\mathbb{T}_L$. Moreover,

$$F(x, x) + G(x, x) = \phi(0) + \gamma(0) = \sum_{v \in L\mathbb{Z}^n} f(v) + g(v) = f(0) + g(0),$$

since $\text{supp } f \cup \text{supp } g \subseteq [-L/2, L/2]^n$. So, take $t = L^n f(0) + L^n g(0)$. Since we can take $L \geq 1$, it follows that $F(x, x) + G(x, x) \leq t$.

It is left to prove that (t, F, G) satisfies the edge and cone constraints of $\mathcal{V}^*(G_L, \text{COP}(\mathbb{T}_L)_c)$. Recall that $G_L = (\mathbb{T}_L, E)$ with $xy \in E$ if and only if $d_L(x, y) \geq 1$. If $d_L(x, y) \geq 1$, then for every $v \in L\mathbb{Z}^n$, $\|x - y + v\| \geq 1$. Hence,

$$F(x, y) + G(x, y) = \sum_{v \in L\mathbb{Z}^n} f(x - y + v) + g(x - y + v) \leq 0.$$

That $F \in \text{PSD}(\mathbb{T}_L)$ and $G \in \text{COP}(\mathbb{T}_L)$ follows from the following. Let μ_L be the Haar measure on $\mathbb{T}_L$, normalized so that $\mu_L(\mathbb{T}_L) = L^n$. Take a function $\rho \in L^2(\mathbb{T}_L)$, and denote the inner product on $L^2(\mathbb{T}_L)$ by $\langle \cdot, \cdot \rangle_{\mathbb{T}_L}$ and the cross-correlation on $\mathbb{T}_L$ by $\star_{\mathbb{T}_L}$. Then, by Equation (12), Lemma 10.2, and because $(\rho \star_{\mathbb{T}_L} \rho) \circ p$ is continuous, bounded, and periodic,

$$\begin{aligned} \langle F, \rho \otimes \rho \rangle_{\mathbb{T}_L} &= \langle \phi, \rho \star_{\mathbb{T}_L} \rho \rangle_{\mathbb{T}_L} = \langle f, (\rho \star_{\mathbb{T}_L} \rho) \circ p \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{T^n} \langle f(\mathbf{1}_{C_T} \star \mathbf{1}_{C_T}), (\rho \star_{\mathbb{T}_L} \rho) \circ p \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{T^n} \int_{C_T^n} \int_{C_T^n} f(x - y) (\rho \star_{\mathbb{T}_L} \rho)(p(x - y)) dy dx \\ &= \lim_{T \rightarrow \infty} \frac{1}{T^n} \int_{C_T^n} \int_{C_T^n} f(x - y) \\ &\quad \int_{\mathbb{T}_L} \rho(p(x) + z) \rho(p(y) + z) d\mu_L(z) dy dx \\ &= \lim_{T \rightarrow \infty} \frac{1}{T^n} \int_{\mathbb{T}_L} \\ &\quad \int_{C_T^n} \int_{C_T^n} f(x - y) \rho(p(x) + z) \rho(p(y) + z) d\mu_L(z) dy dx \\ &\geq 0. \end{aligned}$$

Thus, $F \in \text{PSD}(\mathbb{T}_L)$, and similarly $\langle G, \rho \otimes \rho \rangle_{\mathbb{T}_L} \geq 0$ for all $\rho \in L^2(\mathbb{T}_L)_{\geq 0}$, hence, $G \in \text{COP}(\mathbb{T}_L)$.

Finally, since F is invariant under the torus, to show that $F - J \in \text{PSD}(\mathbb{T}_L)$ it suffices that the Fourier coefficient of ϕ at 0 as a function on $\mathbb{T}_L$ is at least 1. Indeed, since f has compact support, again by Equation (12),

$$\int_{\mathbb{T}_L} \phi(x) d\mu_L(p(x)) = \int_{\mathbb{T}_L} \sum_{v \in L\mathbb{Z}^n} f(x+v) d\mu_L(p(x)) = \int_{\mathbb{R}^n} f(x) dx = \hat{f}(0) = 1.$$

This concludes the proof. $\square$

4. Exactness and convergence of the copositive hierarchy

We now prove exactness of the copositive bound $\vartheta_{\text{SP}}^X(\text{COP}(\mathbb{R}^n)_{\text{inv}})$ for any $C_c^\infty(\mathbb{R}^n) \subset X \subseteq C_0(\mathbb{R}^n)$. As a consequence, we will see that if X is the space of Schwartz functions, the bound is sharp; this is Corollary 10.7. The following inequality does most of the work to show exactness of the copositive formulation of Δ_n and convergence of the copositive hierarchy. Let the set $\mathcal{K}_{\text{fin}}(\mathbb{R}^n)$ be the family of compact subsets of $\mathbb{R}^n$ with nonzero measure, and recall the operators K_K from Chapter 3 given by $K_K f(x, y) = f(x - y)$ for all x and $y \in K$.

LEMMA 10.5. *Let $n \geq 2$ be an integer and $K \in \mathcal{K}_{\text{fin}}(\mathbb{R}^n)$. If $\mathcal{C}(K) \subseteq C_{\text{sym}}(K)$ is a convex cone and $\mathcal{C} \subseteq C_0(\mathbb{R}^n)$ is a convex cone containing $K_K^* \mathcal{C}(K)$, then*

$$\vartheta_{\text{SP}}^{C_c(\mathbb{R}^n)}(\mathcal{C}) \leq \frac{\vartheta^*(G_K, \mathcal{C}(K))}{\lambda(K)}.$$

Proof. If $\vartheta^*(G_K, \mathcal{C}(K))$ is infeasible, the conclusion follows. So, let (t, F, G) be a feasible solution to $\vartheta^*(G_K, \mathcal{C}(K))$, and let $f = K_K^* F$ and $g = K_K^* G$. Then, f and g are in $C_c(\mathbb{R}^n)$. Indeed, the support of both functions is contained in $K + K$. Since $K_K^* = \text{Av}_{\mathbb{R}^n} \text{Res}_K^*$, a straightforward calculation shows that, since F and G are uniformly continuous, f and g are continuous.

If $\|x\| \geq 1$, then

$$f(x) + g(x) = \int_{\mathbb{R}^n} (\text{Res}_K^* F(y, x+y) + \text{Res}_K^* G(y, x+y)) dy \leq 0.$$

That $K_K^* G \in \mathcal{C}(K)$ follows from the assumption on $\mathcal{C}$.

Since $F \in \text{PSD}(K)_c$, for all finite $U \subseteq \mathbb{R}^n$ the matrix $\text{Res}_U \text{Res}_K^* F$ is positive semidefinite. Since $K_K^* = \text{Av}_{\mathbb{R}^n} \text{Res}_K^*$, it follows that

$$\sum_{u, v \in U} f(u-v) x_u x_v = \int_{\mathbb{R}^n} \sum_{u, v \in U+z} \text{Res}_K^* F(u, v) x_{u-z} x_{v-z} dz \geq 0$$

for all finite $U \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^U$. By [45, Proposition 3.35], it follows that $f \in \mathcal{P}(\mathbb{R}^n)$.

Moreover, since $F - J \in \text{PSD}(K)$,

$$0 \leq \langle F - J, J \rangle = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} F(y, x + y) dy dx - \lambda(K)^2 = \widehat{f}(0) - \lambda(K)^2,$$

so that $\widehat{f}(0) \geq \lambda(K)^2$. Finally,

$$f(0) + g(0) = \int_{\mathbb{R}^n} F(z, z) + G(z, z) dz \leq t\lambda(K),$$

so that $(f(0) + g(0))/\widehat{f}(0) \leq t/\lambda(K)$. $\square$

THEOREM 10.6. *For all integers $n \geq 2$, $\Delta_n = \vartheta_{\text{SP}}^{C_c^\infty(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$.*

Proof. The inequality $\vartheta_{\text{SP}}^{C_c^\infty(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}}) \geq \Delta_n$ follows from the inequality (63) and Theorem 10.4.

For the other inequality, the strategy is to use Theorem 7.7 and the expression (62), which says that

$$\Delta_n = \lim_{L \rightarrow \infty} \frac{\alpha(G_{C_L})}{L^n} = \lim_{L \rightarrow \infty} \frac{\vartheta^*(G_{C_L}, \text{COP}(C_L)_c)}{L^n}.$$

By Theorem 3.7 and A.2, $\text{COP}(\mathbb{R}^n)_{\text{inv}} = \overline{\text{cone}} \bigcup_{A \in \mathcal{B}_{\text{fin}}} K_A^* \text{COP}(A)$, thus it satisfies the conditions of Lemma 10.5, and

$$\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}}) \leq \vartheta_{\text{SP}}^{C_c(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}}) \leq \frac{\vartheta^*(G_{C_L}, \text{COP}(C_L)_c)}{L^n}$$

for all L . Taking the limit over L shows that $\Delta_n = \vartheta_{\text{SP}}^{C_c(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$.

It remains to show that every feasible solution (f, g) of $\vartheta_{\text{SP}}^{C_c(\mathbb{R}^n)}(\mathbb{R}^n)$ can be approximated in a suitable way by smooth feasible solutions. To achieve this, make use of Leibniz's rule for differentiation under the integral: if ρ is smooth and ψ continuous with compact support, then for all i ,

$$(\partial/\partial x_i)(\rho \star \psi) = \rho \star ((\partial/\partial x_i)\psi) = -\psi \star ((\partial/\partial x_i)\rho),$$

whence the cross-correlation of a smooth function with a compactly supported continuous function is smooth.

Thus, take a smooth approximate identity $(\phi'_i)_{i \in \mathbb{N}}$ where each ϕ'_i has support in $B_{1/(2i)}(0)$ for $i \geq 2$. For example, the bump functions defined by

$$\phi'_i(x) = \begin{cases} N_i \exp^{-1/(1-2i\|x\|^2)} & \text{if } \|x\| < 1/(2i), \\ 0 & \text{otherwise,} \end{cases}$$

with N_i a suitable normalization constant, satisfy the requirements. Then, the functions $\phi_i = \phi'_i \star \phi'_i$ form a smooth approximate identity where every ϕ_i is supported in $B_{1/i}(0)$.

Recall that if $r \in \mathbb{R}$ and $\rho : \mathbb{R}^n \rightarrow \mathbb{R}$, then $\rho_r(x) = \rho(rx)$. Note that $f_r \in P(\mathbb{R}^n)$ and $g_r \in \text{COP}(\mathbb{R}^n)_{\text{inv}}$ for all $r \in \mathbb{R}$. Let $r_i = 1/(1 - 1/i)$. Then, by Lemma 10.1 the function $r_i^n \phi_i \star f_{r_i}$ is in $P(\mathbb{R}^n)$ and $r_i^n \phi_i \star g_{r_i}$ is in $\text{COP}(\mathbb{R}^n)_{\text{inv}}$. Moreover, all these functions are smooth with compact support, and by Lemma 10.3 they are nonpositive on x such that $\|x\| \geq 1$, converge uniformly to f and g respectively, and $\int r_i^n \phi_i \star f_{r_i}(x) dx = \int f(x) dx$ for all i . This shows the theorem. $\square$

COROLLARY 10.7. *Let $n \geq 2$ be an integer. If X is a vector space such that $C_c^\infty(\mathbb{R}^n) \subseteq X \subseteq C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, then $\Delta_n = \vartheta_{\text{SP}}^X(\text{COP}(\mathbb{R}^n)_{\text{inv}})$. In particular, $\Delta_n = \vartheta_{\text{SP}}^{\mathcal{S}(\mathbb{R}^n)}(\text{COP}(\mathbb{R}^n)_{\text{inv}})$.*

Proof. This follows immediately from Theorem 10.4, Theorem 10.6, and Equation (63). $\square$

We conclude with the convergence of a copositive hierarchy for Δ_n , which is a simple corollary of all the work we have done before.

THEOREM 10.8. *If $n \geq 2$ is an integer, then*

$$\Delta_n = \lim_{r \rightarrow \infty} \vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\check{C}_r(\mathbb{R}^n)_{\text{inv}}) = \lim_{r \rightarrow \infty} \vartheta_{\text{SP}}^{C_c(\mathbb{R}^n)}(\check{C}_r(\mathbb{R}^n)_{\text{inv}}).$$

Proof. Convergence of the copositive hierarchy for compact packing graphs, Theorem 7.8, says that

$$\Delta_n = \lim_{L \rightarrow \infty} \lim_{r \rightarrow \infty} \frac{\vartheta^*(G_{C_L}, C_r(C_L)_c)}{L^n}.$$

By Lemma 3.6 and Theorem A.2, the inclusions $\check{C}_r(\mathbb{R}^n)_{\text{inv}} \supseteq K_K^* C_r(K)$ hold for all $K \in \mathcal{K}_{\text{fin}}(\mathbb{R}^n)$, hence by Lemma 10.5, for all L ,

$$\vartheta_{\text{SP}}^{C_c(\mathbb{R}^n)}(\check{C}_r(\mathbb{R}^n)_{\text{inv}}) \leq \frac{\vartheta^*(G_{C_L}, C_r(C_L)_c)}{L^n}.$$

First taking the limit over r and then over L concludes the proof. $\square$

5. Comparison to other known bounds

Can we use these copositive hierarchies to improve known bounds? A first step would be to see whether it improves on the Cohn-Elkies bound, which in our notation is $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\{0\})$; clearly, it is at most as strong as $\vartheta_{\text{SP}}^{C_0(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)}(\check{C}_r(\mathbb{R}^n)_{\text{inv}})$ for each r .

To improve on the Cohn-Elkies bound, it is not necessary to optimize over all of $\check{C}_r(\mathbb{R}^n)_{\text{inv}}$; it is enough to only take a suitable subset. However, finding feasible solutions that are not feasible for the Cohn-Elkies bound seems to be difficult, and is work in progress.

Conclusion and discussion

Exact completely positive and copositive formulations were introduced for four classes of extremal problems in geometry that are modeled as an independence number problem on a graph. These formulations were used to define a completely positive or copositive hierarchy, and prove they converge to the independence number. Table 2 shows the current status of these results. These hierarchies were then compared to extensions of the moment and block moment hierarchy for two classes of these problems. Two applications were worked out in detail.

Discussion of the theory

Finite measure spaces and density. The exactness statements of the completely positive hierarchy for measurable graphs on a finite countably generated measure space are quite complete. In both the thick as in the homogeneous setting, we used the concept of density points to move from a set that might have a zero-measure subset of edges to an independent set. The use of density seems inevitable. However, the notions of density system and thick set as presented in Chapter 4 are crude, and there might be more refined definitions. This would form an interesting topic of study.

For example, consider the following argument in the homogeneous setting. Let $O(n)$ be the orthogonal group on $\mathbb{R}^n$ with Haar measure μ , let it act on S^{n-1} , and let ν be the pushforward onto S^{n-1} of μ . Clearly, the edge set $E = \{(x, y) : x, y \in S^{n-1}, x^\top y = 0\}$ has measure $\nu^2(E) = 0$, hence it is not thick.

However, we can directly a measure on E that has better properties. Let x_0 and $y_0 \in S^{n-1}$ be any choice of points such that $x_0^\top y_0 = 0$. The orbit of a pair under the diagonal action of $O(n)$ is characterized by the inner product of the coordinates, thus $E = O(n)\{(x_0, y_0)\}$. We can then define a measure on E by

$$\rho(S) = \mu(\{T : (Tx_0, Ty_0) \in S\}).$$

It is fair to say that E is thick with respect to ρ ; Castro-Silva [22] used such arguments to formulate zero-measure-removal lemmas for S^{n-1} and $\mathbb{R}^n$ for similar edge sets. Thus, it seems that thickness should be measured not by a measure on V , but by a measure on E .

Converging hierarchies in the thick setting. In contrast to the completeness of the exactness statements, Table 2 also shows several interesting gaps. For a measurable graph on a finite and countably generated measure space, the view adopted in this thesis is that the programs essentially optimize over a set of p -integrable functions. However, in neither setting does the

	<i>Exact completely positive / copositive formulation</i>	<i>Converging completely positive / copositive hierarchy</i>
<i>Measurable k-uniform hypergraph</i>		
• with a thick edge set	This thesis.	If $k = 2$; this thesis.
• compact homogeneous	If the group admits a right-invariant density metric; this thesis and [36].	If $k = 2$ on a continuous, compact, two-point homogeneous space with real dimension at least 2; this thesis and [12].
<i>Compact packing graph</i>	If the vertex set is a metric space [39].	If the vertex set is a metric space [57, 56].
<i>Distance-avoiding sets in $\mathbb{R}^n$</i>	If $n \geq 2$ [36].	If $n \geq 2$; this thesis.
<i>Sphere packing in $\mathbb{R}^n$.</i>	If $n \geq 2$; this thesis.	If $n \geq 2$; this thesis.

TABLE 2. Known sufficient conditions for the existence of an exact completely positive or copositive formulation a corresponding hierarchy for the independence number of a (measurable) graph, and where they were first described. It is assumed everywhere that the independence number is nonzero, and that underlying measure space of a measurable graph is assumed to be finite and countably generated.

proof of convergence of the completely positive hierarchy extend to k -uniform hypergraphs with $k > 2$.

The shortcoming of the L^p formulation can be understood by thinking of the normalization constraint of ϑ_{big} in the thick setting. For a measurable graph $G = (V, E)$ with thick edge set E , the normalization $\|A\|_{\mathcal{B}^1} = 1$ bounds the feasible region, since it bounds the L^2 -norm, which we used to show that the feasible regions of the programs $\vartheta_{\text{big}}(G, C_r(V)^*)$ lie in a common compact space. However, this result can be executed analogously in the space of trace-class operators $\mathcal{B}^1(L^2(V))$, since the edge constraints can be written in terms of linear functionals in a predual of the trace class: the space of compact operators. Thus, using Banach-Alaoglu, we obtain the same compactness result under the weak* topology on $\mathcal{B}^1(L^2(V))$ with respect to the duality with the space of compact operators, which is a more specialized statement. This seems to be a more natural setting, since we only consider trace class kernels to begin with.

For this reason, it seems that the correct extension to k -uniform hypergraphs with $k \geq 2$ would use a normalization constraint given by a *symmetric nuclear norm*

$$\|A\|_{\mathcal{N}_p} = \inf \left\{ \sum_{n \in \mathbb{N}} |\lambda_n| \|\phi_n\|_p^k : A = \sum_{n \in \mathbb{N}} \lambda_n \phi_n^{\otimes k} \right\}.$$

The infimum is over every such expansion of A with $\lambda_n \in \mathbb{R}$ and $\phi_n \in L^p(V)$ for suitable p . Whether bounding this norm forces all required properties requires closer investigation.

Converging hierarchies in the homogeneous setting. As mentioned in Section 5.4, the convergence of the completely positive hierarchy in the homogeneous setting can be readily extended to the continuous, compact, two-point homogeneous spaces with real dimension at least 2: the sphere, the real, complex and quaternionic projective spaces, and the octonionic projective plane. Analogues on k -uniform homogeneous hypergraphs on these spaces also seem within reach, using the results on configuration-avoiding sets by Castro-Silva [22].

The challenge lies in extending the convergence result to other homogeneous spaces. As explained in Section 5.4 the method we used already fails for homogeneous graphs G on the circle, even though $\vartheta_{\text{small}}(G, C_0(S^1)) = \alpha(G)$.

If the vertex set is a compact group, we can limit the feasible regions to a space spanned by the functions of positive type on the group: the Fourier algebra. This approach is already present in the graph case: when V is compact, the image of the trace class kernels under the averaging operator lies

in the Fourier algebra. It would be interesting to better understand the role of the Fourier algebra, for example to better understand the relation between the behavior of the hierarchies and the representation theory of the group. It is, however, unclear whether this will help to extend convergence of the completely positive hierarchy to a larger class of compact groups. On the positive side, there are k -tensor analogues of the Fourier algebra [99], which could perhaps form a setting for hierarchies for k -uniform hypergraphs on S^{n-1} and related spaces.

Compact packing hypergraphs. How to define compact packing hypergraphs is still an open problem. There are simple conditions on a k -uniform hypergraph with $k > 2$ that would imply all results on copositive programming in this thesis. For example, if V is a metric space with an edge set $E \subseteq \text{Sub}_{=k}(V)$, asking for all $r \geq 0$ that $\mathcal{I}_r = \bigsqcup_{i=0}^r \mathcal{I}_i$ as topological spaces under the restriction of the standard topology would be sufficient, as is the case for compact packing graphs. On the other hand, this definition seems too limiting, as the application studied in Chapter 8 shows. In the case of that particular chapter, simply replacing the standard topology on $\text{Sub}_r(V)$ by the disjoint union topology $\bigsqcup_{i=0}^r \text{Sub}_i(V)$ seems satisfactory, but the precise details need more thought, since such a space differs topologically from $\text{Sub}_r(V)$; most notably, compactness of $\text{Sub}_r(V)$ under this topology is no longer implied by compactness of V . See also the discussion in Section 5.6.

Euclidean space. It is likely that the results for problems on Euclidean space can be extended to problems on k -uniform homogeneous hypergraphs with vertex set $\mathbb{R}^n$ by similar methods. Again, to do this properly, a first step would be to investigate the relation of these optimization problems and the Fourier algebra.

In Chapter 10 we saw that the optimal sphere-packing density can be approximated using copositive Schwartz functions. It is not clear that this also holds when one only considers those Schwartz functions that lie in the cone $\bigcup_r \check{C}_r(\mathbb{R}^n)_{\text{inv}}$. It would be interesting to know whether this is true, since these are the functions that are used in practice.

Discussion of the computations

For distance-avoiding set on $\mathbb{R}^n$, the completely positive hierarchy offers the first bound taking correlations between more than two points into account. Attempts were made to implement a version of a related three-point bound for the 1-avoiding-sets problem on $\mathbb{R}^2$. Although in theory this would lead to bounds that are stronger than the optimization method introduced by

DeCorte, Oliveira, and Vallentin [36], in practice it turns out that within a reasonable runtime, there is no improvement. Whether this is inherent to this optimization problem, or whether it has to do with the way the problem is modeled is not clear. This is still work in progress.

A similar conclusion holds for the sphere-packing problem, see also the discussion in Section 10.5. Although new feasible solutions coming from $\check{C}_r(\mathbb{R}^n)_{\text{inv}}$ were found, none of these make an improvement on the objective value. This too requires more experimentation.

APPENDIX A

Preliminaries from analysis

This appendix contains the preliminaries to this thesis that I consider more elementary, or which did not have a natural place in the relevant chapters.

1. Locally convex spaces, duality, and cones

Our reference for all matters convex is the book *Convexity* by Simon [97]. All vector spaces are over $\mathbb{R}$, all topological vector spaces are Hausdorff.

A *locally convex space* is a topological vector space X with a topology that is generated by a set of seminorms $\{\rho_j\}_{j \in J}$. This means that a net $(x_i)_{i \in I}$ converges to x if and only if for all $j \in J$ the net $(\rho_j(x_i - x))_{i \in I}$ converges to 0. The space X is Hausdorff if and only if $\bigcap_j \{x : \rho_j(x) = 0\} = \{0\}$.

Let X and Y be vector spaces. A *duality* between X and Y is a bilinear map $\langle \cdot, \cdot \rangle : X \times Y \rightarrow \mathbb{R}$ such that if $x \neq 0$ there is a y such that $\langle x, y \rangle \neq 0$, and if $y \neq 0$ there is an x such that $\langle x, y \rangle \neq 0$. If there exists a duality between X and Y , we call (X, Y) a *dual pair* of vector spaces.

Let (X, Y) be a dual pair of vector spaces with duality $\langle \cdot, \cdot \rangle$. Then X is a locally convex space under the topology generated by the set of seminorms $\{|\langle \cdot, y \rangle|\}_{y \in Y}$. This topology is called the $\sigma(X, Y)$ -*topology*. We will call it the *weak topology* when the duality is clear. The seminorms $\{|\langle x, \cdot \rangle|\}_{x \in X}$ make Y into a locally convex space. The induced topology on Y is called the *weak* topology*; it is the same as the $\sigma(Y, X)$ -topology when we interchange X and Y everywhere.

The map $y \mapsto \langle \cdot, y \rangle$ identifies Y with a subspace of X' . Then, $X^* = Y$ under the weak topology [97, Proposition 5.1]. Any locally convex topology on X such that $X^* = Y$ is called an (X, Y) -*dual topology*. The weak topology is by definition the weakest dual topology.

A linear map $A : X_1 \rightarrow X_2$ induces a unique *linear adjoint* A' from X'_2 to X'_1 by $A'(l)(x) = l(A(x))$ for all $x \in X_1$ and $l \in X'_2$. With these notions, continuity of A under weak topologies is reduced to an algebraic condition. If A^* from the following lemma exists and is continuous, it is called the *continuous adjoint* of A . In this case, $(A^*)^* = A$.

LEMMA A.1. *Let (X_1, Y_1) and (X_2, Y_2) be dual pairs of vector spaces with respective dualities $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. Let $A : X_1 \rightarrow X_2$ be a linear map.*

If A is continuous under a dual topology on X_1 and any topology on X_2 such that $\langle \cdot, y \rangle_2$ is continuous for all $y \in Y_2$, then $A'(Y_2) \subseteq Y_1$.

If $A'(Y_2) \subseteq Y_1$, X_1 and Y_2 are equipped with dual topologies, and X_2 and Y_1 are equipped with weak topologies, then both A and the map $A^ : Y_2 \rightarrow Y_1$ given by $A^*(y) = A'(y)$ are continuous.*

Proof. Let A be continuous with respect to an (X_1, Y_1) -dual topology on X_1 and any topology on X_2 . Suppose that for all $y \in Y_1$, the functionals $\langle \cdot, y \rangle_2$ are continuous, and fix $y \in Y_2$. Then, the functional $x \mapsto \langle A(x), y \rangle_2$ is continuous. The definition of A' and the identification $Y_1 = X_1^*$ together warrant $\langle A(x), y \rangle_2 = \langle x, A'(y) \rangle_1$, and $A'(y) \in Y_1$.

If $A'(Y_2) \subseteq Y_1 = X_1^*$, the functional $x \mapsto \langle A(x), y \rangle_2 = \langle x, A'(y) \rangle_1$ is continuous for all $y \in Y_2$. So, if $(x_i)_{i \in I}$ is a converging net in X_1 , then $(A(x_i))_{i \in I}$ converges weakly in X_2 , and A is continuous under a dual topology on X_1 and the weak topology on X_2 .

Continuity of $A^* : Y_2 \rightarrow Y_1$ follows by applying the previous argument to A'' restricted to Y_2 , and $A'' = A$. $\square$

Let X and Y be vector spaces, and suppose X has an (X, Y) -dual topology. Recall that here, all cones are convex. The closure of a convex set $S \subseteq X$ under the chosen dual topology is equal to its closure under the weak topology [97, Theorem 5.2]. Let $S \subseteq X$ be any subset. The *dual cone* of S is

$$S^* = \{ y \in Y : \langle x, y \rangle \geq 0 \text{ for all } x \in S \}.$$

It is a convex cone and closed under all (Y, X) -dual topologies on Y . Since the duality is bilinear and continuous under dual topologies, $S^* = (\overline{\text{cone}} S)^*$. Taking the dual reverses inclusion: if $S_1 \subseteq S_2$ then $S_2^* \subseteq S_1^*$.

We refer to the following identities many times.

THEOREM A.2. *Let (X_1, Y_1) and (X_2, Y_2) be dual pairs of vector spaces, each of X_1 , X_2 , Y_1 , and Y_2 with dual topologies. Let $A : X_1 \rightarrow X_2$ be a continuous linear map with continuous adjoint A^* , let $S_X \subseteq X_1$, let $S_Y \subseteq Y_1$, and let $\{S_i\}_{i \in I}$ be an arbitrary family of subsets of X_1 . Then:*

- (i) $(S_X^*)^* = \overline{\text{cone}} S_X$;
- (ii) $(AS_X)^* = A^{*-1} S_X^*$;
- (iii) $(A^{*-1} \overline{\text{cone}} S_Y)^* = \text{cl } AS_Y^*$;
- (iv) $(\bigcap_{i \in I} S_i^*)^* = \overline{\text{cone}} \bigcup_{i \in I} S_i$.

Proof. (i) This follows from $S_X^* = (\overline{\text{cone}} S_X)^*$ and [97, Proposition 5.5, Example 5.9], since $\overline{\text{cone}} S_X$ is convex and contains 0.

(ii) Denote by $\langle \cdot, \cdot \rangle_i$ be the duality between X_i and Y_i . Let $s \in S_X$ and $y \in Y_2$. From $\langle As, y \rangle_2 = \langle s, A^*y \rangle_1$ it follows immediately that $\langle As, y \rangle_2 \geq 0$ if and only if $A^*y \in S_X^*$. So $(AS_X)^* = A^{*-1}S_X^*$.

(iii) Since A is linear and S_Y^* is a convex cone, $AS_Y^* = \text{cone } AS_Y^*$. Then, by (i) and (ii):

$$\text{cl } AS_Y^* = \overline{\text{cone } AS_Y^*} = ((AS_Y^*)^*)^* = (A^{*-1}(S_Y^*)^*)^* = (A^{*-1} \overline{\text{cone } S_Y})^*.$$

(iv) The inclusion $\overline{\text{cone } \bigcup_i S_i} \subseteq (\bigcap_i S_i^*)^*$ holds if $\bigcup_i S_i \subseteq (\bigcap_i S_i^*)^*$ holds. Indeed, each S_i^* is a closed convex cone, and an intersection of closed convex cones is again a closed convex cone. Then, apply $\overline{\text{cone}}$ to both sides.

Let $j \in I$, $s \in S_j$, and $y \in \bigcap_i S_i^*$. In particular, $y \in S_j^*$, so $\langle s, y \rangle_1 \geq 0$. Thus, $\bigcup_i S_i \subseteq (\bigcap_i S_i^*)^*$ and $\overline{\text{cone } \bigcup_i S_i} \subseteq (\bigcap_i S_i^*)^*$.

Suppose that $y \in Y_1 \setminus S_j^*$ for some $j \in I$. By the definition of the dual there exists an $s \in S_j$ such that $\langle s, y \rangle_1 < 0$. But S_j is contained in $\overline{\text{cone } \bigcup_i S_i}$, so y is not in $(\overline{\text{cone } \bigcup_i S_i})^*$. Then, $(\overline{\text{cone } \bigcup_i S_i})^* \subseteq \bigcap_i S_i^*$. By taking the dual again and applying (i), it follows that $(\bigcap_i S_i)^* = \overline{\text{cone } \bigcup_i S_i}$. $\square$

2. p -integrable functions

All functions and measures here are real valued.

Let $(V, \mathcal{A}, \mu)$ be a measure space with σ -algebra $\mathcal{A}$ and measure μ . The integrals $\|f\|_p = \mu(|f|^p)^{1/p}$ are defined for all $0 < p < \infty$. The *essential supremum* of f is $\|f\|_\infty = \inf\{a : \mu(\{x : f(x) \geq a\}) = 0\}$. It is the smallest upper bound on f outside a set of measure 0.

For $0 < p < \infty$ a function f from V to $\mathbb{R}$ is called *p -integrable* if $\|f\|_p$ is finite. When $\|f\|_\infty$ is finite f is called *essentially bounded*. Sometimes we will just say that f is bounded to mean the same. We say that two measurable functions are μ -*equivalent* if they are equal *almost everywhere*, that is, they are equal outside a set of measure 0. For $0 < p \leq \infty$ the space of equivalence classes of real-valued p -integrable functions on $(V, \mathcal{A}, \mu)$ is denoted $L^p(V, \mathcal{A}, \mu)$ or simply $L^p(V)$.

The map $f \mapsto \|f\|_p$ induces a norm on $L^p(V)$ whenever p is at least 1. The spaces $L^p(V)$ with $1 \leq p \leq \infty$ are Banach spaces under their respective norms. Under this norm, the dual of $L^p(V)$ is $L^q(V)$ with $1 \leq p, q \leq \infty$ such that $1/p + 1/q = 1$, or if $p = 1$ and $q = \infty$. For such p and q , denote the duality between $L^p(V)$ and $L^q(V)$ by $\langle \cdot, \cdot \rangle$. We have the following inequalities.

If $1 \leq p, q \leq \infty$ such that $1/p + 1/q = 1$, then $|\langle f, g \rangle| \leq \|f\|_p \|g\|_q$. This is called *Hölder's inequality*.

If V_1 and V_2 are two σ -finite measure spaces with respective measures μ_1 and μ_2 , and $F \in V_1 \times V_2 \rightarrow \mathbb{R}$ is measurable, then for all $1 \leq p < \infty$, we have

$$\begin{aligned} \left(\int_{V_2} \left| \int_{V_1} F(v, w) d\mu_1(v) \right|^p d\mu_2(w) \right)^{1/p} \\ \leq \int_{V_1} \left(\int_{V_2} |F(v, w)|^p d\mu_2(w) \right)^{1/p} d\mu_1(v), \end{aligned}$$

and similar for $p = \infty$. This is called *Minkowski's inequality*.

Let V be a measure space with measure μ , W a set equipped with a σ -algebra, and $q : V \rightarrow W$ a measurable function. Then q induces a measure $q_*\mu$ on W by $\int_W f(w) dq_*\mu(w) = \int_V f(q(v)) d\mu(v)$, called the *pushforward* of μ under q . This is nothing more than the change of variables formula. A function $f \in L^p(W, q_*\mu)$ if and only if $f \circ q \in L^p(V, \mu)$.

3. Radon measures

Cohn [25] gives a complete overview of elementary measure theory. Let V be a locally compact Hausdorff space. A *Borel measure* on V is a measure on the Borel algebra, which we denote $\mathcal{B}$. A *Radon measure* is a measure on $\mathcal{B}$ that is finite on compact sets, outer regular on Borel sets, and inner regular on open sets. That is, μ is Radon if and only if it is finite on compact sets, for every Borel set A

$$\mu(A) = \inf \{ \mu(U) : A \subseteq U, U \text{ open} \},$$

and if $U \subseteq V$ is open, then

$$\mu(U) = \sup \{ \mu(K) : K \subseteq U, K \text{ compact} \}.$$

We denote the space of Radon measures by $M(V)$. It is isometrically isomorphic to the continuous dual of $C_0(V)$, with duality $\langle f, \mu \rangle = \mu(f)$ for all $\mu \in M(V)$ and $f \in C_0(V)$.

4. Spaces of continuous linear operators

Suppose X and Y are topological vector spaces. Denote the set of continuous linear operators $X \rightarrow Y$ by $\mathcal{L}(X, Y)$. This space has many useful topologies. Here is a list of some of them. If $X = Y$ is a Hilbert space, and the duality is the inner product, the list goes from the weakest topology to the strongest.

When (Y, Z) is a dual pair, the *weak operator topology* is the topology of pointwise convergence under the weak topology on Y . It is the topology generated by the set of seminorms $\{A \mapsto |\langle Ax, z \rangle|\}_{(x, z) \in X \times Z}$. Convergence under

this topology is called *weak convergence*, or we say a net of operators *converges weakly*.

When Y is a normed space with norm $\|\cdot\|$, the *strong operator topology* is the topology of pointwise convergence under the norm. It is the topology generated by the seminorms $\{A \mapsto \|Ax\|\}_{x \in X}$. Convergence under this topology is called *strong convergence*, or we say that a net of operators *converges strongly*.

When X and Y are normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, then A is in $\mathcal{L}(X, Y)$ if and only if it is bounded. That is, if and only if there is a $c \geq 0$ for which $\|Ax\|_Y \leq c\|x\|_X$ for all $x \in X$. Denote the infimum over such c by $\|A\|_{\text{op}}$, the *operator norm* of A . Equivalently,

$$\|A\|_{\text{op}} = \sup\{\|Ax\|_Y : x \in X, \|x\|_X = 1\}.$$

The topology of the operator norm is the topology of uniform convergence on bounded sets.

These topologies are all defined by seminorms, so they turn $\mathcal{L}(X, Y)$ into a locally convex space. When X and Y are normed spaces, denote $\mathcal{L}(X, Y)$ by $\mathcal{B}(X, Y)$. This is a normed space under the operator norm. When Y is a Banach space, $\mathcal{B}(X, Y)$ is a Banach space. Denote $\mathcal{B}(X) = \mathcal{B}(X, X)$.

If $(V, \mathcal{A}, \mu)$ is a measure space, then $L^2(V)$ is a Hilbert space, and the space of bounded operators $\mathcal{B}(L^2(V))$ has two subspaces that are important to us. The first is the space of *Hilbert-Schmidt operators*, which can be identified as a Hilbert space with $L^2(V^2)$. An element of $L^2(V^2)$ is called a *kernel*.

The second space of interest is the *trace class* $\mathcal{B}^1(L^2(V))$. We say that a kernel K is of trace-class if there are orthonormal sets Φ and Ψ and complex numbers $\lambda_{\phi\psi}$ such that

$$K = \sum_{\phi \in \Phi, \psi \in \Psi} \lambda_{\phi\psi} \phi \otimes \psi$$

such that $\sum_{\phi, \psi} |\lambda_{\phi\psi}| < \infty$. If finite, the latter sum is independent of Φ and Ψ , and is known as the *trace norm* of K ; we will denote it $\|K\|_{\mathcal{B}^1}$. It can be shown that $\|K\|_2 \leq \|K\|_{\mathcal{B}^1}$. See [96] for a detailed account

5. Invariant measures

We mostly follow [45]. A *topological group* is a group Γ with a Hausdorff topology such that the multiplication and inversion maps are continuous. A *locally compact group* is a topological group with a locally compact topology.

Denote the Borel algebra of Γ by $\mathcal{B}$. A *left Haar measure* on Γ is a nonzero Radon measure μ on Γ such that for every $E \in \mathcal{B}$ and $g \in \Gamma$, we have $\mu(gE) = \mu(E)$. If Γ is locally compact, a left Haar measure exists and is

unique up to a constant multiple. We denote a choice of left Haar measure by μ . The group is called *unimodular* if the left Haar measure is also right invariant, that is, if, for all $g \in \Gamma$ and all $E \in \mathcal{B}$, we have $\mu(gE) = \mu(Eg) = \mu(E)$. If Γ is σ -compact, $(\Gamma, \mathcal{B}, \mu)$ is σ -finite.

If $H = \text{Stab}(v_0)$ and we identify V with Γ/H , then V is locally compact and Hausdorff. When both Γ and H are unimodular, then there is a Γ -invariant Radon measure on Γ/H . Indeed, if μ is a Haar measure on Γ and ν a Haar measure on H , then $f \mapsto (x \mapsto \int_H f(xh) d\nu(h))$ is a surjective map $C_c(\Gamma) \rightarrow C_c(\Gamma/H)$. Let $p : \Gamma \rightarrow \Gamma/H$ be the quotient map. Theorem 2.51 of [45] implies that if both Γ and H are unimodular, there exists a Γ -invariant Radon measure ω on Γ/H , i.e. for all $g \in \Gamma$ and all measurable subsets $E \subseteq V$, we have $\omega(gE) = \omega(Eg) = \omega(E)$. This measure is unique up to a constant multiple, and is related to μ and ν by the Fubini-type formula

$$\int_{\Gamma} f(g) d\mu(g) = \int_{\Gamma/H} \int_H f(vh) d\nu(h) d\omega(p(v)),$$

for all $f \in C_c(\Gamma)$. The formula extends to all integrable functions f , as Reiter and Stegeman [90, §3.4] discuss in detail.

On the other hand, when H is compact, the quotient map p induces the pushforward measure $p_*\mu$, which is Radon. Since $p_*\mu$ defines a continuous linear functional on $C_c(\Gamma/H)$ and is Γ -invariant, it is a Γ -invariant Radon measure, hence a multiple of ω . We therefore set $\omega = p_*\mu$ if H is compact, and call it the *quotient measure* on Γ/H .

Bibliography

- [1] C. Adams, *Copositivity in infinite dimension*, PhD thesis, Universität Trier, 2018. (Cited on page 15.)
- [2] G. Ambrus, A. Csiszárík, M. Matolcsi, D. Varga, and P. Zsámboki, *The density of planar sets avoiding unit distances*, Math. Program. **207** (2024), no. 1-2, 303–327. (Cited on pages 3 and 147.)
- [3] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, Encyclopedia of Mathematics and its Applications, vol. 71, Cambridge University Press, Cambridge, 1999. (Cited on pages 72, 90, and 119.)
- [4] C. Bachoc, B. Bekker, P. Moustrou, and F. M. de Oliveira Filho, *Obtuse almost-equiaangular sets*, preprint (2025). arXiv:2504.11086 (Cited on pages 3, 111, 121, 123, and 143.)
- [5] C. Bachoc, G. Nebe, F. M. de Oliveira Filho, and F. Vallentin, *Lower bounds for measurable chromatic numbers*, Geom. Funct. Anal. **19** (2009), no. 3, 645–661. (Cited on pages 2 and 77.)
- [6] C. Bachoc and F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. **21** (2008), no. 3, 909–924. (Cited on pages 3, 8, 79, 96, 118, and 120.)
- [7] M. Balko, A. Pór, M. Scheucher, K. Swanepoel, and P. Valtr, *Almost-equidistant sets*, Graphs Combin. **36** (2020), no. 3, 729–754. (Cited on pages 125, 133, and 143.)
- [8] E. Bannai and E. Bannai, *A survey on spherical designs and algebraic combinatorics on spheres*, European J. Combin. **30** (2009), no. 6, 1392–1425. (Cited on pages 136 and 137.)
- [9] R. Barrio, *Rounding error bounds for the Clenshaw and Forsythe algorithms for the evaluation of orthogonal polynomial series*, J. Comput. Appl. Math. **138** (2002), no. 2, 185–204. (Cited on page 91.)
- [10] A. Barvinok, *A Course in Convexity*, Graduate Studies in Mathematics, vol. 54, American Mathematical Society, Providence, RI, 2002. (Cited on page 104.)
- [11] B. Bekka, P. de la Harpe, and A. Valette, *Kazhdan’s property (T)*, New Mathematical Monographs, vol. 11, Cambridge University Press, Cambridge, 2008. (Cited on page 148.)
- [12] B. Bekker, O. Kuryatnikova, F. M. de Oliveira Filho, and J. C. Vera Lizcano, *Optimization hierarchies for distance-avoiding sets in compact spaces*, Trans. Amer. Math. Soc. **379** (2026), no. 1, 33–70. (Cited on pages 2, 15, 20, 23, 48, 65, 70, 77, 82, 103, and 170.)

- [13] B. Bekker and F. M. de Oliveira Filho, *On the convergence of the k -point bound for topological packing graphs*, preprint (2023). arXiv:2306.02725 (Cited on pages 3 and 96.)
- [14] K. Bezdek and Z. Lángi, *Almost equidistant points on S^{d-1}* , vol. 39, 1999, Discrete geometry and rigidity (Budapest, 1999), pp. 139–144. (Cited on pages 111, 113, 121, 126, 134, and 135.)
- [15] D. Bilyk, D. Ferizović, A. Glazyrin, R. W. Matzke, J. Park, and O. Vlasiuk, *Optimal measures for multivariate geometric potentials*, preprint (2023). arXiv:2303.14258 (Cited on pages 112 and 144.)
- [16] ———, *Optimizers of three-point energies and nearly orthogonal sets*, Proc. Amer. Math. Soc. **152** (2024), no. 9, 4015–4033. (Cited on pages 112 and 144.)
- [17] I. M. Bomze, *On standard quadratic optimization problems*, vol. 13, 1998, pp. 369–387. (Cited on page 1.)
- [18] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. J. Quist, and T. Terlaky, *On copositive programming and standard quadratic optimization problems*, vol. 18, 2000, pp. 301–320. (Cited on page 1.)
- [19] I. M. Bomze and E. de Klerk, *Solving standard quadratic optimization problems via linear, semidefinite and copositive programming*, J. Global Optim. **24** (2002), no. 2, 163–185. (Cited on pages 1, 24, and 25.)
- [20] C. Brislawn, *Traceable integral kernels on countably generated measure spaces*, Pacific J. Math. **150** (1991), no. 2, 229–240. (Cited on page 38.)
- [21] B. Bukh, *Measurable sets with excluded distances*, Geom. Funct. Anal. **18** (2008), no. 3, 668–697. (Cited on page 148.)
- [22] D. Castro-Silva, *Geometrical sets with forbidden configurations*, Forum Math. Sigma **11** (2023), Paper No. e44, 45. (Cited on pages 74, 170, and 171.)
- [23] D. Castro-Silva, F. M. de Oliveira Filho, L. Slot, and F. Vallentin, *A recursive Lovász theta number for simplex-avoiding sets*, Proc. Amer. Math. Soc. **150** (2022), no. 8, 3307–3322. (Cited on pages 49 and 112.)
- [24] ———, *A recursive theta body for hypergraphs*, Combinatorica **43** (2023), no. 5, 909–938. (Cited on page 112.)
- [25] D. L. Cohn, *Measure theory*, second ed., Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser/Springer, New York, 2013. (Cited on pages 22 and 178.)
- [26] H. Cohn, *New upper bounds on sphere packings. II*, Geom. Topol. **6** (2002), 329–353. (Cited on page 156.)
- [27] H. Cohn and N. Elkies, *New upper bounds on sphere packings. I*, Ann. of Math. (2) **157** (2003), no. 2, 689–714. (Cited on pages 4, 155, 156, and 157.)
- [28] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, *The sphere packing problem in dimension 24*, Ann. of Math. (2) **185** (2017), no. 3, 1017–1033. (Cited on pages 4, 155, and 156.)
- [29] H. Cohn, D. de Laat, and N. Leijenhorst, *Optimality of spherical codes via exact semidefinite programming bounds*, preprint (2024). arXiv:2403.16874 (Cited on pages 96 and 121.)
- [30] H. Cohn, D. de Laat, and A. Salmon, *Three-point bounds for sphere packing*, (2022). arXiv:2206.15373 (Cited on pages 155 and 156.)
- [31] H. Cohn and A. Salmon, *Sphere packing bounds via rescaling*, (2021). arXiv:2108.10936 (Cited on page 155.)

- [32] J. Conway, *A Course in Functional Analysis*, Graduate Texts in Mathematics, Springer, New York, NY, 2010. (Cited on pages 58, 59, 63, 71, 73, and 153.)
- [33] H. T. Croft, *Incidence incidents*, Eureka, vol. 30, The Archimedean, 1967, pp. 22–26. (Cited on pages 147 and 195.)
- [34] L. Deaett, *The minimum semidefinite rank of a triangle-free graph*, Linear Algebra Appl. **434** (2011), no. 8, 1945–1955. (Cited on pages 111, 113, 132, and 133.)
- [35] E. DeCorte, *The Eigenvalue Method for Extremal Problems on Infinite Vertex-Transitive Graphs*, PhD thesis, Delft University of Technology, 2015. (Cited on page 77.)
- [36] E. DeCorte, F. M. de Oliveira Filho, and F. Vallentin, *Complete positivity and distance-avoiding sets*, Math. Program. **191** (2022), no. 2, 487–558. (Cited on pages 15, 47, 49, 51, 56, 61, 72, 74, 77, 86, 88, 89, 147, 148, 149, 150, 170, and 173.)
- [37] P. Delsarte, *Bounds for unrestricted codes, by linear programming*, Philips Res. Rep. **27** (1972), 272–289. (Cited on page 95.)
- [38] P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geometriae Dedicata **6** (1977), no. 3, 363–388. (Cited on pages 3 and 95.)
- [39] C. Dobre, M. Dür, L. Frerick, and F. Vallentin, *A copositive formulation for the stability number of infinite graphs*, Math. Program. **160** (2016), no. 1-2, 65–83. (Cited on pages 101, 102, 104, 105, 161, and 170.)
- [40] J. L. Doob, *Stochastic processes*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1990, Reprint of the 1953 original, A Wiley-Interscience Publication. (Cited on page 21.)
- [41] R. E. Edwards and G. I. Gaudry, *Littlewood-Paley and Multiplier Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 90, Springer-Verlag, Berlin-New York, 1977. (Cited on pages 20, 21, and 60.)
- [42] P. Erdős, *Problems and results in combinatorial geometry*, Discrete geometry and convexity (New York, 1982), Ann. New York Acad. Sci., vol. 440, New York Acad. Sci., New York, 1985, pp. 1–11. (Cited on pages 3 and 147.)
- [43] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), no. 2, 113–121. (Cited on page 54.)
- [44] M. Fiedler, *Matrices and Graphs in Geometry*, Encyclopedia of Mathematics and its Applications, vol. 139, Cambridge University Press, Cambridge, 2011. (Cited on pages 115 and 124.)
- [45] G. B. Folland, *A course in abstract harmonic analysis*, second ed., Textbooks in Mathematics, CRC Press, Boca Raton, FL, 2016. (Cited on pages 31, 34, 36, 41, 151, 159, 165, 179, and 180.)
- [46] P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica **1** (1981), no. 4, 357–368. (Cited on page 78.)
- [47] H. Furstenberg, Y. Katznelson, and B. Weiss, *Ergodic theory and configurations in sets of positive density*, Mathematics of Ramsey Theory, Algorithms Combin., vol. 5, Springer, Berlin, 1990, pp. 184–198. (Cited on page 149.)
- [48] K. Gatermann and P. A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Algebra **192** (2004), no. 1-3, 95–128. (Cited on page 85.)
- [49] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization*, second ed., Algorithms and Combinatorics, vol. 2, Springer-Verlag, Berlin, 1993. (Cited on page 6.)

- [50] N. Gvozdenović, M. Laurent, and F. Vallentin, *Block-diagonal semidefinite programming hierarchies for 0/1 programming*, Oper. Res. Lett. **37** (2009), no. 1, 27–31. (Cited on pages 8 and 96.)
- [51] D. Handel, *Some homotopy properties of spaces of finite subsets of topological spaces*, Houston J. Math. **26** (2000), no. 4, 747–764. (Cited on pages 66 and 67.)
- [52] G. Kalai, *Some old and new problems in combinatorial geometry I: around Borsuk’s problem*, Surveys in combinatorics 2015, London Math. Soc. Lecture Note Ser., vol. 424, Cambridge Univ. Press, Cambridge, 2015, pp. 147–174. (Cited on pages 2 and 78.)
- [53] R. M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations (Proceedings of a Symposium on the Complexity of Computer Computations, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, 1972) (R. E. Miller and J. W. Thatcher, eds.), The IBM Research Symposia Series, Plenum, New York-London, 1972, pp. 85–103. (Cited on page 1.)
- [54] E. de Klerk and D. V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM J. Optim. **12** (2002), no. 4, 875–892. (Cited on pages 1, 9, and 11.)
- [55] ———, *A linear programming reformulation of the standard quadratic optimization problem*, J. Global Optim. **37** (2007), no. 1, 75–84. (Cited on pages 1 and 9.)
- [56] O. Kuryatnikova, *The many faces of positivity to approximate structured optimization problems*, PhD thesis, Tilburg University, 2019. (Cited on pages 3, 20, 23, 82, 96, 102, 104, and 170.)
- [57] O. Kuryatnikova and J. C. Vera Lizcano, *Approximating the cone of copositive kernels to estimate the stability number of infinite graphs*, Electronic Notes in Discrete Mathematics **62** (2017), 303–308, LAGOS’17 – IX Latin and American Algorithms, Graphs and Optimization. (Cited on pages 3, 20, 23, 96, 103, and 170.)
- [58] D. de Laat, *Moment methods in energy minimization: new bounds for Riesz minimal energy problems*, Trans. Amer. Math. Soc. **373** (2020), no. 2, 1407–1453. (Cited on page 120.)
- [59] D. de Laat and N. Leijenhorst, *Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput. **16** (2024), no. 3, 503–534. (Cited on pages 85, 88, and 121.)
- [60] D. de Laat, F. C. Machado, F. M. de Oliveira Filho, and F. Vallentin, *k -point semidefinite programming bounds for equiangular lines*, Math. Program. **194** (2022), no. 1–2, Ser. A, 533–567. (Cited on pages 8, 9, 96, 100, and 112.)
- [61] D. de Laat, F. C. Machado, and W. de Munck Keizer, *The Lasserre hierarchy for equiangular lines with a fixed angle*, preprint (2023). arXiv:2211.16471 (Cited on pages 8, 96, and 112.)
- [62] D. de Laat, F. M. de Oliveira Filho, and F. Vallentin, *Upper bounds for packings of spheres of several radii*, Forum Math. Sigma **2** (2014), Paper No. e23, 42. (Cited on pages 88 and 158.)
- [63] D. de Laat and F. Vallentin, *A semidefinite programming hierarchy for packing problems in discrete geometry*, Math. Program. **151** (2015), no. 2, Ser. B, 529–553. (Cited on pages 96, 98, 104, and 112.)
- [64] D. G. Larman and C. A. Rogers, *The realization of distances within sets in Euclidean space*, Mathematika **19** (1972), 1–24. (Cited on page 136.)

- [65] J. B. Lasserre, *An explicit exact SDP relaxation for nonlinear 0-1 programs*, Integer programming and combinatorial optimization (Utrecht, 2001), Lecture Notes in Comput. Sci., vol. 2081, Springer, Berlin, 2001, pp. 293–303. (Cited on page 7.)
- [66] ———, *An explicit equivalent positive semidefinite program for nonlinear 0-1 programs*, SIAM J. Optim. **12** (2002), no. 3, 756–769. (Cited on page 7.)
- [67] M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming*, Math. Oper. Res. **28** (2003), no. 3, 470–496. (Cited on pages 7, 8, and 96.)
- [68] M. Laurent and L. F. Vargas, *Exactness of Parrilo’s conic approximations for copositive matrices and associated low order bounds for the stability number of a graph*, Math. Oper. Res. **48** (2023), no. 2, 1017–1043. (Cited on page 28.)
- [69] V. I. Levenšteĭn, *Boundaries for packings in n -dimensional Euclidean space*, Dokl. Akad. Nauk SSSR **245** (1979), no. 6, 1299–1303. (Cited on page 96.)
- [70] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory **25** (1979), no. 1, 1–7. (Cited on page 6.)
- [71] F. C. Machado and F. M. de Oliveira Filho, *Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry*, Exp. Math. **27** (2018), no. 3, 362–369. (Cited on page 85.)
- [72] M. Matolcsi, I. Z. Ruzsa, D. Varga, and P. Zsámboki, *The fractional chromatic number of the plane is at least 4*, preprint (2025). arXiv:2311.10069 (Cited on page 44.)
- [73] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey, Jr., *The Lovász bound and some generalizations*, J. Combin. Inform. System Sci. **3** (1978), no. 3, 134–152. (Cited on page 96.)
- [74] A. Montina, *Communication cost of classically simulating a quantum channel with subsequent rank-1 projective measurement*, Phys. Rev. A **84** (2011), 060303. (Cited on page 78.)
- [75] W. O. J. Moser, *Problems, problems, problems*, vol. 31, 1991, First Canadian Conference on Computational Geometry (Montreal, PQ, 1989), pp. 201–225. (Cited on page 147.)
- [76] T. S. Motzkin and E. G. Straus, *Maxima for graphs and a new proof of a theorem of Turán*, Canadian J. Math. **17** (1965), 533–540. (Cited on page 1.)
- [77] W. de Muinck Keizer, *On zonal stiefel harmonics with applications in discrete geometry*, PhD thesis, Delft University of Technology, 2025. (Cited on pages 79 and 117.)
- [78] O. R. Musin, *Multivariate positive definite functions on spheres*, Discrete geometry and algebraic combinatorics, Contemp. Math., vol. 625, Amer. Math. Soc., Providence, RI, 2014, pp. 177–190. (Cited on pages 8 and 79.)
- [79] A. M. Odlyzko and N. J. A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Combin. Theory Ser. A **26** (1979), no. 2, 210–214. (Cited on page 96.)
- [80] F. M. de Oliveira Filho, *New bounds for geometric packing and coloring via harmonic analysis and optimization*, PhD thesis, University of Amsterdam, 2009. (Cited on pages 2 and 77.)
- [81] ———, *Data files and code related to the upper bound for Witsenhausen’s parameter*, <https://doi.org/10.7910/DVN/TWI1SR>, Harvard Dataverse, 2023, V1. (Cited on pages 88 and 90.)

- [82] F. M. de Oliveira Filho and F. Vallentin, *Fourier analysis, linear programming, and densities of distance avoiding sets in $\mathbb{R}^k$* , J. Eur. Math. Soc. (JEMS) **12** (2010), no. 6, 1417–1428. (Cited on pages 3 and 147.)
- [83] ———, *A quantitative version of Steinhaus’ theorem for compact, connected, rank-one symmetric spaces*, Geom. Dedicata **167** (2013), 295–307. (Cited on page 74.)
- [84] P. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, PhD thesis, California Institute of Technology, 2000. (Cited on pages 1 and 9.)
- [85] J. Peña, J. Vera, and L. F. Zuluaga, *Computing the stability number of a graph via linear and semidefinite programming*, SIAM J. Optim. **18** (2007), no. 1, 87–105. (Cited on pages 1, 9, and 82.)
- [86] J.-P. Pier, *Amenable locally compact groups*, Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1984. (Cited on pages 40 and 43.)
- [87] A. Polyanskii, *On almost-equidistant sets II*, Electron. J. Combin. **26** (2019), no. 2, Paper No. 2.14, 9. (Cited on pages 111, 114, and 137.)
- [88] P. Pudlák, *Cycles of nonzero elements in low rank matrices*, Combinatorica **22** (2002), no. 2, 321–334, Special issue: Paul Erdős and his mathematics. (Cited on page 111.)
- [89] G. Pólya, *Über positive Darstellung von Polynomen*, George Pólya: Collected Papers, Mathematicians of Our Time, vol. II, The MIT Press, Cambridge, 1974, pp. 309–313. (Cited on pages 10 and 11.)
- [90] H. Reiter and J. D. Stegeman, *Classical harmonic analysis and locally compact groups*, second ed., London Mathematical Society Monographs. New Series, vol. 22, The Clarendon Press, Oxford University Press, New York, 2000. (Cited on page 180.)
- [91] M. Rosenfeld, *Almost orthogonal lines in E^d* , Applied geometry and discrete mathematics, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 4, Amer. Math. Soc., Providence, RI, 1991, pp. 489–492. (Cited on pages 111, 113, 121, and 134.)
- [92] I. J. Schoenberg, *Metric spaces and completely monotone functions*, Ann. of Math. (2) **39** (1938), no. 4, 811–841. (Cited on pages 151 and 152.)
- [93] ———, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108. (Cited on pages 73, 79, and 118.)
- [94] A. Schrijver, *A comparison of the Delsarte and Lovász bounds*, IEEE Trans. Inform. Theory **25** (1979), no. 4, 425–429. (Cited on page 96.)
- [95] ———, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory **51** (2005), no. 8, 2859–2866. (Cited on pages 8 and 96.)
- [96] B. Simon, *Trace ideals and their applications*, second ed., Mathematical Surveys and Monographs, vol. 120, American Mathematical Society, Providence, RI, 2005. (Cited on pages 17, 38, and 179.)
- [97] ———, *Convexity*, Cambridge Tracts in Mathematics, vol. 187, Cambridge University Press, Cambridge, 2011, An analytic viewpoint. (Cited on pages 16, 58, 59, 63, 175, and 176.)
- [98] G. Szegő, *Orthogonal polynomials*, fourth ed., American Mathematical Society Colloquium Publications, vol. Vol. XXIII, American Mathematical Society, Providence, RI, 1975. (Cited on pages 74, 90, and 91.)
- [99] I. G. Todorov and L. Turowska, *Multipliers of multidimensional Fourier algebras*, Oper. Matrices **4** (2010), no. 4, 459–484. (Cited on page 172.)

- [100] M. S. Viazovska, *The sphere packing problem in dimension 8*, Ann. of Math. (2) **185** (2017), no. 3, 991–1015. (Cited on pages 4, 155, and 156.)
- [101] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, 1922. (Cited on page 152.)
- [102] H. S. Witsenhausen, *Spherical sets without orthogonal point pairs*, Amer. Math. Monthly **81** (1974), 1101–1102. (Cited on pages 2, 77, 78, and 89.)

Index

- $*$, 33
- $\star$, 33
- $\otimes$, 5, 16
- $\triangle$, 53
- $\preceq$, 20
- $\langle \cdot, \cdot \rangle$, 5, 16, 17, 175
- $\| \cdot \|_{B^1}$, 179
- $\| \cdot \|_{\text{op}}$, 179
- $\| \cdot \|_p$, 177
- $\| \cdot \|_p$, 177
- $\| \cdot \|_\infty$, 177
- A^* , 16
- $\alpha(H)$, 5
- α_n , 77
- $\alpha(n, t)$, 111
- A^+ , 52
- $\mathcal{A}_{\text{fin}}$, 29
- $\alpha(H)$, 49
- algint, 11
- $\text{Aut}(H)$, 5
- $\text{Av}_{\mathfrak{S}_{k+r}}$, 10
- Av_Γ , 34
 - for compact group, 35
 - for trace class, 36
- $\mathcal{B}$, 178
- $\mathcal{B}^1(L^2(V))$, 179
- $\text{blockM}_r(H)$
 - finite graph, 9
 - measurable hypergraph, 67
 - packing graph, 100
- $B_r(p)$, 156
- $\mathcal{B}(X)$, 179
- $\mathcal{B}(X, Y)$, 179
- $C(V)$, 17
- $C_0(V)$, 17
- $C_c(V)$, 17
- $C_c^\infty(\mathbb{R}^n)$, 155
- $\overline{\text{conv}}$, 16
- $\overline{\text{cone}}$, 16
- cone, 16
- conv, 16
- $\text{COP}(\Gamma)_{\text{inv}}$, 40
- $\text{COP}(V)$
 - in finite dimension, 7
- $\text{COP}(V, k)$, 23
 - in finite dimensions, 10
- $\text{COP}(V, k)_c$, 101
- $\text{CP}(\Gamma)_{\text{inv}}$, 40
- $\text{CP}(V)$, 71
 - in finite dimension, 6
- $\text{CP}(V, k)$, 23
 - in finite dimensions, 10
- $\text{CP}(V, k)_m$, 101
- $\check{C}_r(\Gamma)_{\text{inv}}$, 40
- $C_r(V)$, 71
- $C_r(V, k)$, 23
- $\check{C}_r(V, k)$, 29
- $C_r(V, k)^*$, 23
- $\check{C}_r(V, k)^*$, 29
- C_T , 157
- $D(A)$, 53
- $\text{dens}_v(A)$, 53
- d_L , 148
- dx , 147
- $E(H)$, 115
- $\mathcal{E}_{\text{small}}$, 60

- $\mathcal{E}_{\text{big}}$, 56
- E_n , 20
- $E_{\mathcal{P}}$, 20
- $G(D)$, 148
- $G_L(D)$, 149
- $\overline{H}$, 66
- $\mathcal{I}_{=r}$, 66
- $\mathcal{I}_r$, 66
- $I^S(r)$, 24
- $\mathcal{K}_{=r}$, 66
- $\mathcal{K}_r$, 66
- $L^2_{\text{sym}}(V)$, 16
- $L^2_{\text{sym}}(V, k)$, 16
- $L^p(V)$, 177
- $\mathcal{L}(X, Y)$, 178
- M , 60
- $m_D(\mathbb{R}^n)$, 148
- $M_r(H)$
 - finite graph, 8
 - measurable hypergraph, 67
 - packing graph, 100
- $MS(k)$, 126
- $M_{\text{sym}}(V, k)$, 97
- $O(n)$, 79
- Ω_n , 151
- $\text{Orb}(v)$, 32
- $P_{1, \dots, k}$, 21
- $\text{PSD}(V)$, 17
- $q_*\mu$, 178
- $\mathcal{Q}_r(V)$, 82
- $\mathcal{Q}_r(V, k)$, 82
- Res_A , 29
- σ -compact, 31
- $S(k, l)$, 126
- $\sigma(\mathcal{P})$, 16
- $S(\mathbb{R}^n)$, 156
- S_c , 60
- $\mathfrak{S}_k$, 4
- S^{n-1} , 31
- $\text{Stab}(v)$, 32
- $\text{supp } \mu$, 17
- $\text{Sym}(S)$, 4
- $\text{Sym}(S)_{\geq 0}$, 5
- $\text{Sym}(S, k)$, 4
- $\mathcal{T}(\text{CP}(V, k))$, 56
- $\tau(\text{CP}(V, k))$, 60
- $\vartheta(G)$, 6
 - $\vartheta(G, \mathcal{C})$
 - Euclidean space, 150
 - finite graph, 6
 - packing graph, 103
 - $\vartheta_{\text{SP}}^X(\mathcal{C})$, 156
 - $\vartheta^*(G, \mathcal{C})$, 104
 - $\vartheta_{\text{big}}(H, \mathcal{C})$, 51
 - $\vartheta_{\text{small}}(H, \mathcal{C})$, 52
 - $\mathbb{T}_L$, 148
 - $\mathcal{T}_r$, 23
 - Tr , 17
 - $\bar{\delta}(S)$, 147
 - $V_{1/2}$, 156
 - $V(H)$, 115
 - W_k , 132
 - X' , 16
 - X^* , 16
- action, 32
- adjoint
 - continuous adjoint, 175
 - linear adjoint, 175
- algebraic interior, 11
- almost-equiangular set, 111
 - almost-orthogonal set, 111
 - maximum obtuse, 113, 134
- obtuse, 111
- almost everywhere, 177
- anti-triangle-free graph, 113, 124
 - minimal, 142
 - (n, t) -optimal, 142
- approximate identity, 33
- automorphism group of a hypergraph, 5
- averaging operator, 34
- averaging sequence, 40
- basic open set of $\text{Sub}_k(V)$, 66
- big tip of $\text{CP}(V, k)$, 56
- binary-codes problem, 95
- block moment hierarchy
 - finite graph, 9
 - measurable hypergraph, 67
 - packing graph, 100
- BQP(S^{n-1})-inequality, 86
- circumsphere, 115
- clique, 66
- complement hypergraph, 66

- completely positive cone
 - in finite dimension, 6
 - in finite dimensions, 10
 - in $L^2_{\text{sym}}(V, k)$, 23
 - in $M_{\text{sym}}(V, k)$, 101
- completely positive k -tensor, 23
- conditional expectations, 21
- conic hull, 16
- convergence
 - strong convergence, 179
 - weak convergence, 179
- converges
 - strongly, 179
 - weakly, 179
- converging hierarchy, 8
- convex hull, 16
- convolution, 33
- copositive cone
 - in $C_{\text{sym}}(V, k)$, 101
 - in finite dimension, 7
 - in finite dimensions, 10
 - in $L^2_{\text{sym}}(V, k)$, 23
- copositive k -tensor, 23
- correlation
 - auto-correlation, 33
 - cross-correlation, 33
- D -thick set, 53
- density, 53
- density point, 53
- density system, 53
- diagonal action, 32
- distance-avoiding set, 78
- distance graph, 78
 - distance- t graph, 113, 115
- double-cap conjecture, 78
- double-regular simplex, 115
- dual
 - algebraic dual, 16
 - continuous dual, 16
- dual cone, 176
- dual pair, 175
- duality, 175
- essential supremum, 177
- essentially bounded, 177
- exact bound, 6
- extended $(n - 1)$ -rhombus, 132
- finite-rank approximation, 20
- formulation
 - completely positive, 7
 - copositive formulation, 7
- full support, 17
- fundamental domain, 148
- Γ -equivariant function, 32
- Γ -invariant function, 32
- Γ -space, 32
- great circle, 115
- great sphere, 115
- group
 - amenable group, 40
 - locally compact group, 179
 - topological group, 179
 - unimodular group, 180
- Hamming distance, 95
- Hilbert-Schmidt operator, 179
- homogeneous, 32
- homogeneous setting, 47
- hypergraph
 - k -uniform homogeneous hypergraph, 49
 - k -uniform hypergraph, 5
 - k -uniform locally independent hypergraph, 50
 - k -uniform measurable hypergraph, 49
 - k -uniform measurable locally independent hypergraph, 50
 - uniform homogeneous hypergraph, 49
 - uniform measurable hypergraph, 49
- independence number, 5
 - measurable independence number, 49
- independent set, 5
- induced subgraph, 115
- inequality
 - Hölder's inequality, 177
 - Minkowski's inequality, 178
 - Young's inequality, 33
- k -point bound, 8
- k -tensor, 16
 - symmetric k -tensor, 16
- k -tensor, 4
- kernel, 179

- Lasserre hierarchy, 8
- locally convex space, 175
- Lovász theta number, 6
- martingale, 20, 21
 - convergence, 20
- martingale maximal
 - function, 60
 - operator, 60
 - theorem, 60
- maximal center density, 157
- measure
 - Borel measure, 178
 - Haar measure, 179
 - pushforward measure, 178
 - quotient measure, 180
 - Radon measure, 178
 - symmetric measure, 97
- metric
 - density metric, 61
 - right-invariant metric, 61
- moment hierarchy
 - finite graph, 8
 - locally independent hypergraph, 67
 - packing graph, 100
- μ -equivalent, 177
- $(n - 1)$ -sphere, 115
- (n, t) -realizable, 115
- norm
 - L^p norm, 177
 - trace norm, 179
- operator norm, 179
- orbit, 32
- outer approximation, 11
- p -integrable, 177
- packing graph, 95
 - compact packing graph, 95
- periodic function, 149
 - with period L , 149
- periodic set, 149
 - with period L , 149
- Pólya-type cone
 - in finite dimensions, 10
 - on a finite measure space, 23
- Pólya's theorem, 11
- positive semidefinite
 - kernel, 17
 - measure, 97
- positive type, 34
- problem
 - 1-avoiding set, 3
 - almost-equiangular set, 111
 - equiangular-lines, 112
 - kissing number, 2
 - sphere packing, 3, 155
 - spherical codes, 95
 - Witsenhausen's, 2, 77
- projective approximation lemma, 18
- pushforward, 178
- quadrangular graph, 137
- quotient measure, 32
- radial function, 151
- realizable, 115
- refinement of partition, 20
- Reynolds operator, 34
- rhombus, 125
 - base of rhombus, 125
 - k -rhombus, 125
- Schwartz space, 156
- r -set, 5
- sharp bound, 6
- simplex
 - n -simplex, 115
 - regular n -simplex, 115
- slice positive tensor, 50
- small tip of $\text{CP}(V, k)$, 60
- sphere packing, 155
- spherical code, 95
- spindle, 126
 - (k, l) -spindle, 126
 - k -Moser spindle, 126
 - Moser spindle, 126
- split k -cycle, 132
- stabilizer, 32
- support of measure, 17
- support points of
 - $\text{BQP}(S^{n-1})$ -inequality, 86
- theta-prime number, 8
- thick set, 53
- thick setting, 47
- topology

- dual topology, 175
- $\sigma(X, Y)$ -topology, 175
- standard topology, 66
- strong operator topology, 179
- weak operator topology, 178
- weak topology, 175
- weak* topology, 175

trace, 17

trace class, 179

unique optimal construction, 142

upper density, 147

Dankwoord / Acknowledgements

De eerste regels van dit dankwoord gaan naar mijn ouders, uiteraard. Voor de liefde, het geduld en dat jullie er altijd zijn geweest. Ze gaan ook naar Bronke, ook voor je liefde en voor het luisteren, en omdat je ook op jouw beurt geduld hebt moeten hebben, zeker het afgelopen jaar. Ik ga snel weer mee om met de paarden te wandelen (zonder aan wiskunde te denken).

Roos, dankjewel voor het minutieuze werk dat je hebt gestoken in het tekenen van de mooiste kافت van een proefschrift ooit. (Merk op, lezer, dat Hallard de Schildpad niet alleen zeer gedetailleerd is, maar bovendien Crofts schildpadschild-constructie accuraat verbeelt [33]).

Next, a thank you to those that are mathematically closest to me. Fernando, not only for your expert guidance, but also for the (genuine!) fun while editing the same section of text line-by-line multiple times until it was up to your standard, which I hope has become my standard as well. In that spirit I should also thank Jackie Daytona (regular human bartender) for telling it straight when we are talking The Bullshit. Willem and Nando, besides the many insightful discussions we have had, I am lucky to have had such good friends as my closest colleagues. Willem, having COVID together in Norway is in my top three favorite memories of the past four years—a magical time. David, thanks for the many discussions, mathematical and otherwise.

The DMO group is filled with wonderful people, all of whom made my life in Delft enjoyable. In fear of missing names, the following list is purposefully incomplete and in no particular order, but everyone in the group has been a pleasure to be around. My thanks to Naqi and Qiaochu, for the fond memories Casper has of you; to Yuki, for the **gains**; to Niels, for maximizing shareholder value three days per day; to Nicolaas, for making the entire group believe we do not know basic Dutch—*schabouwelijk*; to Esther, our family-style communal bread lunches are something to remember; and to Cindy, Lara, Ananth, and everyone else, for the lunches and coffee breaks that delayed submission of this thesis by several weeks.

Finally, a thank you to the many amazing people of the optimization community that I have been lucky to meet. In particular, Christine and Philippe were kind enough to host me during research visits to Bordeaux and Toulouse, respectively. In all, the community is a fun, friendly, and welcoming bunch, enough so to convince me to stay.

I am not the only person who worked hard on this thesis. First and foremost, Dorothée, I would not have been able to work my way through the bureaucracy that is TU Delft without you. Joffrey, thank you for the extensive support with everything computer related, and maintaining the computational cluster at TU Delft that most of the optimization problems were solved on. Nando and David, thanks for all the help with `ClusteredLowRankSolver.jl`. Fernando, Philippe, and Dion, thank you for proofreading this thesis—not an insignificant feat.

List of publications

B. Bekker, O. Kuryatnikova, F. M. de Oliveira Filho, and J. C. Vera Lizcano, *Optimization hierarchies for distance-avoiding sets in compact spaces*, Trans. Amer. Math. Soc. **379** (2026), no. 1, 33–70

C. Bachoc, B. Bekker, P. Moustrou, and F. M. de Oliveira Filho, *Obtuse almost-equiaangular sets*, preprint (2025). arXiv:2504.11086

B. Bekker and F. M. de Oliveira Filho, *On the convergence of the k -point bound for topological packing graphs*, preprint (2023). arXiv:2306.02725

Curriculum vitae

ABRAHAM JOHANNES FRANCISCUS BEKKER

Sep. 30, 1993	Born in Tilburg, The Netherlands
2014	Graduated VWO Atheneum Staatsexamen
2014–2018	Double BSc in Physics & Mathematics Radboud University Nijmegen
2018–2021	MSc in Mathematics Radboud University Nijmegen
2022–2026	PhD in Mathematics Delft University of Technology

