On the computation of three and four-point bounds in discrete geometry and analytic number theory

Leijenhorst, N.M.

DOI
10.4233/uuid:91af805a-376c-4ef8-aec5-e6ce08ae20a7

Publication date
2025

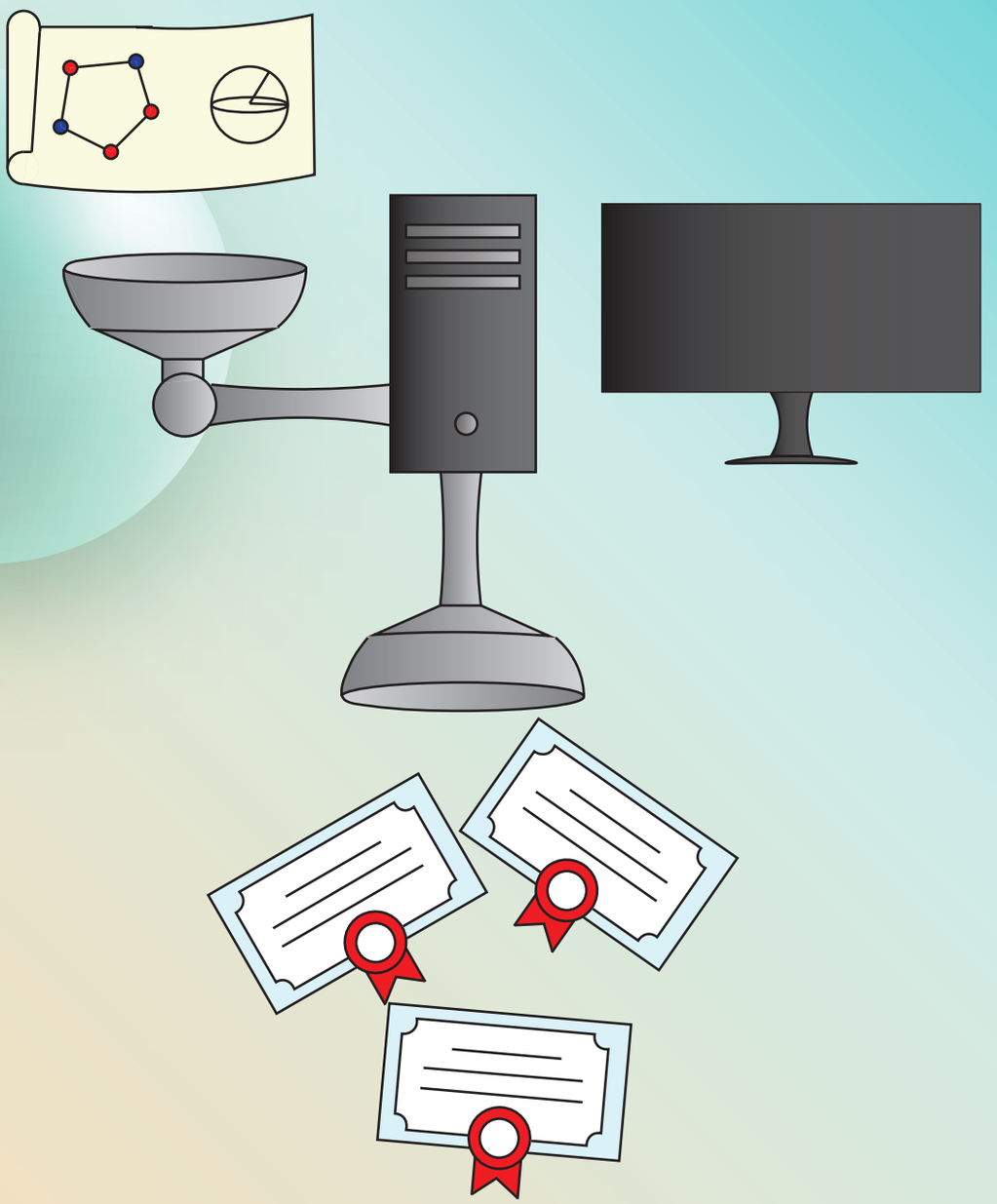
Document Version
Final published version

Citation (APA)
Leijenhorst, N. M. (2025). *On the computation of three and four-point bounds in discrete geometry and analytic number theory*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:91af805a-376c-4ef8-aec5-e6ce08ae20a7

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

On the computation of three and four-point bounds in discrete geometry and analytic number theory

Nando Leijenhorst



On the computation of three and four-point bounds in discrete geometry and analytic number theory Nando Leijenhorst

On the computation of three and four-point bounds in discrete geometry and analytic number theory

On the computation of three and four-point bounds in discrete geometry and analytic number theory

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op donderdag 16 oktober 2025 om 12:30 uur.

door

Nando Matthias Leijenhorst

Master of Science in Applied Mathematics,
Delft University of Technology,
geboren te Delft, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. D.C. Gijswijt	Technische Universiteit Delft, promotor
Dr. D. de Laat	Technische Universiteit Delft, copromotor
Onafhankelijke leden:	
Prof. dr. ir. M.C. Veraar	Technische Universiteit Delft
Prof. dr. M. Laurent	CWI/Tilburg University
Prof. dr. E. de Klerk	Tilburg University
Dr. P. Moustrou	Université Toulouse Jean Jaurès
Prof. dr. ir. C. Vuik	Technische Universiteit Delft, reservelid

Het onderzoek beschreven in dit proefschrift is gefinancierd door de Technische Universiteit Delft.

TU Delft Delft University of Technology

Keywords: Discrete geometry, Lasserre hierarchy, spherical codes, kissing number, energy minimization, covering, polarization, analytic number theory, n-level correlations

Printed by: Ipskamp printing

ISBN 978-94-6473-917-6

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

Contents

Summary

In this thesis, we consider extremal problems in discrete geometry, and in particular, the Lasserre hierarchies for such problems. We give a unifying framework that encompasses the known hierarchies, and lay the foundation to use the second level of such hierarchies for problems on the sphere in practice. We perform explicit harmonic analysis and use polynomial optimization techniques to reduce the problems to a finite-dimensional semidefinite program. We introduce a specialized semidefinite programming solver that uses the structure of the problems, allowing us to use polynomials of significantly higher degree than previously possible, and a much faster rounding procedure to obtain exact optimal solutions to the semidefinite programs. We use this to prove that the D_4 root system is the unique optimal solution to the kissing number problem in dimension 4, and is an optimal spherical code. We also prove there are exactly two optimal spherical codes with 12 points in dimension 4. Furthermore, we show that the spectral embeddings of all known triangle-free strongly regular graphs are optimal spherical codes, as well as certain Kerdock spherical codes. We give numerical evidence that the second level of the Lasserre hierarchy for minimizing harmonic energy is sharp for several infinite families of configurations. We also investigate the strength of the hierarchy for the polarization problem. Finally, we consider triple and quadruple correlation bounds in analytic number theory, which gives new bounds on the fraction of double and triple zeros of the Riemann ζ-function and related functions.

Samenvatting

In dit proefschrift bekijken we extreme problemen in discrete meetkunde, en in het bijzonder de Lasserre-hiërarchieën voor dergelijke problemen. We geven een verenigend raamwerk dat de bekende hiërarchieën omvat, en leggen de basis om het tweede niveau van dergelijke hiërarchieën te berekenen voor problemen op de bol. We voeren de harmonische analyse expliciet uit, en gebruiken polynomiale optimalisatietechnieken om de problemen te reduceren tot een eindig-dimensionaal semidefiniet programma. We introduceren een gespecialiseerd algoritme om deze semidefiniete programma's op te lossen, waarmee we polynomen met significant hogere graad kunnen gebruiken dan voorheen, en geven een snellere afrondingsprocedure om exacte optimale oplossingen te vinden voor de semidefinitieve programma's. We gebruiken dit om te bewijzen dat het D_4-wortelsysteem de unieke optimale configuratie is voor het kusgetal in dimensie 4, en een optimale sferische code is. We bewijzen ook dat er precies twee optimale sferische codes zijn met 12 punten in dimensie 4. Verder laten we zien dat de spectrale inbeddingen van alle bekende driehoeksvrije sterk reguliere grafen optimale sferische codes zijn, evenals bepaalde Kerdock-sferische codes. We leveren numeriek bewijs dat het tweede niveau van de Lasserre-hiërarchie voor het minimaliseren van harmonische energie scherp is voor verschillende oneindige families van configuraties. We onderzoeken ook de sterkte van de hiërarchie voor het polarisatieprobleem. Tot slot bekijken we drievoudige en viervoudige correlatiegrenzen in de analytische getallentheorie, wat nieuwe grenzen geeft aan de fractie van dubbele en driedubbele nulpunten van de Riemann ζ-functie en gerelateerde functies.

CHAPTER 1

Introduction

In discrete geometry, we consider configurations of points under certain constraints. In this thesis we are interested in *extremal* configurations: configurations that additionally maximize or minimize a certain property of the configuration. A famous example is the sphere packing problem: What is the densest packing of non-overlapping spheres in n-dimensional Euclidean space? In 2017, Maryna Viazovska [**140**] solved this problem in dimension 8 , which quickly led to the solution of the problem in dimension 24 [**35**].

A key ingredient of the proof of Viazovska was the linear programming bound of Cohn and Elkies [**33**]. A feasible configuration to the problem gives a lower bound on the density, and the linear programming bound gives an upper bound. The gap between the bounds gives some information about how good the constructions and upper bounds are. In most other dimensions, there is a gap between the best known lower and upper bounds, but due to the exceptional structure of the optimal sphere packings in dimension 8 and 24, the linear programming bound gives exactly the correct density in those cases. To prove optimality, Viazovska gave an *exact optimal* solution to the linear programming bound.

Linear programming bounds have also been developed for the spherical cap packing problem (see Example 1.1) by Delsarte, Goethals and Seidels in 1977 and for the energy minimization problem (see Example 1.2) in 1993 by Yudin. The linear programming bounds are 2-point bounds: they only consider conditions arising from the distribution of distances between pairs of points. For two-point bounds, we can generally compute a very precise approximation of the bound.

EXAMPLE 1.1 (Spherical cap packing). *Given an integer n and angle θ, what is the largest set $C \subseteq S^{n-1}$ such that $\langle x, y\rangle \leq \cos\theta$ for all distinct $x, y \in C$? Here $\langle x, y\rangle$ denotes the Euclidean inner product between two vectors in $\mathbb{R}^n$.*

EXAMPLE 1.2 (Energy minimization). *Given integers n and N, and a function $f : [-1, 1) \to \mathbb{R}$, what configuration $C \subseteq S^{n-1}$ minimizes*

$$E_f(C) = \frac{1}{2} \sum_{\substack{x,y \in C \\ x \neq y}} f(\langle x, y\rangle)$$

over all N-point configurations on S^{n-1}?

For several problems in discrete geometry, k-point bounds for $k > 2$ have also been developed. This was initiated in 2008 when Bachoc and Vallentin gave a three-point bound for the kissing number and spherical cap packing problem [**2**], and Cohn and Woo used similar techniques to give a three-point bound for the energy minimization problem on the sphere in 2012 [**42**]. For sphere packing, a three-point bound was only given in 2022, by Cohn, de Laat, and Salmon [**39**]; the

main difference here is problems considered earlier are restricted to the sphere, a compact set, whereas the sphere packing problem optimizes over configurations in Euclidean space. Three-point bounds are more difficult to compute than the two-point bounds, but can generally be approximated moderately well.

In this thesis, we lay the foundation to compute four-point bounds in practice. The first computations of a four-point bound for such problems have only been done in 2019, for an energy minimization problem in dimension 3 [**85**]. However, the method does not extend easily to higher dimensions, and the approximation is not good enough for the problems considered in this thesis.

This thesis consists of 11 chapters, including this introduction and a final chapter discussing future work. It is divided into two parts. In Part I, we consider techniques to set up the Lasserre hierarchy for problems in discrete geometry, a sequence of semidefinite programming bounds; to reduce such an infinite-dimensional problem to a finite dimensional semidefinite program; and to give an exact solution to it. In Part II, we apply these methods to several problems in discrete geometry, and we apply semidefinite programming bounds to a problem in analytic number theory. Chapters 3 and 5–10 are largely based on papers.

Part I: Techniques

In Chapter 2, we consider one of the known techniques to define such k-point bounds for problems in discrete geometry on a compact set: the (generalized) Lasserre hierarchy. The original hierarchies, also called moment/SOS hierarchies, were defined for polynomial optimization problems [**94, 119, 95**]. De Laat and Vallentin generalized the moment side of the hierarchy to packing problems on compact sets [**92**], after which similar ideas have been used to define hierarchies for the energy minimization problem [**85**] and covering problems [**124**], in compact settings. In this chapter, we state these hierarchies in a unifying framework.

For a maximization problem with optimal value α, such a hierarchy gives a sequence of optimization problems

$$\mathrm{las}_1 \geq \mathrm{las}_2 \geq \cdots \geq \alpha,$$

and for the problems in this thesis, we have finite convergence:

$$\mathrm{las}_N = \alpha$$

for some finite N. Higher levels in the hierarchy give better bounds, but are also significantly more difficult to compute. The k-th level is a $2k$-point bound. The first level is generally equivalent to the linear programming bound for the problem, and the three-point bounds can often be seen as relaxations of the second level of the hierarchy.

In Chapter 3, we develop the harmonic analysis required to compute the second level of the Lasserre hierarchy for problems on the sphere, in any dimension. Importantly, the computational complexity does not depend on the dimension. This is in contrast with the method used in [**85**] to compute the second level of the hierarchy for the energy minimization problem on the sphere in dimension 3.

In Chapter 4, we give a short exposition on polynomial optimization in the context of this thesis. This can result in a semidefinite program with additional structure.

In Chapter 5 we introduce a semidefinite programming solver which exploits this structure to speed up the computations, and uses high-precision arithmetic. This

solver is fundamental for our results. Previously, the largest semidefinite programs in discrete geometry were solved by Machado and Oliveira using the solver `SDPA-GMP` [**103**], and our solver is a factor 28 faster for those problems [**98**].

Summarizing, Chapters 3–5 give a sequence of optimization problems

$$\text{las}_2 \leq \text{las}_{2,d} \leq \text{las}_{2,d,\delta}, \tag{1.1}$$

where the inequalities correspond to a maximization problem in discrete geometry, and are reversed for a minimization problem. The first inequality we obtain through a truncation of the Fourier series of the functions we optimize over in las_2, and the second inequality by reformulating a polynomial optimization problem with positive semidefinite variables to a semidefinite program of finite size using sum-of-squares polynomials.

In Chapter 6, we give a method to obtain an exact optimal solution to such a semidefinite program, from a numerically optimal solution. This method is much faster than the previous methods considered in [**51, 109**]. Such an exact solution is fundamental for proving optimality of configurations. The solution to the sphere packing problem in dimensions 8 and 24 was possible because the linear programming bound was *sharp*: it gives exactly the density of the optimal sphere packing. This can happen more generally: even though the proof of finite convergence shows that $\text{las}_N = \alpha$ for a given N, it may happen that $\text{las}_k = \alpha$ for $k \ll N$. Additionally, for compact spaces such as the sphere, equality may hold throughout (1.1) for a certain d and δ, so that an (exact) optimal solution of a semidefinite program of finite size can give an optimal solution to the k-th level of the hierarchy.

Part II: Applications

In Chapter 7, we consider the spherical code problem. This problem asks for a subset $C \subseteq S^{n-1}$ of size N, a *spherical code*, which minimizes

$$\max_{\substack{x,y \in C \\ x \neq y}} \langle x, y \rangle$$

among all spherical codes of size N in dimension n. We use the three-point bound and the second level of the Lasserre hierarchy to prove optimality of certain spherical codes, and give improved kissing number bounds. The following are our main results. First, we prove that, up to isometry, the D_4 root system is the unique optimal spherical code in dimension 4 of size 24, and the unique optimal solution to the kissing number problem in dimension 4. This is closely related to the sphere packing problem in dimension 4, where the D_4 root lattice is conjectured to be optimal, and gives some indication that the second level of a Lasserre hierarchy for the sphere packing problem may be able to prove optimality of the D_4 root lattice. Second, we prove that there are exactly *two* optimal solutions to the spherical code problem in dimension 4 for spherical codes of size 12, up to isometry. In contrast, many problems have either zero, one or infinitely many solutions.

In Chapter 8, we consider the energy minimization problem on the sphere. The problem asks for a spherical code $C \subset S^{n-1}$ of size N which minimizes

$$E_f(C) = \frac{1}{2} \sum_{\substack{x,y \in C \\ x \neq y}} f(\langle x, y \rangle)$$

for a given potential function f. We prove optimality of one spherical code for the wide class of absolutely monotonic potential functions using the three-point bound, and give numerical evidence that the second level of the Lasserre hierarchy is optimal for three families of spherical codes, with $N = n + 2$, $N = 2n - 1$, and $2n + 2$ points in dimension n, for the harmonic energy potential function

$$f(u) = (2 - 2u)^{-(n-2)/2}.$$

This potential function generalizes the Coulomb potential in dimension 3.

In Chapter 9, we consider the problem of polarization with a threshold: Given a potential function $f : [-1, 1) \to \mathbb{R}$, and a threshold E, find a spherical code C of minimum size such that

$$P_f(C, y) = \sum_{x \in C} f(\langle x, y \rangle) \geq E$$

for every $y \in S^{n-1}$. In other words, what is a configuration with the minimum number of light sources, such that the darkest point on the sphere has at least a certain brightness? A particular instance of this is the problem of covering the sphere with spherical caps of a given radius. In this chapter we generalize the hierarchy for covering problems of Riener, Rolfes and Vallentin [**124**] to this problem, and show for example that the allowed minimum distance has an important influence on the bounds. In particular, taking the limit of the minimum distance to 0 gives a trivial bound, in contrast to the energy minimization problem, where the second level of the hierarchy still gives many sharp bounds after taking this limit.

The linear programming bound introduced by Cohn and Elkies [**33**] has some interesting connections to fields outside discrete geometry. For example, there is a precise connection between the Cohn-Elkies bound and the modular bootstrap in physics [**67**]. In analytic number theory, there is a relation between the feasible region of the bound and the feasible region of bounds on the number of simple zeros of so-called L-functions of certain representations of $\mathrm{GL}_m/\mathbb{Q}$, which generalize the well-known Riemann ζ-function.

This connection is established through the pair-correlation approach of Montgomery [**111**], assuming the Riemann Hypothesis that every zero of the Riemann ζ-function is a negative even integer or of the form $\rho_j = 1/2 + i\gamma_j$ with $\gamma_j \in \mathbb{R}$. A slight generalization of the pair-correlation approach of Montgomery minimizes a linear functional over the feasible region of the Cohn-Elkies bound.

In Chapter 10, we consider bounds using the higher-order correlation functions of Hejhal [**69**] and Rudnick and Sarnak [**126, 127**]. Although the pair-correlation function gives bounds on the number of simple zeros, it turns out that using the k-level correlation functions leads to bounds on the number of zeros with multiplicity at most $k - 1$.

In Chapter 11, we give possible directions for future research. This includes for example possible improvements in the solver and rounding procedure, interesting questions related to our work, and some suggestions how these questions can possibly be handled.

Part I

Techniques

CHAPTER 2

Lasserre hierarchies for problems in discrete geometry

Many problems in discrete geometry have an analogous problem on a finite graph. For example, packing problems are analogous to independent set problems, and covering problems to the set cover problem. One method to solve such problems on a finite graph is to reformulate them as binary polynomial optimization problems, and to apply the moment/SOS hierarchy for polynomial optimization to it [**95**]. In [**92**], this was generalized to infinite packing problems in compact spaces, and later the approach was adapted to energy minimization [**85**] and covering [**124**]. These hierarchies generalize the moment side of the moment/SOS hierarchies, and are often called Lasserre hierarchies. The earlier known bounds, such as the Delsarte-Goethals-Seidels linear programming bound [**47**] and the Bachoc-Vallentin three-point bound [**2**] for spherical codes or Yudin's bound for energy minimization [**144**], can in general be understood as (relaxations of) a level of the Lasserre hierarchy for the corresponding problem.

In this chapter, we give an exposition of the generalized Lasserre hierarchies as used in [**92, 85, 124**], by stating these hierarchies in a more general, unifying framework.

2.1. Preliminaries

In this chapter we assume a basic knowledge of graphs and semidefinite programming. This section provides some of the basic definitions required.

Let $G = (V, E)$ be a graph with vertex set V and edges $E \subseteq \mathrm{Sub}(V, 2)$, where $\mathrm{Sub}(V, k)$ denotes the subsets of V of size k. A subset $C \subseteq V$ is *independent* if $\{x, y\} \notin E$ for every $x, y \in C$.

A semidefinite program is an optimization problem of the form

$$\begin{array}{lll} & \text{maximize} & \langle C, X\rangle, \\ & \text{subject to} & \langle A_i, X\rangle = b_i, \quad i = 1, \dots, m, \\ & & X \succeq 0. \end{array} \tag{2.1.1}$$

Here $\langle A, B\rangle = \mathrm{Tr}(A^{\mathsf{T}} B)$ is the trace inner product for two matrices $A, B \in \mathbb{R}^{n\times m}$, and $X \succeq 0$ denotes that the matrix $X \in \mathbb{R}^{n\times n}$ is *positive semidefinite*: It is symmetric, and

$$a^{\mathsf{T}} X a \geq 0 \quad \forall a \in \mathbb{R}^n,$$

or equivalently, all eigenvalues of X are nonnegative. In particular, the trace inner product between two positive semidefinite matrices is nonnegative.

Problem (2.1.1) is a semidefinite program in *standard form*. The dual problem of this reads

$$\begin{aligned} &\text{minimize} && \langle y, b\rangle, \\ &\text{subject to} && \sum_i y_i A_i - C \succeq 0. \end{aligned}$$

Such a problem is said to be in *geometric form*.

Weak duality holds: for every primal feasible solution X and dual feasible solution y, we have $\langle C, X\rangle \leq \langle y, b\rangle$. This can easily be proven with

$$0 \leq \langle \sum_i y_i A_i - C, X\rangle = \sum_i y_i \langle A_i, X\rangle - \langle C, X\rangle = \langle y, b\rangle - \langle C, X\rangle.$$

Duality can be considered much more generally, see for example [**5**] for a general treatment; we give a short summary in Section 2.4, and use it for our problems.

A continuous kernel $K \in C(V \times V)$ is *positive definite* ($K \succeq 0$) if for every finite set $\{x_1, \ldots, x_N\} \subseteq V$ the $N \times N$ matrix with entries $A_{ij} = K(x_i, x_j)$ is a positive semidefinite matrix.

2.2. Binary polynomial optimization with finitely many variables

A polynomial optimization problem is a problem of the form

$$\begin{aligned} &\text{maximize} && g_0(x) \\ &\text{subject to} && g_i(x) \geq 0, \quad i = 1, \ldots, m. \end{aligned}$$

Here g_i is a polynomial in n variables. For most problems in extremal geometry, we restrict the variables x_j to be binary. This can be done by adding the constraints $x_j^2 - x_j = 0$ (for an equality constraint $g(x) = 0$ we add the two constraints $g(x) \geq 0$ and $-g(x) \geq 0$).

The following examples give formulations for the problems on finite graphs analogue to those considered in this thesis.

Example 2.1 (Packing). *The independent set problem asks for the maximum size of a set I such that for every $x, y \in I$, $\{x, y\} \notin E$. This can be modeled as a binary polynomial optimization problem by setting $g_{\{i,j\}}(x) = 1 - (x_i + x_j)$ for every $\{i, j\} \in E$, and $g_0 = \sum_{i \in V} x_i$. This will be the running example throughout the first part of this thesis.*

Example 2.2 (Energy minimization). *Let G be a complete graph (that is, $E = \mathrm{Sub}(V, 2)$ contains all possible edges) and let $f : E \to \mathbb{R}$ be a weight function on the edges. Given $N \leq |V|$, the energy minimization problem asks for the minimum energy*

$$\frac{1}{2} \sum_{\substack{x,y \in S \\ x \neq y}} f(\{x, y\})$$

over all $S \subseteq V$ of size N. Here we can take $g_1 = \sum_i x_i - N$, $g_2 = N - \sum_i x_i$ and $g_0 = -\sum_{i \neq j} f(\{i, j\}) x_i x_j$.

Example 2.3 (Covering). *The dominating set problem asks for the minimum size of a set $S \subseteq V$ such that every vertex in V is either in S or adjacent to a vertex in S. For this problem we can take $g_0 = -\sum_{i \in V} x_i$, and $g_j = x_j + \sum_{i : \{i,j\} \in E} x_i - 1$ for $j \in V$.*

EXAMPLE 2.4 (Polarization with threshold). *Given a complete graph $G = (V, E)$, a weight function $f : V \times V \to \mathbb{R}$ and a set S, the polarization of a point $i \in V$ is given by*

$$\mathcal{P}_f(S, i) = \sum_{j \in S} f(i, j).$$

The polarization problem with threshold E asks for the minimum size of a set S such that $P_f(S, i) \geq E$ for every $i \in V$, or equivalently, that the minimum polarization is at least E. The objective is $g_0 = -\sum_i x_i$, and the constraints are given by

$$g_i(x) = \sum_j f(i, j)x_j - E$$

for all $i \in V$.

A polynomial $g(x)$ can be written as

$$g(x) = \sum_{\alpha \in \mathbb{N}^n} g_\alpha x^\alpha$$

where α is the exponent vector of the monomial $x^\alpha = \prod_i x_i^{\alpha_i}$ and g_α is the coefficient corresponding to this monomial.

Let $\mathcal{P}_k^n \subseteq \mathbb{N}^n$ denote the exponent vectors α of degree $|\alpha| = \sum_i \alpha_i \leq k$. For $y \in \mathbb{R}^{\mathcal{P}_{2k}^n}$, the *moment matrix* of y is defined by the entries

$$M_k(y)_{\alpha,\beta} = y_{\alpha+\beta}$$

with $\alpha, \beta \in \mathcal{P}_k^n$. The *localizing matrix* $M_k^g(y)$ corresponding to a constraint $g(x) \geq 0$ is given by

$$M_k^g(y)_{\alpha,\beta} = \sum_\gamma g_\gamma y_{\alpha+\beta+\gamma}$$

whenever $|\alpha|, |\beta| \leq k - \lceil \deg(g)/2 \rceil$.

Lasserre introduces in [**94**] the following hierarchy of relaxations for polynomial optimization problems:

$$\begin{aligned}
\text{maximize} \quad & \sum_{\alpha \in \mathcal{P}^n_{\deg(g_0)}} (g_0)_\alpha y_\alpha, \\
\text{subject to} \quad & y_0 = 1, \\
& M_k(y) \succeq 0, \\
& M_k^{g_i}(y) \succeq 0 \qquad\qquad i = 1, \ldots, m, \\
& y \in \mathbb{R}^{\mathcal{P}^n_{2k}}_{\geq 0}.
\end{aligned}$$

This is a semidefinite program in geometric form.

Let x be a feasible solution to the original binary optimization problem, and let y be the moment vector corresponding to a Dirac measure δ_x on x. Then we have

$$y_\alpha = \int z^\alpha d\delta_x(z) = x^\alpha,$$

so that

$$g(x) = \sum_\alpha g_\alpha y_\alpha$$

and

$$y_0 = \int d\delta_x(z) = 1,$$
$$M_k(y) = \int b(z)b(z)^\mathsf{T} d\delta_x(z) \succeq 0,$$
$$M_k^{g_i}(y) = \int g_i(z)b(z)b(z)^\mathsf{T} d\delta_x(z) \succeq 0,$$

where b is a vector of monomials. Hence y is feasible, so the optimal value of each level of the hierarchy gives an upper bound on the optimal value of the binary polynomial optimization problem.

The hierarchy converges as $k \to \infty$ [**94**] under some mild conditions (see also Chapter 4 where the dual problem of optimizing over sum-of-squares polynomials is considered), and in certain cases (such as binary polynomial optimization [**95**]), the hierarchy convergences in a finite number of steps.

When $g_i = -g_j$ for some i, j (that is, we enforce the equality constraint $g_i = 0$), the constraints $M_k^{g_i}(y) \succeq 0$ and $M_k^{-g_i}(y) = -M_k^{g_i}(y) \succeq 0$ imply that $M_k^{g_i}(y) = 0$. This enables us to find a more suitable form of the program to generalize to infinite graphs.

In the following example, we show that the binary constraints imply that we can index the moment vector y by sets $Q \subseteq [n]$ instead of exponent vectors $\alpha \in \mathcal{P}_{2k}^n$, where

$$Q = \operatorname{Supp}(\alpha) = \{i : \alpha_i > 0\}.$$

This was first observed in [**96**].

EXAMPLE 2.5 (Binary variables). *For the constraints $x_j^2 - x_j = 0$, $M_k^{g_i} = 0$ reduces to*

$$y_{\alpha+2e_j} = y_{\alpha+e_j}$$

for every exponent vector α with $|\alpha| \leq 2t-2$. Iterating this gives that $y_\alpha = y_\beta$ whenever

$$\operatorname{Supp}(\alpha) = \operatorname{Supp}(\beta).$$

In other words, we may index the moment vector y and the coefficients and monomials of g by sets instead of exponent vectors, so that

$$M_k^g(y)_{J_1,J_2} = \sum_{Q \subseteq [n]} g_Q y_{J_1 \cup J_2 \cup Q}.$$

Note that if some variables are not binary, we still need to include all possible powers for those variables.

EXAMPLE 2.6 (Cardinality constraints). *Consider the constraint $\sum_{j \in V} x_j - N = 0$. As in the previous example, this implies that the corresponding localizing matrix should be 0, which gives the constraints*

$$\sum_{i \in V} y_{J \cup \{i\}} = N y_J$$

*for all $J \subseteq V$ with $|J| \leq 2t-1$. Lemma 3.1 in [**85**] shows that this the equality constraints*

$$\sum_{S \subseteq \operatorname{Sub}(V,=k)} y_S = \binom{N}{k} y_\emptyset$$

for all $0 \leq k \leq 2t$.

2.3. Generalization to infinite graphs

In [**92**], de Laat and Vallentin generalize the Lasserre hierarchy as described above to infinite but compact spaces: compact topological packing graphs. Such a graph has a compact Hausdorff topological space as vertex set V, and the property that every finite clique is contained in an open clique [**92**, Definition 1]. An open clique is an open subset of the vertex set that is also a clique: every two vertices in the set are adjacent.

Instead of the set of subsets $\mathrm{Sub}(V,k)$ they consider the set $\mathcal{I}_k$ of independent subsets of V of size at most k. Since V is compact and every vertex is contained in an open clique, the maximum size of an independent set is finite. In general, this is needed for finite convergence, but if the problem contains a cardinality constraint (e.g., in energy minimization problems; see Example 2.2), we can often take a limit and avoid the property that every finite clique is contained in an open clique without losing finite convergence, because the cardinality constraint will limit the size of a feasible configuration.

We follow the approach in [**92**, Section 2] to define the topology through the product topology on V^k, the quotient topology for the image of V^k under the map

$$q : (v_1, \dots, v_k) \mapsto \{v_1, \dots, v_k\},$$

and the disjoint union topology to add the empty set. This gives the topology on $\mathrm{Sub}(V,k)$ of subsets of V of size at most k, and $\mathcal{I}_k$ inherits this topology. Then $\mathcal{I}_k$ is Hausdorff and compact [**92**, Lemma 1].

When V is a metric space, this topology is the topology induced by the Hausdorff distance

$$d_H(J, J') = \inf\{\varepsilon : J \subseteq J'_\varepsilon \text{ and } J' \subseteq J_\varepsilon\}$$

where $J_\varepsilon = \bigcup_{x \in J} B(x, \varepsilon)$ is the ε-thickening of J. See the discussion in [**124**, Section 3] and [**92**, Section 2].

With this topology, the connected components of $\mathcal{I}_k$ are the sets $\mathcal{I}_{=0}, \dots, \mathcal{I}_{=k}$, where $\mathcal{I}_{=i}$ contains all independent sets of size i, because of the condition that (V, E) is a topological packing graph. De Laat and Vallentin prove this in Lemma 2 of [**92**], which shows that $\mathcal{I}_{=i}$ is both open and closed in $\mathcal{I}_k$. In particular, the indicator function $\chi_{\mathcal{I}_{=i}}$ is continuous, so the constraint $y_0 = 1$ can be directly translated to $\lambda(\mathcal{I}_0) = 1$.

The generalizations of the Lasserre hierarchy for topological packing graphs replace the moment vector $y \in \mathbb{R}^{\mathcal{I}_{2k}}$ by a measure in $\mathcal{M}(\mathcal{I}_{2k})$, the signed Radon measures on $\mathcal{I}_{2k}$. Since measures are not defined on single points but on Borel sets, it is convenient to generalize the adjoint of the moment and localizing matrices instead of the matrices themselves. Equality constraints can be generalized directly when they can be expressed as $\lambda(g) = b$, with $g \in C(\mathcal{I}_{2k})$. This is the case for the cardinality constraints derived in [**85**, Lemma 3.1] (see also Example 2.6), which are currently the only type of equality constraints used in discrete geometry.

Let $g \geq 0$ be some constraint in the binary polynomial optimization problem. Then the corresponding localizing matrix has entries

$$M_k^g(y)_{J_1,J_2} = \sum_{\substack{Q \subseteq V \\ |Q| \leq \deg(g)}} g_Q y_{J_1 \cup J_2 \cup Q}.$$

with $|J_1 \cup J_2| \leq 2k - \deg(g)$. Set $k_g = k - \lceil \deg(g)/2 \rceil$, and denote by $\mathcal{Q}_k^n$ the subsets of $[n]$ of size at most k. Then $M_k^g(y) \in \mathbb{R}^{\mathcal{Q}_{k_g}^n \times \mathcal{Q}_{k_g}^n}$. Let $K \in \mathbb{R}^{\mathcal{Q}_{k_g}^n \times \mathcal{Q}_{k_g}^n}$ be positive semidefinite. Then the inner product between the localizing matrix and this kernel is given by

$$\begin{aligned}\langle M_k^g(y), K\rangle &= \sum_{J_1,J_2 \in \mathcal{Q}_{k_g}^n} K(J_1,J_2) \sum_{\substack{Q \subseteq [n] \\ |Q| \leq \deg(g)}} g_Q y_{J_1 \cup J_2 \cup Q} \\ &= \sum_{S \in \mathcal{Q}_{2k}^n} y_S \sum_{\substack{Q \in \mathcal{Q}_{\deg(g)}^n, J_1, J_2 \in \mathcal{Q}_{k_g}^n \\ J_1 \cup J_2 \cup Q = S}} g_Q K(J_1,J_2) \\ &= \langle y, A_k^g K\rangle,\end{aligned}$$

where we use kernel notation for indexing the matrix K and define the operator $A_k^g : \mathbb{R}^{\mathcal{Q}_k^n \times \mathcal{Q}_k^n} \to \mathbb{R}^{\mathcal{Q}_{2k}^n}$ by

$$A_k^g K(S) = \sum_{\substack{Q \in \mathcal{Q}_{\deg(g)}^n, J_1, J_2 \in \mathcal{Q}_{k_g}^n \\ J_1 \cup J_2 \cup Q = S}} g_Q K(J_1,J_2).$$

Since $\langle M_k^g(y), K\rangle = \langle y, A_k^g K\rangle$, this is the adjoint of M_k^g. Instead of generalizing M_k^g, we generalize A_k^g: We replace $\mathcal{Q}_k^n$ by $\mathcal{I}_k$, and take a proper generalization of g to a function $g \in C(\mathcal{I}_{2k})$; here evaluating the new function g on a set $Q \in \mathcal{I}_{2k}$ is the analogue to taking the coefficient of the polynomial g corresponding to a set $Q \in \mathcal{Q}_k^n$. Similarly to polynomials, we say that the *degree* of g (denoted by $\deg(g)$) is the maximum ℓ such that g is nonzero on $\mathcal{I}_{=\ell}$. This gives an operator $A_k^g : C(\mathcal{I}_{k_g} \times \mathcal{I}_{k_g}) \to C(\mathcal{I}_{2k})$ defined by

$$A_k^g K(S) = \sum_{\substack{Q \in \mathcal{I}_{\deg(g)}, J_1, J_2 \in \mathcal{I}_{k_g} \\ J_1 \cup J_2 \cup Q = S}} g(Q) K(J_1,J_2),$$

where $k_g = k - \lceil \deg(g)/2 \rceil$. Since $\|A_k^g K\| \leq 2^{2k} \|g\|_\infty \|K\|_\infty$, A_k^g is bounded and thus continuous. That implies that there exists an adjoint $(A_k^g)^* : \mathcal{M}(\mathcal{I}_{2k}) \to \mathcal{M}(\mathcal{I}_{k_g} \times \mathcal{I}_{k_g})$. The generalization of the localizing constraint $M_k^g \succeq 0$ then reads

$$(A_k^g)^*(\lambda) \in \mathcal{M}(\mathcal{I}_{k_g} \times \mathcal{I}_{k_g})_{\succeq 0},$$

where the set $\mathcal{M}(V \times V)_{\succeq 0}$ is defined to be all symmetric, signed Radon measures $\nu \in \mathcal{M}(V \times V)$ such that

$$\langle \nu, K\rangle \geq 0$$

for every positive definite kernel $K \in C(V \times V)_{\succeq 0}$. A signed Radon measure $\mu \in \mathcal{M}(V \times V)$ is symmetric if

$$\mu(A \times B) = \mu(B \times A)$$

for all Borel sets A and B of V. Generally, we will write $A_k = A_k^1$ for the operator used to generalize the moment matrix condition. By $\mathcal{M}(V)_{\geq 0}$ we denote the nonnegative Radon measures on V.

Summarizing, the general form of the Lasserre hierarchy for problems on compact topological packing graphs is given by

$$\begin{array}{lll} \text{maximize} & \lambda(h), & \\ \text{subject to} & (A_k^g)^*(\lambda) \in \mathcal{M}(\mathcal{I}_{k_g} \times \mathcal{I}_{k_g})_{\succeq 0}, & g \in G \cup \{1\}, \\ & \lambda(f) = b_f, & f \in F \cup \{\chi_{\mathcal{I}_0}\}, \\ & \lambda \in \mathcal{M}(\mathcal{I}_{2k})_{\geq 0}, & \end{array} \tag{2.3.1}$$

where $G \subseteq C(\mathcal{I}_{2k})$ is a set generalizing the inequality constraints, and $F \subseteq C(\mathcal{I}_{2k})$ is a set generalizing the equality constraints. Note that the right hand side corresponding to $\chi_{\mathcal{I}_0}$ equals 1.

EXAMPLE 2.1 (Packing, continued). *The moment matrix corresponds to $g = 1$. This gives the operator $A_k : C(\mathcal{I}_k \times \mathcal{I}_k) \to C(\mathcal{I}_{2k})$ defined by*

$$A_k K(S) = \sum_{\substack{J_1, J_2 \in \mathcal{I}_k \\ J_1 \cup J_2 = S}} K(J_1, J_2).$$

Since the indicator function of $\mathcal{I}_{=i}$ is continuous for all i we may use $\lambda(\mathcal{I}_{=1})$ as objective function. The k-th level of the Lasserre hierarchy for the independent set problem for topological packing graphs then becomes (cf. [**92**, Definition 2]*)*

$$\begin{array}{lll} \textit{maximize} & \lambda(\mathcal{I}_{=1}) & \\ \textit{subject to} & A_k^*(\lambda) \in \mathcal{M}(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}, & \\ & \lambda(\mathcal{I}_{=0}) = 1, & \\ & \lambda \in \mathcal{M}(\mathcal{I}_{2k})_{\geq 0}. & \end{array} \tag{2.3.2}$$

2.4. Dualization

To obtain valid bounds, there are two difficulties with problem (2.3.1). In the first place, it is simply difficult to optimize over measures. More importantly, it is a maximization problem where the optimal value gives upper bounds on the quantity we consider. This means that we need an *optimal* solution, with optimality certificate, to give a valid bound. To solve both issues, we dualize the problem. This gives a minimization problem over continuous functions, where each *feasible* solution gives a valid bound. Optimization over continuous functions is easier, and we only need to prove feasibility instead of optimality for a valid bound.

We use the framework of Barvinok [**5**, Chapter IV] for duality theory. Let E, F be topological vector spaces, and consider a non-degenerate bilinear form $\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}$. If every linear functional on E can be written as $\langle \cdot, f \rangle$ for some $f \in F$ and every linear functional on F can be written as $\langle e, \cdot \rangle$ for some $e \in E$, $\langle \cdot, \cdot \rangle$ is called a *duality*. Let $K \subseteq E$ be a (convex) cone. Then the dual cone is given by

$$K^* = \{f \in F : \langle e, f \rangle \geq 0 \, \forall e \in K\}.$$

Now consider vector spaces E_1, E_2, F_1, F_2 with dualities $\langle \cdot, \cdot \rangle_1 : E_1 \times F_1 \to \mathbb{R}$ and $\langle \cdot, \cdot \rangle_2 : E_2 \times F_2 \to \mathbb{R}$. Let $A : E_1 \to E_2$ be a linear transformation, and let $A^* : F_2 \to F_1$ be its adjoint. That is,

$$\langle Ae, f \rangle_2 = \langle e, A^* f \rangle_1$$

for all $e \in E_1, f \in F_2$. Let $K_1 \subseteq E_1$ and $K_2 \subseteq E_2$ be convex cones, take $b \in E$ and $c \in F$. Then the problems

$$\begin{aligned} &\text{minimize} && \langle x, c\rangle, \\ &\text{subject to} && Ax \geq_{K_2} b, \\ &&& x \geq_{K_1} 0, \end{aligned}$$

and

$$\begin{aligned} &\text{maximize} && \langle b, l\rangle, \\ &\text{subject to} && A^* l \leq_{K_1^*} c, \\ &&& l \geq_{K_2^*} 0. \end{aligned}$$

are a primal-dual pair: for every feasible $x \in E$ and $l \in F$ we have $\langle b, l\rangle \leq \langle x, c\rangle$ (weak duality, see [**5**, Theorem IV.6.2]).

Recall that in the general formulation of the Lasserre hierarchy there is a localizing constraint for every $g \in G$ and an equality constraint for every $f \in F$. When G and F are finite, this amounts to taking $F_2 = \mathcal{M}(\mathcal{I}_{2t})$ and

$$F_1 = \left(\bigoplus_{g \in G \cup \{1\}} \mathcal{M}(\mathcal{I}_{k_g} \times \mathcal{I}_{k_g}) \right) \oplus \mathbb{R}^{F \cup \{\chi_{\mathcal{I}_0}\}}$$

with dual spaces $E_2 = C(\mathcal{I}_{2t})$ and

$$E_1 = \left(\bigoplus_{g \in G \cup \{1\}} C(\mathcal{I}_{k_g} \times \mathcal{I}_{k_g}) \right) \oplus \mathbb{R}^{F \cup \{\chi_{\mathcal{I}_0}\}},$$

with the standard dualities $\langle \lambda, f\rangle = \lambda(f)$ on measure spaces and continuous functions and the Euclidean inner product on $\mathbb{R}^{F \cup \{\chi_{\mathcal{I}_0}\}}$.

However, when (for example) G is infinite, the duality becomes more complicated. Suppose $G \cong U$ for some set U with $\deg(g)$ constant for $g \in G$; we denote this degree by $\deg(G)$. Let ω_G be a uniform, finite, positive measure on G (for example, $G \cong S^{n-1}$ with the standard $O(n)$-invariant probability measure ω_n). Define

$$(A_k^G)^*(\lambda) = (A_k^g)^*(\lambda) d\omega_G(g),$$

so that $(A_k^G)^*(\lambda) \in \mathcal{M}(\mathcal{I}_{k_G} \times \mathcal{I}_{k_G} \times G)$, where $k_G = k - \lceil \deg(G)/2 \rceil$. Alternatively, we can view G as the set of functions

$$\{Q \mapsto g(Q, u) : u \in U\},$$

where $g \in C(\mathcal{I}_\ell \times U)$, and we assume G has this form in the remainder of this chapter. Then

$$A_k^G K(S) = \int \sum_{\substack{J_1, J_2 \in \mathcal{I}_{k_G}, Q \in \mathcal{I}_{\deg(G)} \\ J_1 \cup J_2 \cup Q = S}} g(Q, u) K(J_1, J_2, u) d\omega_U(u).$$

A similar approach works when G is isomorphic to a finite, disjoint union of infinite sets.

Example 2.3 (Covering, continued). *Take $V = S^{n-1}$, and suppose we want to cover V with spherical caps*

$$B(y, r) = \{x \in S^{n-1} : \langle x, y\rangle \geq r\}.$$

The covering constraints for a finite graph are of the form

$$g_j = \sum_{i:\{i,j\}\in E} x_i - 1 = \sum_{i\in V} \chi_{B(j)}(i)x_i - 1 \geq 0,$$

where $B(j)$ contains all i such that $\{i,j\} \in E$. For the sphere, this gives a constraint for every $y \in S^{n-1}$, with

$$g(S,y) = \begin{cases} -1 & \text{for } S = \emptyset, \\ 1 & \text{for } S = \{x\} \text{ and } x \in B(y,r), \\ 0 & \text{otherwise.} \end{cases} \tag{2.4.1}$$

Here we took $G \cong S^{n-1}$ with the standard $O(n)$-invariant probability measure ω_n on S^{n-1}. The operator $A_k^G : C(\mathcal{I}_{k-1} \times \mathcal{I}_{k-1} \times S^{n-1}) \to C(\mathcal{I}_{2k})$ is then defined by

$$A_k^G K(S) = \int \sum_{\substack{x\in S^{n-1}, J_1, J_2 \in \mathcal{I}_{k-1} \\ J_1 \cup J_2 \cup \{x\} = S}} \chi_{B(y,r)}(x) K(J_1, J_2, y) - \sum_{\substack{J_1, J_2 \in \mathcal{I}_{k-1} \\ J_1 \cup J_2 = S}} K(J_1, J_2, y) d\omega_n(y).$$

Then the dual problem to problem (2.3.1) becomes

$$\begin{aligned} \text{minimize} \quad & a_0 + \int_F a(f)b(f)d\omega_F(f), \\ \text{subject to} \quad & -A_k K - \int_G A_k^g K'_g d\omega_G(g) \\ & \quad + a_0\chi_{\mathcal{I}_0} + \int_F a(f) f d\omega_F(f) \geq h \qquad \text{on } \mathcal{I}_{2k}, \\ & K' \in C(\mathcal{I}_{k_G} \times \mathcal{I}_{k_G} \times G) \text{ is slice-positive}, \\ & K \in C(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}. \end{aligned} \tag{2.4.2}$$

Here, a function $K \in C(A \times A \times B)$ is slice-positive if for every $b \in B$ the kernel K_b defined by

$$K_b(x,y) = K(x,y,b)$$

is a positive definite kernel; we denote the cone of slice-positive functions by $C(A \times A \times B)_{\succeq 0}$. Note that, if F or G is finite, the uniform measure is the counting measure.

Example 2.1 (Packing, continued). *Dualizing* (2.3.2) *gives*

$$\begin{aligned} \textit{minimize} \quad & a_0, \\ \textit{subject to} \quad & -A_k K + a_0\chi_{\mathcal{I}_0} \geq \chi_{\mathcal{I}_{=1}} \qquad \text{on } \mathcal{I}_{2k}, \\ & K \in C(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}. \end{aligned}$$

The only constraint on a_0 is given by $K(\emptyset,\emptyset) \leq a_0$, and hence we can reformulate the problem as

$$\begin{aligned} \textit{minimize} \quad & K(\emptyset,\emptyset), \\ \textit{subject to} \quad & A_k K \leq -\chi_{\mathcal{I}_{=1}} \qquad \text{on } \mathcal{I}_{2k} \setminus \mathcal{I}_0, \\ & K \in C(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}. \end{aligned} \tag{2.4.3}$$

2.5. Strong duality

Recall that weak duality says that the optimal objective value of the maximization problem of a primal-dual pair is always at most the optimal objective value of the minimization problem. However, it is possible that the inequality is strict. Strong duality means that the inequality is guaranteed to be an equality. Additionally, if the optimal objective function value is finite, it is attained by a primal optimal solution. Theorem 2.7 shows this is true for any reasonable Lasserre hierarchy. This is essentially the same proof as for the original Lasserre hierarchy for packing problems [**92**].

THEOREM 2.7 (Strong duality). *If the k-the level of the Lasserre hierarchy* (2.3.1) *has a feasible solution, strong duality holds.*

PROOF. From Theorem IV.7.2 of [**5**], it follows strong duality holds if the cone

$$\begin{aligned}\{(A_k^*(\lambda) - \mu, (A_k^g)^*(\lambda)d\omega_G(g) - \xi, \lambda(f)d\omega_F(f), \lambda(\mathcal{I}_0), \lambda(h)) :\\ \lambda \in \mathcal{M}(\mathcal{I}_{2k})_{\succeq 0}, \mu \in \mathcal{M}(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0},\\ \xi \in \mathcal{M}(\mathcal{I}_{k_G} \times \mathcal{I}_{k_G} \times G)_{\succeq 0}\}\end{aligned}$$

is closed. The cone is the Minkowski difference of

$$K_1 = \{(A_k^*(\lambda), (A_k^g)^*(\lambda)d\omega_G(g), \lambda(f)d\omega_F(f), \lambda(\mathcal{I}_0), \lambda(h)) : \lambda \in \mathcal{M}(\mathcal{I}_{2k})_{\succeq 0}\}$$

and

$$K_2 = \{(\mu, \xi, 0, 0, 0) : \mu \in \mathcal{M}(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}, \xi \in \mathcal{M}(\mathcal{I}_{k_G} \times \mathcal{I}_{k_G} \times G)_{\succeq 0}\}.$$

By [**78**] and [**49**], P is closed if

- $K_1 \cap K_2 = \{0\}$,
- K_1 and K_2 are closed, and
- K_1 is locally compact.

The first condition holds by Lemma 5 of [**92**]: if $\lambda(\{\emptyset\}) = 0$, $A_k^*(\lambda) \succeq 0$ and $\lambda \geq 0$, then $\lambda = 0$ by the lemma, and therefore the element in K_1 corresponding to the only λ in the intersection equals 0.

The cone K_2 is clearly closed. Following the proof of Lemma 6 of [**92**], the cone K_1 is closed and locally convex. Again, the result for K_1 follows because it extends $(A_k^*(\lambda), \lambda(\mathcal{I}_0))$. □

2.6. Convergence in a finite number of steps

In this section we consider conditions such that the Lasserre hierarchy converges in a finite number of steps.

Let α be the maximum size of an independent set; if F consists of the cardinality constraints defined in Example 2.6 for number N, we assume that there is an optimal N-point configuration which is an independent set in the topological packing graph. In particular this implies $N \leq \alpha$ in that case.

We say that a configuration $R \in \mathcal{I}_k$ is *invalid* if there is a point $u \in U$ such that

$$\sum_{Q \subseteq R'} g(Q, u) < 0.$$

An invalid configuration is not a solution to the original problem. A configuration is *strictly invalid* if additionally there is a neighborhood U' of such a point u and

neighborhoods $T_1, \ldots, T_{|R|}$ of the elements of R such that

$$\sum_{Q \subseteq R'} g(Q, u') < 0$$

for every $R' \in \{\{x_1, \ldots, x_{|R|}\} : x_i \in T_i|\}$ and $u' \in U'$. For continuous functions g, and for geometric problems such as covering the sphere with spherical caps, every invalid configuration is strictly invalid. Intuitively, v is a point that is not 'covered' by R, and small enough perturbations of both R and v retain this property.

THEOREM 2.8. *Suppose $C \in \mathcal{I}_N$ is an optimal configuration for some $N \in \mathbb{N}$, and*

- *either $F = \emptyset$ or F consists of cardinality constraints as in Example 2.6,*
- *$G = \{Q \mapsto g(Q, u) : u \in U\}$ for some piecewise continuous g, and*
- *every invalid configuration is strictly invalid.*

Let $i = \max_{g \in G} \lceil \deg(g)/2 \rceil$. Then $\mathrm{las}_{\alpha+i} = \chi_C(h)$. Moreover, if F consists of cardinality constraints, then $\mathrm{las}_{N+i} = \chi_C(h)$.

Note that the condition on G also includes G finite, with $U = \{1, \ldots, m\}$ for some $m \in \mathbb{N}$.

The proof of this theorem is similar to the proof of Lemma 5.5.3 in [**125**] for the symmetrized covering hierarchy, and uses the following result of [**92**].

LEMMA 2.9 ([**92**, Proposition 1]). *Let $\lambda \in \mathcal{M}(\mathcal{I}_\alpha)$ be such that $\lambda(\{\emptyset\}) = 1$ and $A_\alpha^* \lambda \in \mathcal{M}(\mathcal{I}_\alpha \times \mathcal{I}_\alpha)_{\geq 0}$. Then there exists a unique probability measure $\sigma \in \mathcal{M}(\mathcal{I}_\alpha)_{\geq 0}$ such that*

$$\lambda = \int \chi^R d\sigma(R)$$

where

$$\chi^R = \sum_{Q \subseteq R} \delta_Q.$$

PROOF OF THEOREM 2.8. Since α is the maximum size of an independent set, $\mathcal{I}_{\alpha+1} = \mathcal{I}_\alpha$ and hence $\mathrm{las}_{\alpha+k} = \mathrm{las}_{\alpha+i}$ for all $k \geq i$. Moreover, if F contains cardinality constraints, we have

$$\lambda(\mathcal{I}_{=(N+l)}) = 0$$

for all $l \geq 0$ and hence $\mathrm{las}_{N+k} = \mathrm{las}_{N+i}$ for all $k \geq i$, as long as $\alpha(\varepsilon) \geq N$.

Since C is an optimal configuration, the measure χ_C is a feasible solution of the problem by construction. Hence by Theorem 2.7, there is an optimal measure λ, and by Lemma 2.9, this measure is of the form

$$\lambda = \int \chi^R d\sigma(R)$$

for some probability measure $\sigma \in \mathcal{M}(\mathcal{I}_\alpha)_{\geq 0}$.

Suppose F consists of cardinality constraints. Then

$$1 = \lambda(\mathcal{I}_{=N}) = \int \chi_R(\mathcal{I}_{=N}) d\sigma(R) = \sigma(\mathcal{I}_{=N})$$

so σ is supported on $\mathcal{I}_{=N}$. In particular, we have $\sigma(X) = 0$ where X is given by the set

$$\{R \in \mathcal{I}_t : R \text{ does not satisfy } F\}.$$

Consider a configuration $R = \{x_1, \dots, x_{|R|}\}$ that is invalid (and therefore strictly invalid by assumption). That is, there is a neighborhood U' of some $u \in U'$ and neighborhoods $T_1, \dots T_{|R|}$ of the elements of R such that

$$\sum_{S \subseteq R'} g(S, u') < 0 \tag{2.6.1}$$

for all $u' \in U'$ and $R' \in T_R = \{\{x_1, \dots, x_{|R|}\} : x_i \in T_i\}$. Define $T = \bigcup_i T_i$.

Define the function $\psi_T : \mathcal{I}_\alpha \to \mathbb{R}$ by

$$\psi_R(J) = \begin{cases} (-1)^{|J \setminus T|} & \text{if } |T \cap J| = |R|, \\ 0 & \text{otherwise,} \end{cases}$$

so that

$$\sum_{S \subseteq R'} \psi_R(S) = \begin{cases} 1 & \text{if } |R'| = |R| = |T \cap R| \\ 0 & \text{otherwise,} \end{cases}$$

by the inclusion-exclusion principle. Note that this equals the indicator function of T_R.

Let l_1^k be a sequence of continuous functions weakly converging to ψ_R as $k \to \infty$. For constructing such a sequence, notice that ψ_R can be written as

$$\psi_R(J) = \left(\sum_{i=0}^{\alpha - |R|} (-1)^i 1_{\mathcal{I}_{=|R|+i}}(J) \right) \prod_{x \in R} \left(\sum_{y \in J} 1_{T_i}(y) \right),$$

since the sign purely depends on the cardinality of J, and every neighborhood T_i around a point $x \in R$ can contain at most one element of J and should contain at least one element of J for a nonzero value.

It remains to show that we can approximate $1_{T_i}(y)$ by continuous functions. Recall that V is a compact Hausdorff space, so V is normal. That means we can apply Urysohn's lemma (see, e.g., [**142**]), which gives that any two closed subsets of V can be separated by a continuous function. Let $\{B_k\}_k$ be a sequence of closed sets such that

$$B_1 \subseteq B_2 \subseteq \cdots \subseteq T_i$$

with

$$\bigcup_{k=0}^{\infty} B_k = T_i,$$

and let f_k be a continuous function with $f_k(B_k) = \{1\}$ and $f(V \setminus T_i) = \{0\}$, which exists by Urysohn's lemma. Then since B_k converges to T_i as $k \to \infty$, we have $f_k \to 1_{T_i}$ ν-almost surely for every Radon measure ν.

We will show that $\sigma(T) = 0$. Let $l_2 : V \to \mathbb{R}$ be a nonnegative continuous with support U. Then for every k, we have by feasibility of λ that

$$0 \leq \langle (A^G_{\alpha+i})^*(\lambda), l_1^k \otimes l_1^k \otimes l_2 \rangle.$$

Taking the limit of $k \to \infty$ then gives

$$\begin{aligned}
0 &\leq \int \langle (A^G_{\alpha+i})^*(\chi_R), \psi_R \otimes \psi_R \otimes l_2 \rangle d\sigma(R') \\
&= \iint \sum_{S \subseteq R'} \sum_{\substack{Q \in \mathcal{I}_{\deg(g)}, J_1, J_2 \in \mathcal{I}_\alpha \\ J_1 \cup J_2 \cup Q = S}} g(Q,u)(\psi_R \otimes \psi_R \otimes l_2)(J_1, J_2, u) d\sigma(R') d\omega_U(u) \\
&= \iint \sum_{S \subseteq R'} g(S,u) l_2(u) \sum_{J_1, J_2 \subseteq R'} (\psi_R \otimes \psi_R)(J_1, J_2) d\sigma(R') d\omega_U(u) \\
&= \iint \sum_{S \subseteq R'} g(S,u) l_2(u) 1_T(R') d\sigma(R') d\omega_U(u).
\end{aligned}$$

by the inclusion-exclusion principle and the definition of ψ_R. Using (2.6.1) and the fact that l_2 is positive on U' and zero elsewhere, this is negative unless $\sigma(T_R) = 0$.

The set $\mathcal{X}$ of all configurations R that are invalid equals

$$\bigcup_{R \in \mathcal{X}} T_R.$$

By [**82**, Section 3], the uncountable union of open sets of measure 0 gives a set of measure 0, and since T_R is open, this gives that $\sigma(\mathcal{X}) = 0$.

For every valid configuration R, we have $\chi_R(h) \leq \chi_C(h)$ because C is optimal, and together with $\sigma(\mathcal{X}) = 0$ this gives that

$$\lambda(h) = \int \chi_R(h) d\sigma(R) \leq \chi_C(h) \sigma(\mathcal{I}_\alpha) = \chi_C(h).$$

Furthermore, χ_C is a feasible measure, and therefore $\lambda(h) \geq \chi_C(h)$. Hence $\mathrm{las}_\alpha = \chi_C(h)$. □

2.7. Complementary slackness

Complementary slackness gives conditions that any pair of optimal primal and dual solutions with the same objective value satisfies. That is, given an optimal solution (K, K', a), we can derive conditions any optimal configuration C satisfies, and given an optimal configuration C, we can derive conditions any optimal solution (K, K', a) satisfies. In Part II of this thesis, this allows us to extract proofs that certain configurations are the unique optima, up to isometry; see for instance Section 7.6.2.

Complementary slackness is widely known in the context of conic programming (see, e.g., [**5**, Theorem IV.6.2] and [**19**, Section 5.5.2]).

THEOREM 2.10 (Complementary slackness). *Suppose (K, K', a) is an optimal solution to (2.4.2), and C is an optimal configuration with $a_0 + \int_F a(f)b(f)d\omega_F(f) = \chi_C(h)$. Then*

- *For every $S \subseteq C$ of size at most $2k$, we have*

$$-A_k K(S) - \int_G A^g_k K'(S) d\omega_G(g) + a_0 \chi_{\mathcal{I}_0}(S) + \int_F a(f) f(S) d\omega_F(f) = h(S).$$

- *We have*

$$\sum_{S \subseteq C, |S| \leq 2k} A_k K(S) = 0,$$

and for every measurable set $U \subseteq G$,

$$\int_U \sum_{S\subseteq C, |S|\leq 2k-\deg(g)} A_k^g K_g(S) d\omega_G(g) = 0,$$

where $K_g(J_1, J_2) = K'(J_1, J_2, g)$. Moreover, if K respectively K_g is a sum of positive definite kernels K_λ, then the equations hold for K_λ as well.

Proof. We first prove weak duality. We have

$$\begin{aligned}
0 &\leq \langle \chi_C, (-A_k K - \int_G A_k^g K' d\omega_G(g) + a_0 \chi_{I_0} + \int_F a(f) f d\omega_F(f) - h\rangle \\
&= -\sum_{S\subseteq C, |S|\leq 2k} A_k K(S) - \int_G \sum_{S\subseteq C, |S|\leq 2k-d_g} A_k^g K'(S) d\omega_G(g) \\
&\quad + a_0 + \int_F a(f)\chi_C(f) d\omega_F(f) - \chi_C(h) \\
&\leq a_0 + \int_F a(f) b(f) d\omega_F(f) - \chi_C(h)
\end{aligned}$$

where we used the constraints of both the primal and dual problem. Suppose $\chi_C(h) = a_0 + \int_F a(f)b(f)d\omega_F(f)$. Then we have equality throughout the equation, so in particular for every $S \in \mathcal{I}_{2k}$ if $S \subseteq C$,

$$-A_k K(S) - \int_G A_k^g K'(S) d\omega_G(g) + a_0 \chi_{\mathcal{I}_0}(S) + \int_F a(f) f(S) d\omega_F(f) = h(S).$$

Since K is positive definite, we have

$$\sum_{\substack{S\subseteq C \\ |S|\leq 2k}} A_k K(S) = \sum_{\substack{J_1, J_2 \subseteq C \\ |J_i|\leq k}} K(J_1, J_2) \geq 0,$$

and similarly, since K' is continuous and slice-positive, and C is a feasible configuration, we have

$$\begin{aligned}
&\sum_{\substack{S\subseteq C \\ |S|\leq 2k}} \int_g A_k^g K'(S) d\omega_G(g) \\
&= \int \sum_{\substack{Q\subseteq C \\ |Q|\leq \deg(g)}} g(Q) \sum_{\substack{J_1, J_2 \subseteq C \\ |J_1|, |J_2|\leq k_G}} K'(J_1, J_2, g) d\omega_G(g) \geq 0.
\end{aligned}$$

In particular, because all inequalities in the weak duality proof are equalities this implies that

$$\sum_{\substack{S\subseteq C \\ |S|\leq 2k}} A_k K(S) = \sum_{\substack{S\subseteq C \\ |S|\leq 2k}} \int_g A_k^g K'(S) d\omega_G(g) = 0$$

Now, if K is the sum of positive definite kernels K_λ, we have

$$0 = \sum_{\substack{S\subseteq C \\ |S|\leq 2k}} A_k K(S) \geq \sum_{\substack{S\subseteq C \\ |S|\leq 2k}} A_k K_\lambda(S) \geq 0,$$

and therefore the equality also holds for K_λ.

Similarly, we also have

$$\int_G \sum_{\substack{S\subseteq C\\ |S|\leq 2k}} A_k^g K'_\lambda(S) d\omega_G(g) = 0$$

if K' is a sum of slice-positive kernels K'_λ. Since ω_G is a nonnegative measure, and

$$\sum_{\substack{S\subseteq C\\ |S|\leq 2k}} A_k^g K'_\lambda \geq 0$$

for every g, the integral vanishes if and only if it vanishes on every measurable subset of G. □

CHAPTER 3

Symmetry reduction and harmonic analysis

In Chapter 2, we derived a general hierarchy of optimization problems for problems on a topological packing graph (V, E). The main objects over which the optimization takes place are positive definite kernels $K \in C(\mathcal{I}_k \times \mathcal{I}_k)$ and slice-positive functions in $K' \in C(\mathcal{I}_k \times \mathcal{I}_k \times G)$, where $\mathcal{I}_k$ is the set of independent sets of size at most k in the graph, and $G \subseteq C(\mathcal{I}_{2k})$. In our applications, $G = \emptyset$ or $G \cong V$.

To perform explicit computations on such a hierarchy, we use the symmetry of the problem. Suppose the problem is invariant under a group Γ: if $X \subseteq V$ is a feasible configuration, then γX is also a feasible configuration with the same objective value, for all $\gamma \in \Gamma$. In general, the hierarchy then admits the same or closely related symmetries: if (K, K', a) is a feasible solution, then $(\gamma K, \gamma K', a)$ will also be a feasible solution, where γ acts on K as

$$\gamma K(J_1, J_2) = K(\gamma^{-1} J_1, \gamma^{-1} J_2)$$

and, for example, on K' with $G \simeq V$ as

$$\gamma K'(J_1, J_2, g) = K'(\gamma^{-1} J_1, \gamma^{-1} J_2, \gamma^{-1} g).$$

This implies that averaging also gives a feasible kernel with the same objective, and hence K and K' can be taken to be invariant under the action of Γ. In this chapter we consider such an invariant kernel $K \in C(V \times V)^{\Gamma}_{\succeq 0}$.

Example 2.1 (Packing, continued). *Consider the independent set problem on the sphere S^{n-1}, where distinct points x, y are adjacent if $\langle x, y \rangle > \cos\theta$. Then $\mathcal{I}_k$ is invariant under $O(n)$, where $O(n)$ acts as*

$$\gamma\{x_1, \ldots, x_k\} = \{\gamma x_1, \ldots, \gamma x_k\}.$$

Furthermore, if K is feasible, then γK defined by

$$(\gamma K)(x, y) = K(\gamma^{-1} x, \gamma^{-1} y).$$

Indeed, suppose K satisfies the constraint

$$A_k K(S) \leq -\chi_{\mathcal{I}_{=1}}(S).$$

Then

$$A_k(\gamma K)(S) = A_k K(\gamma^{-1} S) \leq -\chi_{\mathcal{I}_{=1}}(\gamma^{-1} S) = -\chi_{\mathcal{I}_{=1}}(S),$$

since $\mathcal{I}_{=i}$ is an invariant subset of $\mathcal{I}_{2k}$ for every $i = 0, \ldots, 2k$. Furthermore, the kernel γK has the same objective function value. Since $O(n)$ is a compact group, we can define $\overline{K}$ by

$$\overline{K} = \int_{O(n)} \gamma K \, d\gamma$$

as a feasible solution with objective function value $\overline{K}(\emptyset,\emptyset) = K(\emptyset,\emptyset)$*, where the integration uses the Haar measure. This implies that we can restrict to* $O(n)$*-invariant kernels* $K \in C(\mathcal{I}_k \times \mathcal{I}_k)^{O(n)}_{\succeq 0}$.

In Section 3.4, we show that such a kernel can be written in the form

$$K(J_1, J_2) = \sum_{\lambda} \langle K_\lambda, Z_\lambda(J_1, J_2)\rangle,$$

where the entries of Z_λ *are polynomials in the inner products between vectors in* $J_1 \cup J_2$.

Essentially, there are two methods that can be considered for symmetry reduction. The first is based on so-called symmetry adapted bases (see Section 3.2), which is more suited for finite groups acting on finite dimensional vector spaces, while the second uses Γ-equivariant maps (Section 3.3) and is more natural for infinite groups and infinite dimensional vector spaces.

3.1. Preliminaries

In this section we give a quick overview of some important definitions and results in representation theory. A more extensive introduction may be found in, e.g., [**57, 130, 131**].

A *representation* of a group Γ on a vector space V is a group homomorphism $\pi : \Gamma \to \mathrm{GL}(V)$. A subspace W of V is *invariant* if $\pi(\gamma)W \subseteq W$ for all $\gamma \in \Gamma$, and a representation (π, V) is *irreducible* if its only invariant subspaces are $\{0\}$ and V.

Given representations (π_i, V_{π_i}) for $i = 1, 2$, a linear map $A : V_1 \to V_2$ is an intertwining map if

$$A\pi_1(\gamma) = \pi_2(\gamma)A$$

for all $\gamma \in \Gamma$. The representations are called equivalent if A is invertible.

The *direct sum* of two representations $(\pi_1, V_1), (\pi_2, V_2)$ is the representation $(\pi_1 \oplus \pi_2, V_1 \oplus V_2)$ given by

$$(\pi_1 \oplus \pi_2)(\gamma)(v_1, v_2) = (\pi_1(\gamma)v_1, \pi_2(\gamma)v_2).$$

A representation is *completely reducible* if it is equivalent to a direct sum of irreducible representations. Given an inner product on V, a representation (π, V) is *unitary* if $\pi(\gamma)$ is a unitary operator on V for all $\gamma \in \Gamma$.

LEMMA 3.1 (Schur's Lemma (see, e.g., [**131**, Section 2.2]). *Let* $(\pi_1, V_1), (\pi_2, V_2)$ *be irreducible representations on finite-dimensional vector spaces* V_1, V_2*, and let* A *be an intertwining map. Then*

- A *is an isomorphism or* $A = 0$.
- *if* $V_1 = V_2$*, then* $A = \lambda I$ *for some* $\lambda \in \mathbb{C}$.

3.2. Symmetry adapted bases*

Let Γ be a finite group, acting on a finite dimensional vector space V by $L : \Gamma \to \mathrm{GL}(V)$. When V is infinite, we can restrict ourselves to a finite-dimensional

*This section is based on Section 4 of the publication "*N. Leijenhorst and D. de Laat,* Solving clustered low-rank semidefinite programs arising from polynomial optimization, Math. Program. Comput.* **16** *(2024), no. 3, 503-534, doi:10.1007/s12532-024-00264-w*".

Γ-invariant subspace of V. By Maschke's theorem (see, e.g., [**57**, Corollary 1.6]), V admits the decomposition

$$V = \bigoplus_{\pi \in \widehat{\Gamma}} \bigoplus_{i} H_{\pi,i}$$

into irreducible representations, where $\widehat{\Gamma}$ denotes a complete set of irreducible representations $\pi : \Gamma \to \mathrm{GL}(V_\pi)$, and $H_{\pi,i}$ is equivalent to V_π.

Since Γ is finite, there is a finite basis $e_{\pi,1}, \dots, e_{\pi,d_\pi}$ of V_π for each irreducible representation (π, V_π), in which $\pi(\gamma)$ are unitary matrices. A symmetry adapted basis is a basis $\{e_{\pi,i,j}\}_{\pi,i,j}$ of V such that $\{e_{\pi,i,j}\}_j$ is a basis of the invariant subspace $H_{\pi,i}$, and the matrix $L(\gamma)$ in this basis equals $\pi(\gamma)$. Typically, we call $\{e_{\pi,i,j}\}$ a symmetry adapted system if V is a finite-dimensional invariant subspace of an infinite space.

Such a basis exists since $H_{\pi,i}$ is equivalent to V_π, so there are Γ-equivariant isomorphisms $T_{\pi,i} \colon V_\pi \to H_{\pi,i}$ and we can define $e_{\pi,i,j} = T_{\pi,i} e_{\pi,j}$. Then it follows that $L(\gamma) e_{\pi,i,j} = \sum_k \pi(\gamma)_{k,j} e_{\pi,i,k}$. As described in [**131**, Section 2.7] a symmetry-adapted basis can be constructed by defining the operators

$$p^\pi_{j,j'} = \frac{d_\pi}{|\Gamma|} \sum_{\gamma \in \Gamma} \pi(\gamma^{-1})_{j',j} L(\gamma), \tag{3.2.1}$$

and then choosing bases $\{e_{\pi,i,1}\}_i$ of $\mathrm{Im}(p^\pi_{1,1})$ and setting $e_{\pi,i,j} = p^\pi_{j,1} e_{\pi,i,1}$. Then,

$$\begin{aligned} L(\tilde{\gamma}) e_{\pi,i,j} &= \frac{d_\pi}{|\Gamma|} \sum_{\gamma \in \Gamma} \pi(\gamma^{-1})_{1,j} L(\tilde{\gamma}\gamma) e_{\pi,i,1} = \frac{d_\pi}{|\Gamma|} \sum_{\gamma \in \Gamma} \pi(\gamma^{-1}\tilde{\gamma})_{1,j} L(\gamma) e_{\pi,i,1} \\ &= \frac{d_\pi}{|\Gamma|} \sum_{\gamma \in \Gamma} \sum_{k=1}^{d_\pi} \pi(\gamma^{-1})_{1,k} \pi(\tilde{\gamma})_{k,j} L(\gamma) e_{\pi,i,1} = \sum_{k=1}^{d_\pi} \pi(\tilde{\gamma})_{k,j} e_{\pi,i,k}. \end{aligned}$$

That is, one can explicitly build a symmetry adapted basis for a given finite dimensional space, when given a description of the irreducible representations.

Let K be a positive definite kernel on V. Since V is finite dimensional, we can express K in a basis b to get a matrix of size $\dim(V) \times \dim(V)$. An element $x \in V$ can be expressed in the basis as $c(x)^* b$, where $c(x) \in \mathbb{C}^{\dim(V)}$ is a vector of coefficients and * denotes the conjugate transpose. We have

$$L(\gamma) x = c(x)^* L(\gamma) b,$$

so that an invariant kernel then has the property that

$$c(x)^* K c(y) = c(x)^* L(\gamma) K L(\gamma)^* c(y)$$

for all $x, y \in V$.

Proposition 3.2. *Suppose $K \in C(V \times V)_{\succeq 0}$ is invariant. Expressing K in a symmetry adapted basis $b = (e_{\pi,i,j})_{\pi,i,j}$ block-diagonalizes K as*

$$K = \bigoplus_{\pi \in \widehat{\Gamma}} \frac{1}{d_\pi} K_\pi \otimes I_{d_\pi}$$

for some Hermitian positive semidefinite matrices $K_\pi \in \mathbb{C}^{m_\pi \times m_\pi}$. Here m_π is the multiplicity of π in V.

PROOF. Note that expressing $L(\gamma)$ in the symmetry adapted basis b gives

$$L(\gamma) = \bigoplus_{\pi \in \widehat{\Gamma}} I_{m_\pi} \otimes \pi(\gamma).$$

Then for $x = c(x)^* b \in V, \gamma \in \Gamma$, we have

$$L(\gamma)x = c(x)^* L(\gamma) b$$

and hence

$$c(x)^* K c(y) = c(L(\gamma)x)^* K c(L(\gamma)y) = c(x)^* L(\gamma) K L(\gamma)^* c(y)$$

for every $x, y \in V$ and $\gamma \in \Gamma$. In particular, this gives for all $\gamma \in \Gamma$ that

$$K = L(\gamma) K L(\gamma)^*.$$

Writing K in block form $K = (K_{(\pi,i),(\pi',i')})$ then gives, using the expression of $L(\gamma)$ in this basis, that

$$K_{(\pi,i),(\pi',i')} \pi'(\gamma) = \pi(\gamma) K_{(\pi,i),(\pi',i')},$$

for all $\gamma \in \Gamma$: the blocks of K are intertwining operators. By Schur's lemma (Lemma 3.1), the block $K_{(\pi,i),(\pi',i')}$ is a multiple of the identity if $\pi = \pi'$, and 0 otherwise, as we wanted to prove. □

EXAMPLE 3.3 (Schoenberg's theorem [**128**]). *Let $V = \mathrm{Pol}(S^{n-1})_{\leq d}$ be the space of polynomials on S^{n-1} of degree at most d, and consider the action of $O(n)$ on V defined by*

$$\gamma p(x) = p(\gamma^{-1} x).$$

where $O(n)$ acts on S^{n-1} naturally. The representations of $O(n)$ are given by the spaces

$$\mathrm{Harm}_k^n = \{p \in \mathbb{R}[x_1, \ldots, x_n] : p \text{ is homogeneous}, \deg p = k, \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} p = 0\}$$

of homogeneous, harmonic polynomials of degree k. The representations have dimension

$$h_k^n = \binom{n+k-1}{n-1} - \binom{n+k-3}{n-1}.$$

A symmetry adapted basis for V is given by the spherical harmonics, which immediately shows that $m_\pi = 1$ for every representation. Using Proposition 3.2, we get

$$K(x,y) = \sum_k a_k \frac{1}{h_k^n} \sum_{j=1}^{h_k^n} e_{k,j}(x) e_{k,j}(y).$$

It is known that here the addition formula holds (see, e.g., [**1**, Chapter 9.6]*:*

$$\frac{1}{h_k^n} \sum_{j=1}^{h_k^n} e_{k,j}(x) e_{k,j}(y) = P_k^n(x \cdot y),$$

where P_k^n are the Gegenbauer polynomials with parameter $n/2 - 1$. So every kernel of the form

$$K(x,y) = \sum_k a_k P_k^n(x \cdot y).$$

with $a_k \geq 0$ is $O(n)$-invariant and positive definite, and in fact every continuous, $O(n)$-invariant, positive definite kernel on S^{n-1} is the uniform limit of such kernels; this result is known as Schoenberg's Theorem [**128**].

EXAMPLE 3.4 (Sum of squares polynomials). *Consider $V = \mathbb{C}[x_1, \ldots, x_n]_{\leq d}$, and suppose Γ acts on $\mathbb{C}^n$ by γx. The natural action of Γ on V is then given by*

$$\gamma p(x) = p(\gamma^{-1} x).$$

Suppose p is a sum-of-squares polynomial; that is, we can write

$$p(x) = \sum_i (c_i^* b)^* (c_i^* b) = b^* A b$$

for some Hermitian positive semidefinite matrix A, where b is a basis of V. If b is a symmetry adapted basis, we have $L(\gamma)b(x) = b(\gamma^{-1}x)$ for all x, γ. If p is Γ-invariant, we can take the average of p over the group to obtain

$$p(x) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} b(\gamma^{-1}x)^* A b(\gamma^{-1}x) = b(x)^* B b(x),$$

where

$$B = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} L(\gamma)^* A L(\gamma)$$

is a positive semidefinite, Γ-invariant kernel on V. By Proposition 3.2, we can write

$$B = \bigoplus_{\pi \in \widehat{\Gamma}} B_\pi \otimes I_{d_\pi},$$

and hence

$$\begin{aligned} p(x) &= \langle B, b(x)b(x)^* \rangle \\ &= \left\langle \bigoplus_\pi \frac{1}{d_\pi} C_\pi \otimes I_{d_\pi}, b(x)b(x)^* \right\rangle \\ &= \sum_\pi \sum_{j=1}^{d_\pi} \sum_{i,i'=1}^{m_\pi} \frac{1}{d_\pi} (C_\pi)_{i,i'} e_{\pi,i,j}(x) e_{\pi,i',j}(x)^* \\ &= \sum_\pi \big\langle C_\pi, E_\pi(x) \big\rangle, \end{aligned}$$

where we defined

$$E_\pi(x)_{i,i'} = \frac{1}{d_\pi} \sum_{j=1}^{d_\pi} e_{\pi,i,j}(x) e_{\pi,i',j}(x)^*.$$

3.3. Invariant kernels on infinite sets

In this section we consider infinite spaces. We give a brief overview of the topic; for a more thorough treatment, see [**84**, Chapter 3] and [**113**].

Suppose Γ is a compact group acting continuously on a compact set V. Let $K \in C(V \times V)$ be a positive definite, Γ-invariant kernel. Let (π, W) be an irreducible representation of Γ, and denote the space of continuous, Γ-equivariant maps from V to W by $\mathrm{Hom}_\Gamma(V, W)$. That is, for $\psi \in \mathrm{Hom}_\Gamma(V, W)$ we have

$$\psi(\gamma v) = \pi(\gamma)\psi(v).$$

Given a family $\{\psi_{\pi,l}\}$ of elements in this space, define $Z_\pi(x,y)$ by

$$Z_\pi(x,y)_{l_1,l_2} = \langle \psi_{\pi,l_1}(x), \psi_{\pi,l_2}(y)\rangle$$

for $x, y \in V$, where the inner product on W is used.

Comparing with Section 3.2, the inner product

$$\langle \psi_{\pi,l_1}(x), \psi_{\pi,l_2}(y)\rangle$$

essentially corresponds with the sum

$$\sum_{j=1}^{d_\pi} e_{\pi,l_1,j}(x) e_{\pi,l_2,j}(y)^*.$$

The matrix Z_π has the following properties:

- Z_π is Γ invariant: we have
$$\begin{aligned}\gamma Z_\pi(x,y)_{l_1,l_2} &= \langle \psi_{\pi,l_1}(\gamma^{-1}x), \psi_{\pi,l_2}(\gamma^{-1}y)\rangle \\ &= \langle \pi(\gamma)\psi_{\pi,l_1}(x), \pi(\gamma)\psi_{\pi,l_2}(y)\rangle \\ &= Z_\pi(x,y)_{l_1,l_2}\end{aligned}$$
since $\pi(\gamma)$ is unitary.
- The kernel $(x,y) \mapsto \langle K_\pi, Z_\pi(x,y)\rangle$ is positive definite if K_π is a positive semidefinite matrix. Indeed, given points $\{x_i\}_{i=1}^N \subseteq V$, the submatrix of K corresponding to these points has entries
$$\langle K_\pi, Z_\pi(x_i,x_j)\rangle = \sum_{l_1,l_2} (K_\pi)_{l_1,l_2} \langle \psi_{\pi,l_1}(x_i), \psi_{\pi,l_2}(x_j)\rangle,$$
The matrix with entries $(K_\pi)_{l_1,l_2}\langle \psi_{\pi,l_1}(x_i), \psi_{\pi,l_2}(x_j)\rangle$ is a entrywise product of a positive semidefinite matrix with a Gram matrix, and as such, is positive semidefinite. In particular this implies that the submatrix of K is positive semidefinite.

Hence any kernel of the form

$$(x,y) \mapsto \sum_\pi \langle K_\pi, Z_\pi(x,y)\rangle$$

where the positive semidefinite matrices K_π have a finite number of nonzero entries, is a continuous, Γ-invariant, positive definite kernel. Under certain assumptions on V, the action of Γ, and $\{\psi_{\pi,l}\}$, it can be proven that every continuous, Γ-invariant positive definite kernel can be uniformly approximated by kernels of this form. Since this thesis focuses on explicit computations of semidefinite programming bounds, we do not consider the topic of uniform density of such kernels in the cone of Γ-invariant positive definite kernels. For conditions under which this is true, see, e.g., [**85, 113**].

To construct a family $\{\psi_{\pi,l}\}$, we take the following approach. Intuitively, the space V is close to isomorphic to $V/\Gamma \times \Gamma/\mathrm{Stab}_\Gamma(V)$: every element $v \in V$ can be written as γR where R is a representative of the orbit of v, for some $\gamma \in \Gamma$, which is unique up to elements in the stabilizer subgroup of R. The idea then is to identify $C(V)$ with $C(V/\Gamma) \otimes C(\Gamma/\mathrm{Stab}_\Gamma(V))$: the first factor is Γ-invariant by definition, so the harmonic analysis only concerns the second factor. For the second factor, consider a representation (π, W) of Γ. Since $C(\Gamma/\mathrm{Stab}_\Gamma(V))$ is in particular $\mathrm{Stab}_\Gamma(V)$-invariant, we only need to consider the subspace $W^{\mathrm{Stab}_\Gamma(V)}$ of W. In general, V is not exactly isomorphic to $V/\Gamma \times \Gamma/\mathrm{Stab}_\Gamma(V)$: some orbits can have a larger stabilizer subgroup. These discrepancies need to be corrected, which happens

in Section 3.4 by ensuring that the certain functions $\psi_{\pi,l}$ evaluate to 0 on such orbits.

3.4. An explicit construction on the sphere[†]

In this section, we consider an explicit construction for a family of equivariant functions in the context of the Lasserre hierarchy on the sphere. We will use this in Part II to compute the second level of the Lasserre hierarchy on the sphere for several problems.

Let $\varepsilon > 0$, and consider the graph with vertex set S^{n-1}, where distinct vertices $x, y \in S^{n-1}$ are adjacent if $d(x, y) < \varepsilon$. Here $d(x, y) = \sqrt{2 - 2\langle x, y\rangle}$ is the chordal distance. Take $V = \mathcal{I}_2$, the independent sets in this graph of size at most 2. Then V is invariant under $\Gamma = O(n)$, as is the Lasserre hierarchy for the problems we consider. This means that we need $O(n)$-invariant positive definite kernels on $\mathcal{I}_2$ to perform computations.

To describe these kernels, we use the method outlined at the end of Section 3.3. This requires in particular the construction of $O(n-k)$-invariant subspaces of representations of $O(n)$. For this we use the construction by Gross and Kunze [**63**], who describe the spaces $W^{O(n-k)}$ induced by representations of $\mathrm{GL}(k)$. The space $\mathcal{I}_{=2}$ has an additional S_2 symmetry, and for the orbit containing sets of the form $\{x, -x\}$ for some $x \in S^{n-1}$, the stabilizer group is $O(n-1)$ instead of $O(n-2)$. In Section 3.4.3, we give a subspace of $W^{O(n-k)}$ and an additional factor such that the resulting zonal matrices are given by polynomials in the inner products of vectors on the sphere, which we prove in Section 3.4.4.

3.4.1. Representations of the general linear group. We start by briefly recalling some facts about the representations of the general linear group, which may be found, e.g., in [**57**, Chapter 15]. The irreducible representations of $\mathrm{GL}(t)$ are indexed by their signature $\lambda = (\lambda_1, \ldots, \lambda_t)$, which is a tuple of integers satisfying $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_t$. The polynomial, irreducible representations are those with $\lambda_t \geq 0$.

For kernels on $\mathcal{I}_2$, we require an explicit description of the irreducible, polynomial representations of $\mathrm{GL}(t)$ for $t = 2$. They are given by

$$W = \mathrm{Sym}^{\lambda_2}(\wedge^2 U) \otimes \mathrm{Sym}^m(U),$$

where $U = \mathbb{C}^2$ is the tautological representation that sends a matrix to itself with basis e_1, e_2, the signature $\lambda = (\lambda_1, \lambda_2)$ satisfies $\lambda_1 \geq \lambda_2 \geq 0$, and we define $m = \lambda_1 - \lambda_2$. We denote the corresponding group homomorphism by $\rho\colon \mathrm{GL}(2) \to \mathrm{GL}(W)$. A basis of this representation is given by

$$w_k = (e_1 \wedge e_2)^{\lambda_2} e_1^{m-k} e_2^k,$$

where $k = 0, 1, \ldots, m$. We give W the inner product such that $\langle w_{k_1}, w_{k_2}\rangle = \delta_{k_1,k_2}$. With this choice, we have

$$\langle w_{k_1}, \rho(A) w_{k_2}\rangle = \det(A)^{\lambda_2} \sum_{l=0}^{m-k_1} \binom{m-k_2}{l}\binom{k_2}{m-k_1-l} A_{11}^l A_{21}^{m-k_2-l} A_{12}^{m-k_1-l} A_{22}^{k_2-(m-k_1-l)}. \tag{3.4.1}$$

[†]This section is taken from the publication "*D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer,* Optimality and uniqueness of the D_4 root system*, 2024, arXiv:2404.18794*"

For brevity, we shall use the notation $\rho(A)_{k_1,k_2} = \langle w_{k_1}, \rho(A)w_{k_2}\rangle$. Let $c_j(k)$ denote the number of times e_j occurs in the tensor w_k. Concretely, we have $c_1(k) = \lambda_2 + m - k$ and $c_2(k) = \lambda_2 + k$. For a diagonal matrix D, we have

$$\rho(D)w_k = D_{11}^{c_1(k)} D_{22}^{c_2(k)} w_k.$$

We will occasionally refer to the representation as ρ_λ when it is convenient to make the dependence on λ explicit.

For later use, we also record here a formula for the differential $d\rho$ at the identity I evaluated at

$$X = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Using the product rule (see, e.g., [**57**, Chapter 8]), we obtain

$$d\rho(X)w_{k_2} = -(m-k_2)w_{k_2+1} + k_2 w_{k_2-1}$$

and hence $d\rho(X)_{k_1,k_2} = -(m-k_2)\delta_{k_1,k_2+1} + k_2\delta_{k_1,k_2-1}$.

3.4.2. Invariants of the orthogonal group. Let $n \geq 2t$. Denote by $O(n,K)$ the group of $n \times n$ matrices g with entries in the field K satisfying $g^{\mathsf{T}}g = I$. We see the group $O(n-t,K)$ as the subgroup of $O(n,K)$ which fixes the first t standard basis vectors. We will denote $O(n,\mathbb{R})$ by $O(n)$.

Following Gross and Kunze [**63**], we now define certain representations of $O(n)$ induced by representations of $\mathrm{GL}(t)$. Let (ρ, W) be the polynomial, irreducible representation of $\mathrm{GL}(t)$ with signature λ. Define the complex $t \times n$ matrix

$$\omega = \begin{bmatrix} I_t & iI_t & 0 \end{bmatrix}$$

and the $n \times t$ matrix

$$\epsilon = \begin{bmatrix} I_t \\ 0 \end{bmatrix}.$$

For each $w \in W$, define a function $f_w \colon O(n,\mathbb{C}) \to W$ by

$$f_w(\gamma) = \rho(\omega\gamma\epsilon)w. \tag{3.4.2}$$

Define the vector space of right translates of such functions by

$$V = \mathrm{span}\left\{R_g f_w \mid g \in O(n,\mathbb{C}),\, w \in W\right\},$$

where $R_g f_w(\gamma) = f_w(\gamma g)$. This space is a representation of $O(n,\mathbb{C})$ by right translation. A representation of $O(n)$ is obtained by restricting $O(n,\mathbb{C})$ to $O(n)$. We shall refer to this representation of $O(n)$ by (π, V).

Let $\Psi \colon W \to V$ be the map sending w to f_w, and consider the space of invariants

$$V^{O(n-t)} = \{v \in V \mid \pi(h)v = v \text{ for all } h \in O(n-t)\}.$$

Since $h\epsilon = \epsilon$ for $h \in O(n-t)$, we have $\Psi(W) \subseteq V^{O(n-t)}$.

On V, we define the inner product

$$\langle f_1, f_2\rangle = \int_{O(n)} \langle f_1(\gamma), f_2(\gamma)\rangle \, d\gamma.$$

By standard properties of the Haar measure, this makes V a unitary representation of $O(n)$. It may be shown that with the inner product chosen in Section 3.4.1, the numbers

$$\langle \Psi(w_i), \pi(g)\Psi(w_j)\rangle = \int_{O(n)} \langle \Psi(w_i)(\gamma), \Psi(w_j)(\gamma g)\rangle \, d\gamma$$

are real; see [**90**, Section 3].

In this thesis, it is only required that V is a representation of the orthogonal group and $\Psi(W) \subseteq V^{O(n-t)}$. However, it follows from the results in [**63**] that the above description is complete in the following sense. For $n > 2t$, the representations of $O(n)$ defined above are irreducible and we have equality $\Psi(W) = V^{O(n-t)}$. Moreover, all irreducible representations of $O(n)$ with nontrivial invariants under $O(n-t)$ are of this form for a unique λ. For $n = 2t$, a complete characterization of the irreducible representations and invariants is given in [**63**, Section 8], and using this it may be shown that the description of kernels in our approach is also complete in the case $n = 2t$. The exact statement and proof can be found in the PhD thesis [**113**].

3.4.3. Equivariant functions. In this section, we define a family of $O(n)$-equivariant functions from $\mathcal{I}_2$ to the representation V as constructed in Section 3.4.2. The definition of these functions depends on the choice of representatives of the orbits of $\mathcal{I}_2$ under the action of $O(n)$. Let

$$\begin{aligned} p_j(\{x,y\}) &= \langle x,y\rangle^j, \\ q_1(\{x,y\}) &= \sqrt{2+2\langle x,y\rangle}, \\ q_2(\{x,y\}) &= \sqrt{2-2\langle x,y\rangle}. \end{aligned}$$

For the orbit $\mathcal{I}_{=0}$ the representative is $\emptyset$ and for the orbit $O(n)J$ with $|J| \geq 1$ we choose the representative

$$\left\{ \left(\frac{q_1(J)}{2}, \frac{q_2(J)}{2}, 0, \ldots, 0 \right), \left(\frac{q_1(J)}{2}, -\frac{q_2(J)}{2}, 0, \ldots, 0 \right) \right\}. \tag{3.4.3}$$

In particular, this means that the standard basis vector e_1 is the representative for the orbit $\mathcal{I}_{=1}$.

The equivariant functions will be indexed by so-called admissible tuples. If $i = 0$, we call the tuple (λ, i, j, k) admissible if $\lambda = (0,0)$, $j = 0$, and $k = 0$. If $i = 1$, we call the tuple admissible if $\lambda_2 = 0$, $j = 0$, and $k = 0$. Finally, if $i = 2$, we call the tuple admissible for any $\lambda_1 \geq \lambda_2 \geq 0$, $j \geq 0$, and $0 \leq k \leq \lambda_1 - \lambda_2$ with $\lambda_2 + k$ even.

For each admissible tuple (λ, i, j, k), we now define the function

$$\psi_{\lambda,(i,j,k)}(J) = \xi_{\lambda,i,j,k}(J)\pi(s(J))\Psi(w_k),$$

where

$$\xi_{\lambda,i,j,k}(J) = \begin{cases} 1 & \text{if } i = |J| < 2, \\ p_j(J)q_1(J)^{c_1(k)}q_2(J)^{c_2(k)} & \text{if } i = |J| = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Here $s\colon \mathcal{I}_2 \to O(n)$ is a function such that $s(J)R = J$, where R is the orbit representative of the orbit $O(n)J$. To such a function s we shall refer as a section. Once the orbit representatives are fixed, the construction of the functions does not depend on the choice of the section s.

Let us give a brief motivation for these formulae. Firstly, the subscript i indicates the connected component $\mathcal{I}_{=i}$ on which the equivariant function is not identically zero. The space $\mathcal{I}_{=i}$ is homeomorphic to a quotient of $\mathcal{I}_{=i}/O(n) \times O(n)/O(n-i)$. For $i = 2$, the first factor is homeomorphic to an interval and the second factor to a Stiefel manifold. The function p_j may be viewed as a function on the factor $\mathcal{I}_{=2}/O(n)$ and $\pi(s(J))\Psi(w_k)$ as a function on the factor $O(n)/O(n-2)$. These functions are then multiplied to obtain functions on the whole space. The functions

q_1 and q_2 serve two purposes. Namely, they will ensure that we have compatibility with the additional quotient concerning the endpoints of $\mathcal{I}_{=2}/O(n)$, and that we obtain polynomial expressions.

LEMMA 3.5. *For admissible (λ, i, j, k), the function $\psi_{\lambda,(i,j,k)}$ is equivariant.*

PROOF. Since $\psi_{\lambda,(i,j,k)}$ is supported on $\mathcal{I}_{=i}$, and since the action of $O(n)$ on $\mathcal{I}_2$ preserves the cardinality of the sets, we only need to show equivariance for the restriction of $\psi_{\lambda,(i,j,k)}$ to $\mathcal{I}_{=i}$. Let J be an element in $\mathcal{I}_{=i}$ and let R be the orbit representative of $O(n)J$. For $g \in O(n)$, we have $s(gJ)R = gJ$ and $gs(J)R = gJ$, so $s(gJ) = gs(J)h$ for some h in the stabilizer subgroup $\mathrm{Stab}_{O(n)}(R)$. Hence,

$$\begin{aligned}\psi_{\lambda,(i,j,k)}(gJ) &= \xi_{\lambda,i,j,k}(gJ)\pi(s(gJ))\Psi(w_k)\\ &= \xi_{\lambda,i,j,k}(J)\pi(gs(J)h)\Psi(w_k)\\ &= \xi_{\lambda,i,j,k}(J)\pi(g)\pi(s(J))\pi(h)\Psi(w_k).\end{aligned}$$

We will complete the proof by showing that unless $\psi_{\lambda,(i,j,k)}(J)$ and $\psi_{\lambda,(i,j,k)}(gJ)$ are both zero, $\pi(h)\Psi(w_k) = \Psi(w_k)$, which shows

$$\psi_{\lambda,(i,j,k)}(gJ) = \pi(g)\psi_{\lambda,(i,j,k)}(J).$$

For this, we consider the cases $i = 0, 1, 2$ separately. The $i = 0$ case is immediate since $\mathcal{I}_{=0}$ consists of a single element, and since $\lambda = 0$, V is one dimensional. If $i = 1$, then $k = 0$, and the stabilizer subgroup of $O(n)$ with respect to R is $O(n-1)$. By formula (3.4.1), the dependence of $\rho(\omega\gamma h\epsilon)w_0$ on $\omega\gamma h\epsilon$ is only in the first column, which is equal to the first column of $\omega\gamma\epsilon$, so

$$\pi(h)\Psi(w_0)(\gamma) = \rho(\omega\gamma h\epsilon)w_0 = \rho(\omega\gamma\epsilon)w_0 = \Psi(w_0)(\gamma).$$

If $i = 2$ and the points in J are not antipodal, then the stabilizer subgroup of $O(n)$ with respect to R is $S_2 \times O(n-2)$, where S_2 is the two-element group generated by the matrix r which maps e_2 to $-e_2$ and fixes the orthogonal complement of e_2. By construction (see Section 3.4.2), we have $\pi(h)\Psi(w_k) = \Psi(w_k)$ for $h \in O(n-2)$. The matrix $\omega\gamma r\epsilon$ is the same as $\omega\gamma\epsilon$, except that the second column gets multiplied by -1. Since $\lambda_2 + k$ is even, it follows again from formula (3.4.1) that $\rho(\omega\gamma r\epsilon)w_k = \rho(\omega\gamma\epsilon)w_k$, and thus that $\pi(r)\Psi(w_k)(\gamma) = \Psi(w_k)(\gamma)$ holds.

If $i = 2$ and the points in J are antipodal, then the stabilizer subgroup is $O(n-1)$ and $q_1(J) = 0$. If $c_1(k) > 0$, then $q_1(J)^{c_1(k)} = 0$, so both $\psi_{\lambda,(i,j,k)}(J)$ and $\psi_{\lambda,(i,j,k)}(gJ)$ are zero. If $c_1(k) = 0$, then $\lambda_2 = 0$ and $k = \lambda_1$, and according to (3.4.1), $\rho(\omega\gamma h\epsilon)w_k$ only depends on the second column of $\omega\gamma h\epsilon$, which is equal to the second column of $\omega\gamma\epsilon$, so

$$\pi(h)\Psi(w_k)(\gamma) = \rho(\omega\gamma h\epsilon)w_k = \rho(\omega\gamma\epsilon)w_k = \Psi(w_k)(\gamma). \qquad \square$$

3.4.4. Zonal matrices. We now define the zonal matrix Z_λ by

$$Z_\lambda(J_1, J_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)} = \langle \psi_{\lambda,(i_1,j_1,k_1)}(J_1), \psi_{\lambda,(i_2,j_2,k_2)}(J_2)\rangle,$$

where the rows and columns range over all admissible tuples. It follows from equivariance of the function $\psi_{\lambda,(i,j,k)}$ and unitarity of the inner product that the zonal matrices are $O(n)$ invariant.

In the remainder of this section, we use invariant theory to give a short argument showing that the entries of the zonal matrices are polynomials in the inner products between the vectors in $J_1 \cup J_2$. Note that this fact also follows from the direct construction in terms of inner products as given in Section 3.4.5.

LEMMA 3.6. *Let (λ, i, j, k) be admissible. For fixed $w \in W$, the expression*

$$\langle w, \psi_{\lambda,(i,j,k)}(\{x_1, \dots, x_i\})(\gamma)\rangle$$

is a polynomial in the entries of the orthogonal matrix γ and the vectors $x_1, \dots, x_i$.

PROOF. Given the choice of representatives, we have that for $J = \{x_1\}$, the first column of $s(J)$ is equal to x_1, for $J = \{x_1, x_2\}$ with $\langle x_1, x_2\rangle \neq \pm 1$, the first column of $s(J)$ is $(x_1 + x_2)/q_1(J)$ and the second column is either $(x_1 - x_2)/q_2(J)$ or $(x_2 - x_1)/q_2(J)$, and for $J = \{x_1, -x_1\}$ the second column of $s(J)$ is either x_1 or $-x_1$. By the choice of admissible tuples, it will turn out that the resulting expressions do not depend on the sign of the second column.

We prove the lemma for each i separately. For $i = 0$, the expression is a constant. If $i = 1$, then $\lambda_2 = j = k = 0$, and we have

$$\langle w, \psi_{\lambda,(1,0,0)}(\{x_1\})\rangle = \xi_{\lambda,1,0,0}(\{x_1\})\langle w, \pi(s(\{x_1\}))\Psi(w_k)(\gamma)\rangle.$$

Here $\xi_{\lambda,1,0,0}(\{x_1\}) = 1$ and

$$\langle w, \pi(s(\{x_1\}))\Psi(w_0)(\gamma)\rangle = \langle w, \rho(\omega\gamma s(\{x_1\})\epsilon)w_0\rangle.$$

From the expression (3.4.1) for the matrix coefficients of ρ, it follows that the right-hand side is a polynomial in the entries in the first column of $\omega\gamma s(\{x_1\})\epsilon$, which is a polynomial in the entries of γ and x_1.

Now let $i = 2$ and set $J = \{x_1, x_2\}$. We will show that

$$\psi_{\lambda,(i,j,k)}(J) = \langle x_1, x_2\rangle^j \rho(\omega\gamma\begin{bmatrix} x_1 + x_2 & x_1 - x_2\end{bmatrix})w_k. \tag{3.4.4}$$

For $\langle x_1, x_2\rangle \neq \pm 1$, we then have

$$\begin{aligned}\psi_{\lambda,(i,j,k)}(J) &= \xi_{\lambda,i,j,k}(J)\pi(s(J))\Psi(w_k)\\ &= p_j(J)q_1(J)^{c_1(k)}q_2(J)^{c_2(k)}\rho\Big(\omega\gamma\begin{bmatrix}\frac{x_1+x_2}{q_1(J)} & \frac{x_1-x_2}{q_2(J)}\end{bmatrix}\Big)w_k.\end{aligned}$$

Here we used that the expression does not depend on the sign of the second column since $\lambda_2 + k$ is even, i.e., we have

$$\rho\big(\omega\gamma\begin{bmatrix}u & v\end{bmatrix}\big)w_k = \rho\big(\omega\gamma\begin{bmatrix}u & -v\end{bmatrix}\big)w_k$$

for all orthonormal u and v. Since

$$\rho\Big(\omega\gamma\begin{bmatrix}\frac{x_1+x_2}{q_1(J)} & \frac{x_1-x_2}{q_2(J)}\end{bmatrix}\Big) = \rho(\omega\gamma\begin{bmatrix}x_1 + x_2 & x_1 - x_2\end{bmatrix})\rho\left(\begin{bmatrix}1/q_1(J) & 0\\ 0 & 1/q_2(J)\end{bmatrix}\right),$$

it follows that identity (3.4.4) holds whenever $\langle x_1, x_2\rangle \neq \pm 1$.

We will show (3.4.4) also holds for the case $x_1 = -x_2$. We have

$$\begin{aligned}\psi_{\lambda,(i,j,k)}(J) &= \xi_{\lambda,i,j,k}(J)\pi(s(J))\Psi(w_k)\\ &= p_j(J)q_1(J)^{c_1(k)}q_2(J)^{c_2(k)}\rho(\omega\gamma s(J)\epsilon)w_k.\end{aligned}$$

We may now substitute $s(J)\epsilon$ with $\begin{bmatrix}c & x_1\end{bmatrix}$ for any unit vector c orthogonal to x_1, to obtain

$$\begin{aligned}\psi_{\lambda,(i,j,k)}(J) &= \langle x_1, x_2\rangle^j 0^{c_1(k)}2^{c_2(k)}\rho\big(\omega g\begin{bmatrix}c & x_1\end{bmatrix}\big)w_k\\ &= \langle x_1, x_2\rangle^j \rho\big(\omega\gamma\begin{bmatrix}c & x_1\end{bmatrix}\big)\rho\left(\begin{bmatrix}0 & 0\\ 0 & 2\end{bmatrix}\right)w_k\\ &= \langle x_1, x_2\rangle^j \rho(\omega\gamma\begin{bmatrix}x_1 + x_2 & x_1 - x_2\end{bmatrix})w_k.\end{aligned}$$

A similar argument can be used to show (3.4.4) holds for the case $x_1 = x_2$. Together this shows

$$\langle w, \psi_{\lambda,(i,j,k)}(\{x_1, x_2\})(\gamma)\rangle$$

is a polynomial in the entries of γ, x_1, and x_2. □

PROPOSITION 3.7. *Fix i_1 and i_2 and let $J_1 = \{x_1, \dots, x_{i_1}\}$ and $J_2 = \{y_1, \dots, y_{i_2}\}$. For admissible tuples (λ, i_1, j_1, k_1) and (λ, i_2, j_2, k_2),*

$$Z_\lambda(J_1, J_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)}$$

is a polynomial in the inner products between the vectors $x_1, \dots, x_{i_1}, y_1, \dots, y_{i_2}$.

PROOF. By the definition of the inner product on V we have

$$Z_\lambda(J_1, J_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)} = \int_{O(n)} \langle \psi_{\lambda,(i_1,j_1,k_1)}(J_1)(\gamma), \psi_{\lambda,(i_2,j_2,k_2)}(J_2)(\gamma)\rangle \, d\gamma.$$

Since the vectors $w_0, \dots, w_{\lambda_1-\lambda_2}$ form an orthonormal basis of W, this is equal to

$$\int_{O(n)} \sum_{l=0}^{\lambda_1-\lambda_2} \langle \psi_{\lambda,(i_1,j_1,k_1)}(J_1)(\gamma), w_l\rangle \langle w_l, \psi_{\lambda,(i_2,j_2,k_2)}(J_2)(\gamma)\rangle \, d\gamma.$$

By Lemma 3.6, this is a polynomial in the entries of the vectors $x_1, \dots, x_{i_1}, y_1, \dots, y_{i_2}$.

By Lemma 3.5, the functions $\psi_{\lambda,(i_1,j_1,k_1)}$ and $\psi_{\lambda,(i_2,j_2,k_2)}$ are equivariant, so by unitarity of the inner product on V it follows that

$$Z_\lambda(gJ_1, gJ_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)} = Z_\lambda(J_1, J_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)}$$

for all $g \in O(n)$. In other words, this is an $O(n)$-invariant polynomial in the vectors $x_1, \dots, x_{i_1}, y_1, \dots, y_{i_2}$. By invariant theory (see, e.g., [**57**, §F.1]), it follows that $Z_\lambda(J_1, J_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)}$ is a polynomial in the inner products between these vectors. □

3.4.5. Efficient computation of the zonal matrices. In this section, we explain how we compute the zonal matrices from Section 3.4.4. Throughout we assume $t = 2$, but we will sometimes write t instead of 2 to make explicit the dependence on t. Compared to the construction of the zonal matrices in [**90**], we give a much more efficient approach, which is crucial to be able to perform computations with the truncation degree required to get a sharp bound for the D_4 root system.

We have

$$\begin{aligned}
&Z_\lambda(J_1, J_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)} \\
&\quad = \int_{O(n)} \langle \psi_{\lambda,(i_1,j_1,k_1)}(J_1)(\gamma), \psi_{\lambda,(i_2,j_2,k_2)}(J_2)(\gamma)\rangle \, d\gamma \\
&\quad = \xi_{\lambda,i_1,j_1,k_1}(J_1)\xi_{\lambda,i_2,j_2,k_2}(J_2) \int_{O(n)} \langle \rho(\omega\gamma s(J_1)\epsilon)w_{k_1}, \rho(\omega\gamma s(J_2)\epsilon)w_{k_2}\rangle \, d\gamma \\
&\quad = \xi_{\lambda,i_1,j_1,k_1}(J_1)\xi_{\lambda,i_2,j_2,k_2}(J_2)P_{k_1,k_2}(s(J_1)^{\mathsf{T}}s(J_2)),
\end{aligned}$$

where we define

$$P_{k_1,k_2}(S) = \int_{O(n)} \langle \rho(\omega\gamma\epsilon)w_{k_1}, \rho(\omega\gamma S\epsilon)w_{k_2}\rangle \, d\gamma. \tag{3.4.5}$$

Here $P_{k_1,k_2}(S)$ is a polynomial in the entries of the $n \times n$ matrix S. In this section we show how to compute $P_{k_1,k_2}(S)$ efficiently and how to use this to obtain $Z_\lambda(J_1, J_2)$ as a polynomial in the inner products.

3.4.5.1. *Additional symmetries.* To compute $P_{k_1,k_2}(S)$ using the matrix entries of ρ, one could directly use the expression

$$P_{k_1,k_2}(S) = \sum_{l=0}^{m} \int_{O(n)} \langle \rho(\omega\gamma\epsilon) w_{k_1}, w_l \rangle \langle w_l, \rho(\omega\gamma S\epsilon) w_{k_2} \rangle \, d\gamma,$$

where $m = \lambda_1 - \lambda_2$. In this section, we describe additional symmetries under the action of the circle group $O(2)$, which allows us to compute this more efficiently.

Denote by ρ_λ the representation of GL(2) with signature λ. For any matrix M, we have $\rho_\lambda(M) = \det(M)^{\lambda_2} \rho_{(m,0)}(M)$, and hence

$$\begin{aligned} P_{k_1,k_2}(S) &= \int_{O(n)} \langle \rho_\lambda(\omega\gamma\epsilon) w_{k_1}, \rho_\lambda(\omega\gamma S\epsilon) w_{k_2} \rangle \, d\gamma \\ &= \int_{O(n)} \det(\overline{\omega\gamma\epsilon})^{\lambda_2} \langle \rho_{(m,0)}(\omega\gamma\epsilon) w_{k_1}, \rho_{(m,0)}(\omega\gamma S\epsilon) w_{k_2} \rangle \det(\omega\gamma S\epsilon)^{\lambda_2} \, d\gamma, \end{aligned}$$

where $\overline{A}$ denotes the entrywise complex conjugate of A. We introduce some notation to conveniently describe and manipulate expressions such as the one above. We will refer to the representation $\rho_{(m,0)}$ as ρ in this section. Let $\alpha = (\alpha_1, \ldots, \alpha_{\lambda_2})$ be a vector with

$$\alpha_i = (\alpha_{i1}, \alpha_{i2}) \in \{(1,1), (1,2), (2,1), (2,2)\},$$

and let e be the vector with $e_i = (1,2)$ for all i. We also define the matrices $A = \omega\gamma\epsilon$ and $B = \omega\gamma S\epsilon$. Denote by $[\alpha]$ the orbit of α under the action of the symmetric group S_{λ_2} on the λ_2 components. For such α and $0 \le l_1, l_2 \le m$ we consider the following polynomial in the entries of S:

$$J_{l_1,l_2,[\alpha]} = \int_{O(n)} \det(\bar{A})^{\lambda_2} \rho(A)^*_{k_1,l_1} \rho(B)_{l_2,k_2} \prod_{i=1}^{\lambda_2} B_{\alpha_{i1},1} B_{\alpha_{i2},2} \, d\gamma, \tag{3.4.6}$$

where ρ^* is the adjoint of ρ.

For each signature λ that we need and each $0 \le k_1, k_2 \le m$, we will show there are coefficients $c_{l_1,k_2,[\sigma]}$, independent of S and k_1, such that

$$J_{l_1,l_1,[\sigma]} = c_{l_1,k_2,[\sigma]} J_{0,0,[e]} \tag{3.4.7}$$

for all $0 \le l_1 \le m$ and $\sigma \in \{(1,2),(2,1)\}^{\lambda_2}$. We will compute these coefficients by solving a linear system for each λ and k_2. By expanding both the inner product and the determinant involving B, the polynomial $P_{k_1,k_2}(S)$ may then be computed as

$$P_{k_1,k_2}(S) = \sum_{l_1,\sigma} (-1)^{s(\sigma)} J_{l_1,l_1,[\sigma]},$$

where the sum is over $0 \le l_1 \le m$ and all tuples $\sigma \in \{(1,2),(2,1)\}^{\lambda_2}$, and $s(\sigma)$ is the number of times the pair $(2,1)$ occurs in σ. By grouping terms and using (3.4.7) we can write this as

$$P_{k_1,k_2}(S) = J_{0,0,[e]} \sum_{l_1,[\sigma]} (-1)^{s(\sigma)} \binom{\lambda_2}{s(\sigma)} c_{l_1,k_2,[\sigma]}.$$

In summary, for fixed λ, k_1 and k_2, we need to compute only one integral of the form (3.4.6) using this approach.

We now show how to compute these coefficients. For $g \in O(t)$, we may substitute γ with $(g \oplus g \oplus I_{n-2t})\gamma$, and this leaves the expression $J_{l_1,l_2,[\alpha]}$ invariant by the invariance property of the Haar measure of $O(n)$. We have $\omega(g \oplus g \oplus I_{n-2t})\gamma = g\omega\gamma$ and hence we may substitute gA for A and gB for B. This gives

$$J_{l_1,l_2,[\alpha]} = \sum_{l_3,l_4,[\beta]} \det(\bar{g})^{\lambda_2} \rho(\bar{g})_{l_1,l_3} J_{l_3,l_4,[\beta]} \rho(g)_{l_2,l_4} \sum_{\zeta \in [\beta]} \prod_{i=1}^{\lambda_2} g_{\alpha_{i1},\zeta_{i1}} g_{\alpha_{i2},\zeta_{i2}}. \tag{3.4.8}$$

We now phrase this in terms of a representation.

Recall that the representation $\mathrm{Sym}^{\lambda_2}(\wedge^2 U) \cong \mathbb{C}$ is given by multiplication by $\det(g)^{\lambda_2}$. Also recall the representation on $\mathrm{End}(W)$ given by

$$g \cdot M = \rho(g) M \rho(g)^*.$$

For this representation, a basis is given by $w_{l_1} \otimes w_{l_2}^*$. Finally, let $U = \mathbb{C}^2$ be the representation with the standard action of $O(2)$ and consider the representation $(\phi, \mathrm{Sym}^{\lambda_2}(U^{\otimes 2}))$. The vectors

$$e_{[\beta]} = \prod_{i=1}^{\lambda_2} e_{\beta_{i1}} \otimes e_{\beta_{i2}}$$

form a basis. We consider the inner product such that this basis is orthonormal. We then consider the dual representation $\phi(g^*)^*$. We have

$$\phi(g^*) e_{[\alpha]} = \prod_{i=1}^{\lambda_2} g^* e_{\alpha_{i1}} \otimes g^* e_{\alpha_{i2}} = \sum_{[\beta]} \sum_{\zeta \in [\beta]} \prod_{i=1}^{\lambda_2} g_{\alpha_{i1},\zeta_{i1}} g_{\alpha_{i2},\zeta_{i2}} e_\beta$$

and hence

$$\langle e_\alpha, \phi(g^*)^* e_\beta \rangle = \langle \phi(g^*) e_\alpha, e_\beta \rangle = \sum_{\zeta \in [\beta]} \prod_{i=1}^{\lambda_2} g_{\alpha_{i1},\zeta_{i1}} g_{\alpha_{i2},\zeta_{i2}}.$$

Tensoring the above representations gives the representation

$$(\Phi, \mathrm{Sym}^{\lambda_2}(\wedge^2 U) \otimes \mathrm{End}(W) \otimes \mathrm{Sym}^{\lambda_2}(U^{\otimes 2}))$$

and a basis is given by $e_{l_1,l_2,[\alpha]} = w_{l_1} \otimes w_{l_2}^* \otimes e_{[\alpha]}$. We get

$$\langle e_{l_1,l_2,[\alpha]}, \Phi(g) e_{l_3,l_4,[\beta]} \rangle = \det(g)^{\lambda_2} \rho(g)_{l_1,l_3} \rho(g)_{l_2,l_4} \sum_{\zeta \in [\beta]} \prod_{i=1}^{\lambda_2} g_{\alpha_{i1},\zeta_{i1}} g_{\alpha_{i2},\zeta_{i2}}. \tag{3.4.9}$$

Using (3.4.8) and (3.4.9) we have

$$\begin{aligned}
&\Phi(g) \sum_{l_3,l_4,[\beta]} J_{l_3,l_4,[\beta]} e_{l_3,l_3,[\beta]} \\
&= \sum_{l_1,l_2,[\alpha]} \sum_{l_3,l_4,[\beta]} J_{l_3,l_4,[\beta]} \langle e_{l_1,l_2,[\alpha]}, \Phi(g) e_{l_3,l_4,[\beta]} \rangle e_{l_1,l_2,[\alpha]} \\
&= \sum_{l_1,l_2,[\alpha]} \sum_{l_3,l_4,[\beta]} \det(\bar{g})^{\lambda_2} \rho(\bar{g})_{l_1,l_3} J_{l_3,l_4,[\beta]} \rho(g)_{l_2,l_4} \sum_{\zeta \in [\beta]} \prod_{i=1}^{\lambda_2} g_{\alpha_{i1},\zeta_{i1}} g_{\alpha_{i2},\zeta_{i2}} e_{l_1,l_2,[\alpha]} \\
&= \sum_{l_1,l_2,[\alpha]} J_{l_1,l_2,[\alpha]} e_{l_1,l_2,[\alpha]}.
\end{aligned}$$

Defining

$$J = \sum_{l_1,l_2,[\alpha]} J_{l_1,l_2,\alpha} e_{l_1,l_2,[\alpha]},$$

this equation is expressed as $\Phi(g)J = J$ for all $g \in O(2)$. Using the exponential map, this is equivalent to the condition $d\Phi(X)J = 0$, where

$$X = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and $\Phi(g_0)J = J$, where g_0 is an orthogonal matrix with $\det(g_0) = -1$. This follows from the fact that X spans the Lie algebra $\mathfrak{so}(2)$. The additional condition with g_0 comes from the fact that the equation has to hold for all orthogonal matrices and not merely for the special orthogonal matrices.

We now write out the system $d\Phi(X)J = 0$ in components. Let $g(t)$ be a curve of special orthogonal matrices such that $g(0) = I$ and $g'(0) = X$. To obtain the components of $d\Phi(X)$, we plug $g(t)$ into (3.4.9) and take the derivative. Using the product rule, one obtains

$$\begin{aligned}&\langle e_{l_1,l_2,[\alpha]}, d\Phi(X) e_{l_3,l_4,[\beta]}\rangle \\ &\quad = d\rho(X)_{l_1,l_3}\delta_{l_2,l_4}\delta_{[\alpha],[\beta]} + \delta_{l_1,l_3} d\rho(X)_{l_2,l_4}\delta_{[\alpha],[\beta]} + \delta_{l_1,l_3}\delta_{l_2,l_4} G'(0),\end{aligned}$$

where we have defined

$$G(t) = \sum_{\zeta \in [\beta]} \prod_{i=1}^{\lambda_2} g(t)_{\alpha_{i1},\zeta_{i1}} g(t)_{\alpha_{i2},\zeta_{i2}}.$$

A formula for $d\rho(X)$ can be found in Section 3.4.1. One may further verify that each term of $G'(0)$ is zero unless α_{ij} and ζ_{ij} differ for exactly one ij, in which case the term equals $X_{\alpha_{ij},\zeta_{ij}}$. Together this gives explicit formulas for the linear constraints on the coefficients $J_{l_1,l_2,[\alpha]}$ arising from $d\Phi(X)J = 0$.

We now work out the condition $\Phi(g_0)J = J$. For this, we let $d_j(\alpha)$ be the total number of occurrences of j in α. Recall the signature of ρ is $(m, 0)$, so that we have $c_1(l) = m - l$ and $c_2(l) = l$.

LEMMA 3.8. *If $\lambda_2 + c_2(l_1) + c_2(l_2) + d_2(\alpha)$ is odd, then $J_{l_1,l_2,[\alpha]} = 0$.*

PROOF. Let $g_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. We then have

$$\langle e_{l_1,l_2,[\alpha]}, \Phi(g_0) e_{l_3,l_4,[\beta]}\rangle = \delta_{l_1,l_3}\delta_{l_2,l_4}\delta_{[\alpha],[\beta]}(-1)^{\lambda_2 + c_2(l_1) + c_2(l_2) + d_2(\alpha)}$$

and hence from $J = \Phi(g_0)J$ we obtain

$$J_{l_1,l_2,[\alpha]} = (-1)^{\lambda_2 + c_2(l_1) + c_2(l_2) + d_2(\alpha)} J_{l_1,l_2,[\alpha]}. \qquad \square$$

We give additional conditions under which $J_{l_1,l_2,[\alpha]}$ vanishes.

LEMMA 3.9. *Let $j \in \{1, 2\}$. If $c_j(l_1) + \lambda_2 - (c_j(l_2) + d_j(\alpha)) \neq 0$, then $J_{l_1,l_2,[\alpha]} = 0$.*

PROOF. Let $R(\theta)$ be the matrix rotating the j and $j + t$ rows of γ by

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

We then have $\omega R(\theta) = A(\theta)\omega$, where $A(\theta)$ is the diagonal matrix with $e^{i\theta}$ at the jth diagonal entry and 1 at the other diagonal entry. The matrix $R(\theta)$ is orthogonal

and by a similar argument as before we may substitute ω with $A(\theta)\omega$. We then obtain that $J_{l_1,l_2,[\alpha]}$ is equal to

$$\sum_{l_3,l_4,[\beta]} \det(\overline{A(\theta)})^{\lambda_2} \rho(\overline{A(\theta)})_{l_1,l_3} J_{l_3,l_4,[\beta]} \rho(A(\theta))_{l_2,l_4} \prod_{i=1}^{\lambda_2} A(\theta)_{\alpha_{i1}\beta_{i1}} A(\theta)_{\alpha_{i2}\beta_{i2}}.$$

Working this out gives

$$J_{l_1,l_2,[\alpha]} = e^{-i\theta(c_j(l_1)+\lambda_2-(c_j(l_2)+d_j(\alpha)))} J_{l_1,l_2,[\alpha]}.$$

Since this equation holds for all θ, we have $J_{l_1,l_2,[\alpha]} = 0$. □

We thus have the system $d\Phi(X)J = 0$ and certain components of J vanish due to Lemmas 3.8 and 3.9. As a final step, which is necessary to ensure the solution space is one-dimensional, we add the following relations. By expanding into the monomials B_{11}, B_{12}, B_{21}, and B_{22}, there are coefficients $a_{l_2,k_2,[\alpha],\mu}$ such that

$$\rho(B)_{l_2,k_2} \prod_{i=1}^{\lambda_2} B_{\alpha_{i1},1} B_{\alpha_{i2},2} = \sum_{\mu} a_{l_2,k_2,[\alpha],\mu} B^{\mu}.$$

With

$$K_{l_1,\mu} = \int_{O(n)} \det(\bar{A})^{\lambda_2} \rho(A)^*_{k_1,l_1} B^{\mu} \, d\gamma$$

we have

$$J_{l_1,l_2,[\alpha]} = \sum_{\mu} a_{l_2,k_2,[\alpha],\mu} K_{l_1,\mu}.$$

We now enlarge the linear system by introducing new variables for the $K_{l_1,\mu}$, and for each l_1, l_2, and α we add the above constraint on the variables $J_{l_1,l_2,[\alpha]}$ and $K_{l_1,\mu}$.

Finally, we project the linear space satisfying all of the above relations to the space

$$\operatorname{span}\{e_{l_1,l_1,[\sigma]} \mid 0 \le l_1 \le m,\, \sigma \in \{(1,2),(2,1)\}^{\lambda_2}\}.$$

For this, we consider the homogeneous linear system given by the constraints discussed above. We order the columns so that the variables corresponding to $J_{l_1,l_1,[\sigma]}$ are at the end, and $J_{0,0,[e]}$ corresponds to the final column. Then we perform row reduction using rational arithmetic and find that the final column is the only free variable among the columns corresponding to the variables $J_{l_1,l_1,[\sigma]}$. From this, we find the coefficients $c_{l_1,k_2,[\sigma]}$ for which (3.4.7) holds.

3.4.5.2. *Real parts.* As shown in Section 3.4.5.1, to compute the zonal matrices we need to compute the quantity

$$J_{0,0,[e]} = \int_{O(n)} \det(\overline{\omega}\gamma\epsilon)^{\lambda_2} \rho(\omega\gamma\epsilon)^*_{k_1,0} \rho(\omega\gamma S\epsilon)_{0,k_2} ((\omega\gamma S\epsilon)_{1,1}(\omega\gamma S\epsilon)_{2,2})^{\lambda_2} \, d\gamma.$$

In this section we will show that for $\lambda_1 > 0$, this is equal to

(3.4.10)

$$2\int_{O(n)} \mathcal{R}\big(\det(\overline{\omega}\gamma\epsilon)^{\lambda_2} \rho(\omega\gamma\epsilon)^*_{k_1,0}\big) \mathcal{R}\big(\rho(\omega\gamma S\epsilon)_{0,k_2} ((\omega\gamma S\epsilon)_{1,1}(\omega\gamma S\epsilon)_{2,2})^{\lambda_2}\big) \, d\gamma,$$

where $\mathcal{R}(z)$ denotes the real part of $z \in \mathbb{C}$. This yields a factor two speedup in the most expensive part of the generation of the zonal matrices.

For a matrix M and a vector a of natural numbers of the same size, let us adopt the notation

$$M^a = \prod_{i,j} M_{i,j}^{a_{i,j}}.$$

By multilinearity and the formula for the matrix coefficients of the representations of GL(2), it suffices to show

$$\int_{O(n)} (\overline{\omega}\gamma\epsilon)^a (\omega\gamma S\epsilon)^b \, d\gamma = 2 \int_{O(n)} \mathcal{R}\left((\overline{\omega}\gamma\epsilon)^a\right) \mathcal{R}\left((\omega\gamma S\epsilon)^b\right) \, d\gamma \tag{3.4.11}$$

for all $a, b \in \mathbb{N}^{2\times 2}$ with $|a| = |b| = |\lambda|$.

To show this, we introduce the variables R_{11}, R_{12}, R_{21}, and R_{22}, the matrices

$$R_k^+ = \begin{bmatrix} R_{k1} I_2 & R_{k2} I_2 & 0 \end{bmatrix},$$

and the vectors $R_k = \begin{bmatrix} R_{k1} & R_{k2} \end{bmatrix}$ for $k = 1, 2$. We then consider the polynomial

$$\begin{aligned} \int_{O(n)} (R_1^+\gamma\epsilon)^a (R_2^+\gamma S\epsilon)^b \, d\gamma &= \sum_{|u|=|v|=|\lambda|} I_{u,v} R_1^u R_2^v \\ &= \sum_{|s|=2|\lambda|} \sum_{\substack{u+v=s \\ |u|=|v|=|\lambda|}} I_{u,v} R_{11}^{u_1} R_{12}^{u_2} R_{21}^{v_1} R_{22}^{v_2}, \end{aligned} \tag{3.4.12}$$

where the real numbers $I_{u,v}$ are obtained by working out brackets and gathering terms.

Substituting $R_{11} = 1$, $R_{12} = -i$, $R_{21} = 1$ and $R_{22} = i$ gives the left-hand side of (3.4.11), which is a real number by [**90**, Section 3]. So the sum over all terms with $u_2 + v_2$ odd vanishes. Since $|s|$ is even, $u_1 + v_1$ is restricted to be even too. We reparametrize the sum and obtain that the left-hand side of (3.4.11) is given by

$$\sum_{|s|=|\lambda|} (-1)^{s_2} \sum_{\substack{u+v=2s \\ |u|=|v|=|\lambda|}} (-1)^{u_2} I_{u,v}. \tag{3.4.13}$$

A similar reasoning shows that the right-hand side of (3.4.11) is equal to

$$2 \sum_{|s|=|\lambda|} (-1)^{s_2} \sum_{\substack{u+v=2s \\ u_2 \text{ even} \\ |u|=|v|=|\lambda|}} I_{u,v}.$$

We now substitute $R_1 = R_2$ in (3.4.12) to obtain the polynomial

$$\begin{aligned} \int_{O(n)} (R_1^+\gamma\epsilon)^a (R_1^+\gamma S\epsilon)^b \, d\gamma &= \sum_{|u|=|v|=|\lambda|} I_{u,v} R_1^u R_1^v \\ &= \sum_{|s|=2|\lambda|} \sum_{\substack{u+v=s \\ |u|=|v|=|\lambda|}} I_{u,v} R_1^s. \end{aligned} \tag{3.4.14}$$

Similarly as before, we may substitute γ with $(g \oplus g \oplus I_{n-2t})\,\gamma$, and this leaves the polynomial (3.4.14) invariant by the invariance property of the Haar measure of $O(n)$. We have $R_1^+ (g \oplus g \oplus I_{n-2t})\,\gamma = g R_1^+ \gamma$. Hence polynomial (3.4.14) is a polynomial in $R_{11}^2 + R_{12}^2$ by invariant theory. Since it is also a homogeneous polynomial of total degree $2|\lambda|$, it must be linearly proportional to the polynomial

$$\left(R_{11}^2 + R_{12}^2\right)^{|\lambda|}. \tag{3.4.15}$$

Hence the s which occur in the sum in (3.4.14) must have even entries, and the polynomial (3.4.14) may be written as

$$\sum_{|s|=|\lambda|} \sum_{\substack{u+v=2s \\ |u|=|v|=|\lambda|}} I_{u,v} R_1^{2s} = \sum_{|s|=|\lambda|} c_s R_1^{2s}.$$

Furthermore, since it must be linearly proportional to (3.4.15), we have

$$c_s = \binom{|\lambda|}{s_2} c_0$$

by the binomial theorem. We now rearrange terms to obtain

$$\begin{aligned}
\sum_{u+v=2s} (-1)^{u_2} I_{u,v} &= \sum_{\substack{u+v=2s \\ u_2 \text{ even}}} (-1)^{u_2} I_{u,v} + \sum_{\substack{u+v=2s \\ u_2 \text{ odd}}} (-1)^{u_2} I_{u,v} \\
&= 2 \sum_{\substack{u+v=2s \\ u_2 \text{ even}}} (-1)^{u_2} I_{u,v} - \sum_{u+v=2s} I_{u,v} \\
&= 2 \sum_{\substack{u+v=2s \\ u_2 \text{ even}}} (-1)^{u_2} I_{u,v} - c_s,
\end{aligned}$$

where in each sum we implicitly assume $|u| = |v| = |\lambda|$. Using (3.4.13), we now see that we may write the left-hand side of (3.4.11) as

$$\begin{aligned}
\sum_{|s|=|\lambda|} (-1)^{s_2} \sum_{u+v=2s} (-1)^{u_2} I_{u,v} &= 2 \sum_{|s|=|\lambda|} (-1)^{s_2} \sum_{\substack{u+v=2s \\ u_2 \text{ even}}} I_{u,v} - \sum_{|s|=|\lambda|} (-1)^{s_2} c_s \\
&= 2 \sum_{|s|=|\lambda|} (-1)^{s_2} \sum_{\substack{u+v=2s \\ u_2 \text{ even}}} I_{u,v},
\end{aligned}$$

since

$$\sum_{|s|=|\lambda|} (-1)^{s_2} c_s = c_0 \sum_{s_2=0}^{|\lambda|} \binom{|\lambda|}{s_2} (-1)^{s_2} = c_0 (1-1)^{|\lambda|} = 0$$

whenever $|\lambda| > 0$. Recall that the sum over even u_2 equals the integral of the product of the real parts. Hence we have shown equation (3.4.11), which is what we wanted to show.

3.4.5.3. *Inner products.* In this section, we describe how to compute

$$Z_\lambda(J_1, J_2)_{(i_1,j_1,k_1),(i_2,j_2,k_2)}$$

efficiently as a polynomial in the inner products between the vectors in $J_1 \cup J_2$.

We have

$$\rho(\omega\gamma h\epsilon) w_{k_1} = \rho(\omega\gamma\epsilon) w_{k_1}$$

for all $h \in O(n-2)$, where as before we view h as a matrix in $O(n)$ fixing the first 2 coordinates. Let P_{k_1,k_2} be the polynomial defined in (3.4.5). By the invariance property of the Haar measure, we have

$$P_{k_1,k_2}(hS) = P_{k_1,k_2}(S)$$

for all $h \in O(n-2)$. Let S_1 be the top-left 2×2 block of S and S_2 the bottom-left $(n-2) \times 2$ block of S. Using invariant theory (see, e.g., [**57**, §F.1]), we see that $P_{k_1,k_2}(S)$ must be a polynomial in the entries of S_1 and the inner products between the columns of S_2.

Consider the ideal $\mathcal{J}$ in $\mathbb{R}[S]$ generated by the entries of $S^\mathsf{T}S - I$ and the monomials $S_{i,j}$ for $i > 2 + j$. Consider the polynomials in $\mathcal{J}$ given by

$$S_{i,j} \text{ with } i > 2 + j \text{ and } j \leq 2, \quad 1 - \sum_{i=1}^{4} S_{i,2}^2, \quad \sum_{i=1}^{3} S_{i,1}S_{i,2}, \quad 1 - \sum_{i=1}^{3} S_{i,1}^2.$$

We now use the above polynomials to perform Euclidean division on P_{k_1,k_2} and show the remainder p_{k_1,k_2} is a polynomial in the entries of S_1. Here we use a lexicographical term order where the variables are ordered such that $S_{4,2} > S_{3,2} > S_{3,1}$ and $S_{i,j} < S_{3,1}$ if $i \leq 2$ and $S_{i,j} > S_{4,2}$ if $i > 2+j$. This results in the following concrete procedure. We first remove the terms in P_{k_1,k_2} that contain a variable $S_{i,j}$ with $i > 2 + j$, after which we obtain a polynomial of the form

$$\sum_{\alpha,a,b,c} C_{\alpha,a,b,c} S_1^\alpha S_{3,1}^{2a} (S_{3,1}S_{3,2})^b (S_{3,2}^2 + S_{4,2}^2)^c$$

for some C, α, a, b, and c. In this polynomial, we first replace every occurrence of $S_{4,2}^2$ with $1 - S_{1,2}^2 - S_{2,2}^2 - S_{3,2}^2$, then every occurrence of $S_{3,1}S_{3,2}$ with $-S_{1,1}S_{1,2} - S_{2,1}S_{2,2}$, and finally every occurrence of $S_{3,1}^2$ with $1 - S_{1,1}^2 - S_{2,1}^2$. This gives a polynomial p_{k_1,k_2} in the entries of S_1. At each step we have subtracted elements of $\mathcal{J}$, so

$$P_{k_1,k_2} - p_{k_1,k_2} \in \mathcal{J}.$$

From formula (3.4.1) for the matrix coefficients of the representation ρ of GL(2), it follows that every monomial in the expansion of $P_{k_1,k_2}(S)$ contains $c_1(k_2)$ variables from the first column of S and $c_2(k_2)$ variables from the second column of S. Furthermore, in each step of the procedure to obtain p_{k_1,k_2} from P_{k_1,k_2}, the number of variables in each monomial from a given column stays the same or drops by an even number. This shows that in each monomial in $p_{k_1,k_2}(S)$, the number of variables from column $l \in \{1, 2\}$ is at most $c_l(k_2)$, and differs from this by an even number.

We would like to say something similar about the number of variables from each row. For all $g \in O(n)$ we have

$$\begin{aligned} P_{k_1,k_2}(g) &= \langle \Psi(w_{k_1}), \pi(g)\Psi(w_{k_2}) \rangle \\ &= \langle \pi(g^\mathsf{T})\Psi(w_{k_1}), \Psi(w_{k_2}) \rangle \\ &= P_{k_2,k_1}(g^\mathsf{T}), \end{aligned} \tag{3.4.16}$$

where we used the fact that the inner product is unitary and $P_{k_2,k_1}(g^\mathsf{T})$ is real. For each $g \in O(n)$, there is an element $h \in O(n-2)$ such that $(hg)_{i,j} = 0$ for $i > 2 + j$. Hence,

$$P_{k_1,k_2}(g) = P_{k_1,k_2}(hg) = p_{k_1,k_2}(hg) = p_{k_1,k_2}(g),$$

where the second equality holds because hg lies in the vanishing locus of $\mathcal{J}$. Using (3.4.16), it follows that $p_{k_1,k_2}(g) = p_{k_2,k_1}(g^\mathsf{T})$ for all $g \in O(n)$. The only variables which occur in $p_{k_1,k_2}(S)$ and $p_{k_2,k_1}(S^\mathsf{T})$ are from the top-left 2×2 block of S and so we may view them as functions on $\mathbb{R}^4$. Consider the subset of $\mathbb{R}^4$ given by projecting $O(n)$ to the top-left 2×2 block. Since $p_{k_1,k_2}(g) = p_{k_2,k_1}(g^\mathsf{T})$ for all $g \in O(n)$, the functions agree on this subset. This subset has a nonempty interior, and hence the polynomials agree on $\mathbb{R}^4$. Hence we have equality of polynomials: $p_{k_1,k_2}(S) = p_{k_2,k_1}(S^\mathsf{T})$. This shows that in each monomial in $p_{k_1,k_2}(S)$, the number of variables from row $l \in \{1, 2\}$ is at most $c_l(k_1)$, and differs from this by an even number.

Since $c_l(k_1) = 0$ for $l > i_1$ and $c_l(k_2) = 0$ for $l > i_2$, it follows that $p_{k_1,k_2}(S)$ is a polynomial in the top-left $i_1 \times i_2$ block. As discussed in the proof of Lemma 3.6, the (l_1, l_2) entry, with $1 \leq l_1 \leq i_1$ and $1 \leq l_2 \leq i_2$, of $s(J_1)^{\mathsf{T}} s(J_2)$ has denominator $q_{l_1}(J_1)q_{l_2}(J_2)$. To obtain the zonal matrix entry, we may replace each monomial S^a in $p_{k_1,k_2}(S)$ with

$$\begin{aligned}&\xi_{\lambda,i_1,j_1,k_1}(J_1)\xi_{\lambda,i_2,j_2,k_2}(J_2)(s(J_1)^{\mathsf{T}} s(J_2))^a \\ &\quad = q_1(J_1)^{c_1(k_1)} q_2(J_1)^{c_2(k_1)} q_1(J_2)^{c_1(k_2)} q_2(J_2)^{c_2(k_2)} (s(J_1)^{\mathsf{T}} s(J_2))^a,\end{aligned}$$

and by the properties of p_{k_1,k_2} as discussed above, this is a polynomial in the entries of the vectors in $J_1 \cup J_2$. From this we can easily read of the polynomial in terms of the inner products between these vectors.

We now describe additional techniques to speed up the implementation. By Section 3.4.5.1 and 3.4.5.2, the integrand of $P_{k_1,k_2}(S)$ may be replaced by the product of

$$\mathcal{R}(\det(\overline{\omega}\gamma\epsilon)\rho(\overline{\omega}\gamma\epsilon)^*_{k_1,0}) \tag{3.4.17}$$

and

$$\mathcal{R}(\rho(\omega\gamma S\epsilon)_{0,k_2}((\omega\gamma S\epsilon)_{1,1}(\omega\gamma S\epsilon)_{2,2})^{\lambda_2}). \tag{3.4.18}$$

The integration over $O(n)$ and the substitution procedure described above may be swapped. We first compute (3.4.18) explicitly as a polynomial in the variables $S_{i,j}$ with $i \leq 2 + j$ and $j \leq 2$ and the top-left 4×4 block of γ. We then perform the above substitution procedure. Since we know that after integration over $O(n)$ all terms with variables from S_2 will vanish, we remove those terms. This gives a polynomial in the top-left $i_1 \times i_2$ block of S and the top-left 4×4 block of γ.

Whenever $\sum_i a_{ij}$ or $\sum_i a_{ji}$ is odd for any j, we have

$$\int_{O(n)} \gamma^a \, d\gamma = 0.$$

This means that we do not have to work out the product of the whole polynomial (3.4.18) with (3.4.17). Instead, we only multiply terms that produce monomials in γ which do not immediately vanish. We then integrate each monomial in γ using the recursion formulas of [**61**]. This enables us to explicitly compute $p(S)$, from which we obtain the zonal matrix entry as explained above.

CHAPTER 4

Polynomial optimization

In this chapter we consider problems of the form

$$\begin{aligned} &\text{maximize} && \langle C, X\rangle \\ &\text{subject to} && \langle A_i(x), X\rangle \le b_i(x), \quad x \in \Delta_i, \quad i = 1, \dots, m, \\ & && X \succeq 0. \end{aligned}$$

where the optimization variable X is a positive semidefinite matrix, $A_i(x)$ is a matrix with polynomials as entries, $b_i(x)$ is a polynomial, and $\Delta_i \subseteq \mathbb{R}^{n_i}$ is a basic closed semialgebraic set on which constraint i should hold. Here $\langle A, B\rangle = \mathrm{Tr}(A^\mathsf{T} B)$ denotes the trace inner product.

Using explicit constructions of Γ-invariant polynomial kernels, the hierarchies considered in Chapter 2 can be written in this form after truncating the Fourier series of the kernels. Recall that a feasible solution to the problem is enough for a valid bound: we may, if needed, restrict the problem to a subset of feasible solutions to obtain a computable problem.

Example 2.1 (Packing, continued). *Recall that we can write our constraints as*

$$A_k K(S) \le -\chi_{\mathcal{I}_{=1}}(S)$$

for S in $\mathcal{I}_{2k} \setminus \mathcal{I}_0$, with K of the form

$$K(J_1, J_2) = \sum_\lambda \langle K_\lambda, Z_\lambda(J_1, J_2)\rangle,$$

where the entries of Z_λ are polynomials in the inner products between vectors in $S = J_1 \cup J_2$. Since $\mathcal{I}_{2k} \setminus \mathcal{I}_0$ is the disjoint union of $\mathcal{I}_{=i}$ for $i = 1, \dots, 2k$, we can split the constraint into $2k$ constraints where constraint i concerns $\mathcal{I}_{=i}$. Restricting the Fourier series of the kernel to partitions λ with $|\lambda| = \sum_i \lambda_i \le d_1$ and the (infinite) matrix Z_λ to polynomials of degree at most $d_2 \ge d_1$, we obtain a problem as introduced at the start of this chapter. Here the semialgebraic sets are

$$\Delta_i = \{u \in \mathbb{R}^{\binom{i}{2}} : p(u_1) \ge 0, \dots, p(u_{\binom{i}{2}}) \ge 0, G(u) \succeq 0\},$$

where $p(u) = (u+1)(\cos\theta - u)$ and the matrix $G(u)$ denotes a square $i \times i$ matrix with 1 on the diagonal and u on the off-diagonal, in a chosen order. The matrix $G(u)$ represents the Gram matrix of i points on the sphere. By Sylvester's criterion, $G(u) \succeq 0$ is equivalent to requiring that all principal minors of $G(u)$ are nonnegative, and we replace $G(u) \succeq 0$ by these constraints. Note that the 1×1 principal minors equal 1, and the 2×2 principal minors are nonnegative since $p(u_j) \ge 0$ for $j = 1, \dots, \binom{i}{2}$, so these can be ignored in the description.

4.1. Putinar's Positivstellensatz

In this section, we focus on a single polynomial constraint of the form

$$\langle A(x), X\rangle \le b(x), \quad x \in \Delta,$$

where $\Delta = \{x \in \mathbb{R}^n : g_i(x) \ge 0, i = 1, \dots, m\}$.

Consider the quadratic module

$$Q(g_1, \dots, g_m) = \{p_0^2 + \sum_{i=1}^{m} g_i p_i^2 : p_0, \dots, p_m \in \mathbb{R}[x_1, \dots, x_n]\}.$$

Typically, we will set $g_0 = 1$, so that we can write the elements as a single sum over $i = 0, \dots, m$. Let Q_d denote elements of $Q(g_1, \dots, g_m)$ such that each term $g_i p_i^2$ is a polynomial of degree at most d. Clearly, if $p \in Q(g_1, \dots, g_m)$ we have $p(x) \ge 0$ for all $x \in \Delta$, since $g_i(x) \ge 0$ and $p_i(x)^2 \ge 0$ for all $i = 0, \dots, m$. Under some conditions, a partial converse statement also holds. A quadratic module $Q(g_1, \dots, g_m)$ is called *Archimedean* if there is an $N \ge 0$ such that $N - \sum_i x_i^2 \in Q(g_1, \dots, g_m)$.

THEOREM 4.1 (Putinar's Positivstellensatz [**123**]). *Suppose $Q(g_1, \dots, g_m)$ is Archimedean and $p > 0$ on Δ. Then $p \in Q(g_1, \dots, g_m)$.*

Note that here the quadratic module is not restricted to a finite degree.

To use this for our problems, we set $p(x) = b(x) - \langle A(x), X\rangle$. Every solution X such that $p \in Q_d(g_1, \dots, g_m)$ is feasible, and if the program has a strictly feasible solution (that is, an X such that every constraint is strictly satisfied on the whole semialgebraic set corresponding to the constraint), the theorem ensures that the optimal value of the problem can be approximated arbitrarily well by increasing the degree d. Since a sum-of-squares polynomial is of the form

$$\sum_i p_i(x)^2 = \langle q(x)q(x)^\mathsf{T}, Y\rangle \tag{4.1.1}$$

where $Y \succeq 0$ and $\{q_i\}_i$ is a basis of the space of polynomials up to the degree of the sum-of-squares polynomial, this reduces the problem to the form

$$\begin{aligned} \text{maximize} \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i(x), X\rangle = b_i(x), \quad i = 1, \dots, m, \\ & X \succeq 0. \end{aligned}$$

where X and A are block-diagonal, and some blocks of $A_i(x)$ are now of the form $g_j(x)q(x)q(x)^\mathsf{T}$.

4.2. Semidefinite programming constraints

Consider a constraint

$$\langle A(x), X\rangle = b(x),$$

where $X \succeq 0$ is the variable of the problem, and b and the entries of A are polynomials $b, A_{i,j} \in \mathbb{R}[x_1, \dots, x_n]$. Let W denote $\operatorname{Span}\{A_{i,j}(x), b_i(x)\}_{i,j}$, and let $\{w_i\}_i$ be a basis of W. The standard technique to reformulate the polynomial constraint as semidefinite programming constraints is to equate the coefficients of $\langle A(x), X\rangle$ and $b(x)$ in the basis $\{w_i\}_{i=1}^{\dim W}$. Let $\{A_i\}$ be the matrices of coefficients of the entries of $A(x)$ in the basis $\{w_i\}$. Decomposing the constraint in this basis then gives

$$\sum_i \langle A_i, X\rangle w_i = \sum_i b_i w_i$$

which is equivalent to the constraints

$$\langle A_i, X\rangle = b_i, \quad i = 1, \dots, \dim W.$$

Different bases of W have different advantages. Here we highlight one particular type of basis, considered in [**102**]. We call a set of points $Z \subseteq \mathbb{R}^n$ *unisolvent* for a subspace W if for all $p \in W$ we have $p(x) = 0$ for all $x \in Z$ if and only if $p = 0$ as a polynomial. A unisolvent set Z is *minimal* if $|Z| = \dim W$. Consider the basis of W indexed by Z such that $w_x(y) = \delta_{xy}$ for $y \in Z$, where δ_{xy} is the Kronecker delta function. When $n = 1$, this is called the Lagrange basis, and is often used for interpolation problems. It has the property that the coefficients c_x of a polynomial p are given by $c_x = p(x)$. This implies that the semidefinite constraints corresponding to a sum-of-squares constraint

$$\langle qq^\mathsf{T}, X\rangle = b$$

are given by

$$\langle q(x)q(x)^\mathsf{T}, X\rangle = b(x), \quad x \in Z.$$

Note in particular that the constraint matrices are of rank 1. Example 3.4 shows that, after applying symmetry reduction to the sum-of-squares polynomials, the resulting constraint matrices will have at most rank d_π, where π is an irreducible representation of the symmetry group. In Chapter 5, we will see that the low-rank structure can be used to solve the semidefinite programs faster.

4.3. Invariant polynomials*

Recall from Example 3.4 that, given an action of a group Γ on $\mathbb{R}^n$, the natural action of Γ on $\mathbb{R}[x_1, \dots, x_n]$ is given by

$$\gamma p(x) = p(\gamma^{-1}x).$$

In this section we assume Γ to be a finite group. Consider the weighted sum-of-squares constraint

$$\langle A(x), X\rangle + \sum_{i=0}^{m} g_i(x)\sigma_i(x) = b(x) \tag{4.3.1}$$

where again $g_0 = 1$. Suppose the polynomials b, g_i and all entries of A are Γ-invariant. Then acting with Γ on the constraint gives

$$\begin{aligned} b(x) &= b(\gamma^{-1}x) \\ &= \langle A(\gamma^{-1}x), X\rangle + \sum_i g_i(\gamma^{-1}x)\sigma_i(\gamma^{-1}x) \\ &= \langle A(x), X\rangle + \sum_i g_i(x)\sigma_i(\gamma^{-1}x) \end{aligned}$$

Hence taking $\tilde{\sigma}_i = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\gamma\sigma_i$ also gives a feasible solution. This implies that we may take σ_i to be Γ-invariant sum-of-squares polynomials, which can be parametrized using Examples 3.4. Such Γ-invariant sum-of-squares polynomials were considered by Gatermann and Parillo in [**58**].

Often, the description of $\Delta = \{x \in \mathbb{R}^n : g_i(x) \geq 0, i = 1, \dots, m\}$ is not in terms of Γ-invariant polynomials, even when Δ itself is Γ-invariant. One example is when

*Part of this section is adapted from the publication "*N. Leijenhorst and D. de Laat,* Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput.* ***16*** *(2024), no. 3, 503-534, doi:10.1007/s12532-024-00264-w*".

Δ corresponds to the possible off-diagonal entries of a Gram matrix of points on S^{n-1}: with 3 points this equals

$$\Delta = \{(u,v,t) : \begin{bmatrix} 1 & u & v \\ u & 1 & t \\ v & t & 1 \end{bmatrix} \succeq 0\}.$$

The description in terms of polynomials is usually given by the minors of the Gram matrix, since by Sylvester's criterion all principle minors are nonnegative if and only if the matrix is positive semidefinite. This description is not S_3 invariant (the 2×2 minors are not invariant under permuting the variables), but the semialgebraic set itself is S_3-invariant, and there does exist an S_3-invariant description of Δ (see, e.g., [**103**, Lemma 3.1]).

Suppose $\{g_1, \dots, g_m\} = \Gamma g$ is an orbit under Γ. Note that, if g is invariant under a subgroup of Γ, m can be strictly smaller than $|\Gamma|$. We will construct a set of Γ-invariant polynomials such that $\Delta(g_1, \dots, g_m) = \Delta(\tilde{g}_1, \dots, \tilde{g}_m)$.

Define the polynomials

$$\tilde{g}_k = \sum_{\substack{J \subseteq [m] \\ |J|=k}} \prod_{i \in J} g_i.$$

Then

$$\gamma \tilde{g}_k = \sum_{\substack{J \subseteq [m] \\ |J|=k}} \prod_{i \in J} \gamma g_i = \sum_{\substack{J \subseteq [m] \\ |J|=k}} \prod_{i \in J} g_{\sigma(i)} = \sum_{\substack{J \subseteq [m] \\ |J|=k}} \prod_{i \in \sigma(J)} g_i = \tilde{g}_k$$

for some $\sigma \in S_m$, since $\{g_1, \dots, g_m\}$ is an orbit of Γ. Hence $\tilde{g}_k$ is Γ-invariant. Furthermore, let $x \in \Delta(g_1, \dots, g_m)$, then $\tilde{g}_k(x) \geq 0$ for all k because it is a sum of products of $g_i(x) \geq 0$, so $\Delta(g_1, \dots, g_m) \subseteq \Delta(\tilde{g}_1, \dots, \tilde{g}_m)$.

For the reverse inclusion, we suppose there is a point x with $g_1(x) < 0$ and $\tilde{g}_k(x) \geq 0$ for $k \geq 2$. We will argue that $\tilde{g}_1(x) < 0$.

Define

$$T_k = \sum_{\substack{J \subseteq [m] \setminus \{1\} \\ |J|=k}} \prod_{i \in J} g_i.$$

Then $T_m(x) \leq 0$, since J cannot have size m. If $T_k(x) \leq 0$, then $\tilde{g}_k(x) = g_1(x) T_{k-1}(x) + T_k(x) \geq 0$ implies $g_1(x) T_{k-1}(x) \geq 0$ and therefore $T_{k-1}(x) \leq 0$. By induction, we have $T_1(x) \leq 0$. Hence

$$\tilde{g}_1(x) = g_1(x) + T_1(x) < 0,$$

so $\Delta(g_1, \dots, g_m) = \Delta(\tilde{g}_1, \dots, \tilde{g}_m)$.

EXAMPLE 2.1 (Packing, continued). *The set Δ_i is S_i-invariant, where S_i acts on $u \in \mathbb{R}^{\binom{i}{2}}$ by permuting the rows and columns of the Gram matrix $G(u)$. The polynomials used for the description, however, are not. For example, the set Δ_3 is given by*

$$\{u \in \mathbb{R}^3 : p(u_1) \geq 0, p(u_2) \geq 0, p(u_3) \geq 0, 1 + 2u_1u_2u_3 - u_1^2 - u_2^2 - u_3^2 \geq 0\}$$

The non-invariant polynomials $g_i(u) = p(u_i) = (u_i + 1)(\cos\theta - u_i)$ *form an orbit under* S_3*, and the corresponding polynomials* $\tilde{g}_i$ *are given by*

$$\begin{aligned}\tilde{g}_1(u) &= p(u_1) + p(u_2) + p(u_3),\\ \tilde{g}_2(u) &= p(u_1)p(u_2) + p(u_1)p(u_3) + p(u_2)p(u_3),\\ \tilde{g}_3(u) &= p(u_1)p(u_2)p(u_3),\end{aligned}$$

cf. [**103**, Lemma 3.1].

CHAPTER 5

Solving semidefinite programs*

Many solvers are available to solve semidefinite programs (see, for example, [**7, 138, 143, 132**]). For applications in discrete geometry, the corresponding semidefinite programs have seemingly unavoidable bad numerical conditioning. Because we additionally need a solution of high precision, we require a second-order interior point method using high-precision arithmetic. In practice, computations are therefore performed using the general purpose semidefinite programming solvers `SDPA-QD` and `SDPA-GMP` [**143**], which both use high-precision numerics, and computations regularly take weeks to complete (see, e.g., [**103**]).

In this chapter, we introduce a high-precision primal-dual interior point method that uses additional low-rank structure to speed up the computations. As mentioned in Section 4.2, such low-rank constraint matrices can be present in semidefinite programs arising in polynomial optimization. We follow the structure of the solver `SDPB` [**132**], which in turn builds on `SDPA` [**143**]. We generalize the specialization to a very general low-rank structure (see (5.1.2) and (5.1.3)), and we show how this can be exploited in the computation of the Schur complement matrix in a way that fast matrix-matrix multiplication can be employed (which is especially beneficial because we use high-precision arithmetic). Similar low-rank structures are also used in other solvers, such as [**7, 138**], which do not use high-precision arithmetic.

Because our applications consist of problems in discrete geometry, which typically have few clusters and a large number of constraints within a cluster, our parallelization strategy is different from `SDPB`. The interior point method uses the XZ search direction [**70, 80, 110**], the predictor-corrector step due to Mehrotra [**105**], and the Lanczos algorithm for computing step lengths [**137**]. The algorithm is implemented as a Julia package, and can be found at `github.com/nanleij/ClusteredLowRankSolver.jl`.

5.1. Clustered low-rank semidefinite programs

When translating polynomial constraints into semidefinite constraints (see Section 4.2), one obtains for each polynomial constraint a number of semidefinite constraints which use the same positive semidefinite matrix variables. Using sampling, it is possible to keep the rank of the constraint matrices low [**102**]. Together, this leads to a clustered low-rank semidefinite program, with clusters of constraints using the same positive semidefinite variables, and low-rank constraint matrices. We assume these clusters are connected only through free scalar variables.

*This chapter is based on Section 2 of the publication "*N. Leijenhorst and D. de Laat,* Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput.* **16** *(2024), no. 3, 503-534, doi:10.1007/s12532-024-00264-w*".

We therefore consider semidefinite programs with J clusters of the form

$$\begin{array}{lll} \text{maximize} & \displaystyle\sum_{j=1}^{J} \langle C^j, Y^j \rangle + \langle c, y \rangle & \\ \text{subject to} & \left\langle A^j_*, Y^j \right\rangle + B^j y = b^j, & j = 1, \ldots, J \\ & Y^j \succeq 0, & j = 1, \ldots, J, \end{array} \tag{5.1.1}$$

where we optimize over the vector of free variables y and the positive semidefinite block matrices $Y^j = \operatorname{diag}(Y^{j,1}, \ldots, Y^{j,L_j})$. Here $\langle c, y \rangle$ is the Euclidean inner product, and we use the notation

$$\langle A^j_*, Y^j \rangle = \left(\langle A^j_t, Y^j \rangle \right)_{t \in T_j},$$

where $\langle A^j_t, Y^j \rangle$ is the trace inner product.

The semidefinite program is defined by the symmetric matrices C^j and A^j_t, the matrices B^j, and the vectors $c \in \mathbb{R}^N$ and $b^j \in \mathbb{R}^{T_j}$. We assume the matrix A^j_t is of the form

$$A^j_t = \bigoplus_{l=1}^{L_j} \sum_{r,s=1}^{R_j(l)} A^{j,l}_t(r,s) \otimes E^{R_j(l)}_{r,s}, \tag{5.1.2}$$

where $A^{j,l}_t(r,s)$ can be a matrix of low rank and $A^{j,l}_t(r,s)^{\mathsf{T}} = A^{j,l}_t(s,r)$. Here $E^n_{r,s}$ is the $n \times n$ matrix with a one at position (r,s) and zeros otherwise. The implementation also allows for dense blocks, although the type (low-rank or dense) of the blocks should be consistent per (j,l) over the subblocks and the constraints: $A^{j,l}_t(r,s)$ is of the same type as $A^{j,l}_{t'}(r',s')$ but can be of different type than $A^{j',l'}_{t'}(r',s')$ when $(j,l) \neq (j',l')$.

Example 5.1. *In the problems we solve in this thesis, we typically have two types of blocks $A^{j,l}$: the zonal matrices Z_λ from Chapter 3, which are typically 'dense' blocks, and the ones corresponding to the (symmetric) sum-of-squares polynomials with $R_j(l) = 1$ and rank d_π. In these problems, since the zonal matrices are present in all constraints, we have $J = 1$.*

Internally, we represent the blocks $A^{j,l}_t(r,s)$ in the form

$$\sum_i \lambda_i v_i w_i^{\mathsf{T}}, \tag{5.1.3}$$

where we do not require the rank 1 terms to be symmetric (even if the block $A^{j,l}_t(r,s)$ itself is symmetric). Allowing for nonsymmetric matrices in the rank 1 decomposition is more general than what is done in [**132, 138, 7**]. Although in practice one can use a spectral decomposition and thus set $v_i = w_i$, this is not necessarily possible when additionally rounding the numerical solution to an exact solution, for which an exact problem description is necessary, see Chapter 6.

We interpret (5.1.1) as the dual of the semidefinite program

$$\begin{aligned} &\text{minimize} && \sum_{j=1}^{J} \langle b^j, x^j \rangle \\ &\text{subject to} && \sum_{j=1}^{J} (B^j)^{\mathsf{T}} x^j = c \\ & && X^j = \sum_{t \in T_j} x_t^j A_t^j - C^j \succeq 0, \qquad j = 1, \ldots, J, \end{aligned} \tag{5.1.4}$$

where we optimize over the vectors of free variables x^j and the positive semidefinite block matrices $X^j = \operatorname{diag}(X^{j,1}, \ldots, X^{j,L_j})$.

Using the notation X for the block matrix $\operatorname{diag}(X^1, \ldots, X^J)$ and Y for the block matrix $\operatorname{diag}(Y^1, \ldots, Y^J)$, the duality gap for primal feasible (x, X) and dual feasible (y, Y) is given by

$$b^{\mathsf{T}} x - \langle C, Y \rangle - c^{\mathsf{T}} y = \langle X, Y \rangle.$$

We assume strong duality holds, so that if (x, X) and (y, Y) are optimal, then $\langle X, Y \rangle = 0$, and hence $XY = 0$.

5.2. A general primal-dual algorithm

We follow [**132**] for the main steps of the algorithm.

The primal-dual algorithm starts with infeasible primal solutions (x, X) and dual solutions (y, Y), where X and Y are positive definite. At each iteration, a Newton direction (dx, dX, dy, dY) is computed for the system of primal and dual linear constraints and the centering condition $XY = \beta_p \mu I$. Here μ is the surrogate duality gap $\langle X, Y \rangle$ divided by the size of the matrices, and

$$\beta_p = \begin{cases} 0 & \text{primal and dual feasible,} \\ \beta_{infeasible} & \text{otherwise.} \end{cases}$$

That is, we attempt a step to the end of the central path if we start from a primal-dual feasible solution, and try to get a factor $\beta_{infeasible}$ closer to the end otherwise. This is the predictor step.

For the corrector step, the system is solved with $(x+dx, X+dX, y+dy, Y+dY)$ and β_c instead of β_p. The parameter β_c is among others determined by the change of the surrogate duality gap, modeled after the choice in `SDPA` and `SDPB`, see also [**132**, Section 2.4.4]. This results in a new search direction (dx, dX, dy, dY). We then take a step in the direction of the corrector step: (x, X, y, Y) is replaced by $(x + \gamma s_p dx, X + \gamma s_p dX, y + \gamma s_d dy, Y + \gamma s_d dY)$ for some step sizes s_p, s_d such that $X + s_p dX$ and $Y + s_d dY$ are positive definite, with $\gamma \in (0, 1)$. When the primal and dual solutions are feasible, we take $s_d = s_p$. In the following sections we discuss the process of finding the search direction, since exploiting the special form (5.1.2) happens in this part of the algorithm.

5.2.1. Computing the search direction. To compute the Newton search direction we replace the variables (x, X, y, Y) by $(x + dx, X + dX, y + dy, Y + dY)$

in the primal and dual constraints, which gives

$$X^j + dX^j = \sum_{t \in T_j} (x_t^j + dx_t^j) A_t^j - C^j, \tag{5.2.1}$$

$$\sum_{j=1}^{J} (B^j)^\mathsf{T} (x^j + dx^j) = c, \tag{5.2.2}$$

$$\left\langle A_*^j, Y^j + dY^j \right\rangle + B^j(y + dy) = b^j. \tag{5.2.3}$$

Then we apply the same substitution in the centering condition and linearize to get

$$X^j Y^j + X^j dY^j + dX^j Y^j = \mu I. \tag{5.2.4}$$

Substituting the expression for dX^j from (5.2.1) into (5.2.4) and then the expression for dY^j from (5.2.4) into (5.2.3) gives

$$\left\langle A_*^j, Y^j + (X^j)^{-1}\Big(\mu I - X^j Y^j - \Big(\sum_{t \in T_j} (x_t^j + dx_t^j) A_t^j - C^j - X^j\Big) Y^j\Big)\right\rangle + B^j(y + dy) = b^j.$$

Together with constraint (5.2.2) (which is responsible for the last row in the system) this can be written as the following linear system in dx and dy:

$$\begin{bmatrix} S^1 & 0 & \cdots & 0 & -B^1 \\ 0 & S^2 & \cdots & 0 & -B^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & S^J & -B^J \\ (B^1)^\mathsf{T} & (B^2)^\mathsf{T} & \dots & (B^J)^\mathsf{T} & 0 \end{bmatrix} \begin{bmatrix} dx^1 \\ dx^2 \\ \vdots \\ dx^J \\ dy \end{bmatrix} = \begin{bmatrix} -b^1 - \langle A_*^1, Z^1 - Y^1 \rangle + B^1 y \\ -b^2 - \langle A_*^2, Z^2 - Y^2 \rangle + B^2 y \\ \vdots \\ -b^J - \langle A_*^J, Z^J - Y^J \rangle + B^J y \\ c - \sum_{j=1}^J (B^j)^\mathsf{T} x^j \end{bmatrix},$$

Here $Z^j = (X^j)^{-1}((\sum_t x_t^j A_t^j - C^j) Y^j - \mu I)$ and the blocks S^j that form the Schur complement matrix $S = \operatorname{diag}(S^1, \dots, S^J)$ have entries

$$S_{ab}^j = \left\langle A_a^j, (X^j)^{-1} A_b^j Y^j \right\rangle.$$

The above system can be solved to obtain dx and dy. From this dX and dY can be computed, where instead of computing dY as $X^{-1}(\mu I - XY - dXY)$ we set

$$dY = \frac{X^{-1}(\mu I - XY - dXY) + (X^{-1}(\mu I - XY - dXY))^\mathsf{T}}{2}$$

so that Y stays symmetric.

In general, the computation of the Schur complement matrix S and solving the above linear system are the main computational steps.

Due to the clusters, the matrix S is block-diagonal, so that the Cholesky factorization $S = LL^\mathsf{T}$ can be computed blockwise. By using the decomposition

$$\begin{bmatrix} S & -B \\ B^\mathsf{T} & 0 \end{bmatrix} = \begin{bmatrix} L & 0 \\ B^\mathsf{T} L^{-\mathsf{T}} & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & B^\mathsf{T} L^{-\mathsf{T}} L^{-1} B \end{bmatrix} \begin{bmatrix} L^\mathsf{T} & -L^{-1}B \\ 0 & I \end{bmatrix},$$

we can solve the system by solving several triangular systems. The inner matrix $B^\mathsf{T} L^{-\mathsf{T}} L^{-1} B$ is positive definite, so we can again use a Cholesky decomposition.

5.2.2. Speedups for low-rank matrices. Due to the low-rank constraint matrices, we can compute the blocks S^j more efficiently. Suppose for simplicity the constraint matrices are of the form

$$A_t^{j,l} = \sum_{r=1}^{\eta_t^{j,l}} \lambda_{t,r}^{j,l} v_{t,r}^{j,l} (w_{t,r}^{j,l})^{\mathsf{T}},$$

where $\eta_t^{j,l}$ is the rank of the matrix $A_t^{j,l}$. Then we can write

$$\begin{aligned} S_{ab}^j &= \sum_{l=1}^{L_j} \langle A_a^{j,l}, (X^{j,l})^{-1} A_b^{j,l} Y^{j,l} \rangle \\ &= \sum_{l=1}^{L_j} \sum_{r_1=1}^{\eta_a^{j,l}} \sum_{r_2=1}^{\eta_b^{j,l}} \lambda_{a,r_1}^{j,l} \lambda_{b,r_2}^{j,l} \left\langle v_{a,r_1}^{j,l} (w_{a,r_1}^{j,l})^{\mathsf{T}}, (X^{j,l})^{-1} v_{b,r_2}^{j,l} (w_{b,r_2}^{j,l})^{\mathsf{T}} Y^{j,l} \right\rangle \\ &= \sum_{l=1}^{L_j} \sum_{r_1=1}^{\eta_a^{j,l}} \sum_{r_2=1}^{\eta_b^{j,l}} \lambda_{a,r_1}^{j,l} \lambda_{b,r_2}^{j,l} \Big((w_{a,r_1}^{j,l})^{\mathsf{T}} (X^{j,l})^{-1} v_{b,r_2}^{j,l} \Big) \Big((w_{b,r_2}^{j,l})^{\mathsf{T}} Y^{j,l} v_{a,r_1}^{j,l} \Big), \end{aligned}$$

which shows we can compute S_{ab}^j efficiently by precomputing the bilinear pairings

$$(w_{a,r_1}^{j,l})^{\mathsf{T}} (X^{j,l})^{-1} v_{b,r_2}^{j,l} \quad \text{and} \quad (w_{b,r_2}^{j,l})^{\mathsf{T}} Y^{j,l} v_{a,r_1}^{j,l}.$$

In the implementation, we use similar techniques for the more general constraint matrices of the form (5.1.2). Since we use high-precision arithmetic, it is beneficial to use matrix-matrix multiplication with subcubic complexity, and therefore we compute the above pairings efficiently by first creating the matrices $V^{j,l}$ and $W^{j,l}$ with the columns $v_{a,r}^{j,l}$ and $w_{a,r}^{j,l}$, respectively, and then performing fast matrix multiplication to compute

$$(W^{j,l})^{\mathsf{T}} (X^{j,l})^{-1} V^{j,l} \quad \text{and} \quad (W^{j,l})^{\mathsf{T}} Y^{j,l} V^{j,l}.$$

Due to the block structures, the algorithm is relatively easy to parallelize. The best way to parallelize, however, depends on both the problem characteristics and the type of computing system used. The `SDPB` solver specializes in problems with large amounts of clusters with similar-sized blocks, which can be distributed over different nodes in a multi-node system in which there is communication latency between the nodes [**132, 93**].

Problems in discrete geometry typically consist of few clusters, and have a large variation in both the number of blocks per cluster and in the size of the blocks; the blocks corresponding to the sum-of-squares polynomials of Chapter 4 are much larger than the blocks corresponding to the positive definite kernels of Section 3.4. The majority of the workload can be due to a single cluster, hence distributing clusters over nodes in a multi-node system is not a good parallelization strategy in this case. Instead, we focus on distributing the workload over multiple cores in a single-node shared-memory system.

Most of the matrix operations can be done block-wise. We distribute the blocks over the cores such that the workload for each core is about equal. Since the matrices in the products $(W^{j,l})^{\mathsf{T}} (X^{j,l})^{-1} V^{j,l}$ and $(W^{j,l})^{\mathsf{T}} Y^{j,l} V^{j,l}$ can be very large, we split these multiplications into several parts which we distribute over the cores.

CHAPTER 6

Rounding to exact solutions*

6.1. Introduction

The optimality and uniqueness proofs in this thesis employ the complementary slackness conditions of Theorem 2.10. However, for this we require an exact solution to (2.4.2), or its semidefinite programming relaxation. The primary technical obstacle is how to convert the approximate, floating-point output of a semidefinite programming solver to an exact optimal solution. Monniaux and Corbineau [**109**] and Dostert, de Laat, and Moustrou [**51**] propose rounding methods for this using the LLL algorithm in different ways, but both of these methods are too slow for the sizes of semidefinite programs we consider in this thesis.

The reason the previous approaches become too slow for large semidefinite programs stems mainly from the following two bottlenecks. Firstly, both approaches operate with the full matrix defining the affine constraints of the semidefinite program (the first to write the semidefinite program in LMI form and the second to project into the affine space). Secondly, both approaches apply the LLL algorithm to a lattice of dimension linear in the size of the matrix variable in the semidefinite program. In this chapter we propose a method which avoids both bottlenecks.

In this chapter we consider a semidefinite program in primal form

$$\begin{aligned} &\text{minimize} && \langle C, X\rangle \\ &\text{subject to} && \langle A_i, X\rangle = b_i \text{ for } i = 1, \dots, m \text{ and} \\ & && X \in \mathbb{S}^n_+. \end{aligned} \tag{6.1.1}$$

Here $\mathbb{S}^n_+$ is the cone of positive semidefinite $n \times n$ matrices over $\mathbb{R}$ (for comparison, $\mathbb{S}^n$ will denote the set of symmetric $n \times n$ matrices, and $\mathbb{S}^n_{++}$ the cone of positive definite $n \times n$ matrices), and $\langle A, B\rangle = \operatorname{tr}(A^\mathsf{T} B)$ is the trace inner product. The problem is specified by the symmetric $n \times n$ matrices $C, A_1, \dots, A_m$ and real scalars $b_1, \dots, b_m$, and we generally assume the affine constraints are linearly independent.

All matrices and scalars in the problem definition are assumed to be defined over an algebraic field of low degree. This is the case for most problems in this thesis, but for example in Chapter 9, some entries of the matrices A_i are linear combinations of powers of π, due to integration of polynomials in inner products over a sphere.

*This chapter is based on Section 2 of the publication "*H. Cohn, D. de Laat and N. Leijenhorst,* Optimality of spherical codes via exact semidefinite programming bounds*, 2024, arXiv:2403.16874*".

The corresponding dual semidefinite program reads

$$\begin{aligned} &\text{maximize} && \langle b, y \rangle \\ &\text{subject to} && C - \sum_{i=1}^{m} y_i A_i \in \mathbb{S}^n_+. \end{aligned}$$

We assume the existence of strictly feasible points for both the primal and dual semidefinite programs, which are solutions where the matrix is positive definite. By Slater's criterion, this condition guarantees the existence of optimal primal and dual solutions, and these solutions have the same objective value (see, for example, Section 5.3.2 in [**19**]). We also assume the existence of a strictly complementary solution, which is a pair (X, y) of primal and dual solutions such that

$$\operatorname{rank}(X) + \operatorname{rank}\left(C - \sum_{i=1}^{m} y_i A_i \right) = n.$$

Under this assumption, the central path (which is followed by interior-point methods used in practice) converges to the analytic center of the optimal face [**64**]. This means we can use a solver to find a numerical approximation of a point in the relative interior of the optimal face.

Although there exist rational semidefinite programs for which the following is not the case [**116**], for many problems that we encounter in practice the affine hull of the optimal face is given by affine equations whose coefficients lie in a common algebraic number field of low degree. The rounding method from [**51**] finds and uses such a description to round a numerical solution to an exact optimal solution over the field of algebraic numbers. In this section, we explain this heuristic, and we explain our changes that make it much faster.

In practice, the semidefinite programs we consider are in block form (where the cone $\mathbb{S}^n_+$ is replaced by a product of positive semidefinite matrix cones), but here we only consider the case of one block, as the extension to multiple blocks is immediate. We also only consider the case of rounding to rationals and postpone the case of rounding to a number field to Section 6.5.

One simple approach is to compute the solution to high precision (using an arbitrary precision solver) and then use the LLL lattice basis reduction algorithm [**99**] to round each entry of the numerical solution X^* to a nearby algebraic number. This method does work for some examples, but in our experiments, it often does not work when the optimal solution is not unique. For example, we have not been able to find the number field for the analytic center of the semidefinite program we used in the optimality proof of the 56-point spherical code in 20 dimensions (see Chapter 7), even after solving the semidefinite program to extremely high precision. It seems the analytic center, in this case, requires an algebraic number field of high degree or with large coefficients.

Simply projecting the solution into the affine space given by the constraints $\langle A_i, X \rangle = b_i$ for $i = 1, \ldots, m$ also often does not work. The issue here is that the dimension of the optimal face of a semidefinite program is usually smaller than the dimension of this affine space. This dimension mismatch implies that if we project a numerical solution into the affine space, the projected solution will generally no longer be positive semidefinite.

The rounding procedure we use consists of three steps. First we find a suitable description of the affine hull of the optimal face (Section 6.2), then we transform the

problem (Section 6.3), and finally we round the numerical solution to a solution of the transformed problem (Section 6.4). For each of the three steps, we first describe the approach from [**51**] and then give our new approach.

6.2. Describing the optimal face

Let $\mathcal{S}$ be the set of feasible solutions of the primal semidefinite program (6.1.1), and $\mathcal{F}$ the face of $\mathcal{S}$ consisting of the optimal solutions. As shown in [**72**], the minimal face of $\mathbb{S}^n_+$ containing a matrix $X \in \mathbb{S}^n_+$ is given by

$$\{Y \in \mathbb{S}^n_+ : \ker X \subseteq \ker Y\}.$$

This implies that if X lies in the relative interior of $\mathcal{F}$, then

$$\mathcal{F} = \{Y \in \mathbb{S}^n_+ : \langle A_i, Y\rangle = b_i \text{ for } i = 1, \ldots, m \text{ and } \ker X \subseteq \ker Y\}$$

Note that $\ker X$ is independent of the choice of X, as long as X lies in the relative interior of $\mathcal{F}$. We assume this kernel has a nice description over $\mathbb{Q}$, by which we mean it is the solution set of some affine functions whose coefficients are rational numbers of reasonably low bit size (i.e., not excessively large numerators or denominators).

The rounding procedure starts by using an interior-point method to find a solution X^* close to the analytic center of the optimal face. Since the analytic center lies in the relative interior of $\mathcal{F}$, there exists a basis of its kernel of small bit size. We want to use X^* to find such a basis.

In [**109**] and [**51**] the LLL algorithm is used to find such basis vectors. In the former the LLL algorithm is applied directly to the columns of the matrix

$$\begin{pmatrix} I \\ \alpha X^* \end{pmatrix},$$

for some large scalar α, so that the first few reduced vectors will be some of the basis vectors we want to find. They give an iterative procedure to find all basis vectors. In [**51**], the following procedure is used instead. First, a basis for the kernel of the numerical solution matrix X^* is computed numerically (for this the dual solution of the semidefinite program could also be used), and the basis vectors are listed as the rows of a matrix. Then the LLL algorithm is used similarly as above to compute integer relations between the columns of this matrix. Once sufficiently many relations have been found, the coefficient vectors of these relations are listed as the rows of an integral matrix and a basis for the kernel of this matrix is computed in exact arithmetic. The problem with these approaches is that the LLL algorithm is too time-consuming for large matrices: in both cases the complexity of one application of the LLL algorithm scales as n^6, where X^* is of size $n \times n$.

We therefore find the kernel vectors as follows. We list the k numerically computed kernel vectors as the rows of a matrix $M \in \mathbb{R}^{k\times n}$, and let M' be its reduced row-echelon form. Since the reduced row-echelon form is unique, and since X^* approximates the analytic center of the optimal face, whose kernel vectors are assumed to be definable over $\mathbb{Q}$, the entries of M' are close approximations of rational numbers. We then replace each entry in M' by the corresponding rational number.

Although this approach works in principle (given that we use high enough floating point precision) and is much faster than the corresponding procedure in [**51**], we notice that the obtained kernel vectors can have significantly larger bit size, which becomes an issue in Section 6.3 and 6.4.

We would like to use the LLL algorithm to reduce the bit size of the vectors (note that this scales as $k^5 n$ instead of the n^6 discussed above). When applied to the rows of M', the algorithm finds integer linear combinations of the rows which are closer to being orthogonal to each other and have smaller norms. To make these outcomes correspond to a smaller bit size, M' first needs to be converted to an integral matrix. If we do so by clearing the denominators per row, then the pivot columns of M' become columns with exactly one large nonzero entry, so that these entries cannot be reduced by the LLL algorithm. We therefore convert M' to an integral matrix using a different method.

Let $N \in \mathbb{Q}^{n\times(n-k)}$ be a matrix of full column rank such that $M'N = 0$, which is easily obtained since M' is of full row rank and in reduced row-echelon form, and let N' be the matrix obtained from N by clearing denominators in each column. We then compute the row Hermite normal form

$$\begin{pmatrix} H \\ 0_{k\times n} \end{pmatrix} = \begin{pmatrix} T_H \\ T_0 \end{pmatrix} N', \tag{6.2.1}$$

where

$$T = \begin{pmatrix} T_H \\ T_0 \end{pmatrix}$$

is a unimodular transformation matrix with $T_0 \in \mathbb{Z}^{k\times n}$. It follows that the rows of T_0 form a basis of the row space of M, and thus an integral basis of the kernel. We use this matrix as the integral version of M', and we can now further improve the bit size of the rows using the LLL algorithm.

In the implementation, we also use this reduction approach to check heuristically whether we are using enough precision to compute X^* and to round the entries of M'. In practice, the vectors obtained after reduction will still be approximate kernel vectors of X^* if and only if we used high enough precision.

6.3. Transforming the problem

We now describe the affine hull of $\mathcal{F}$ using a coordinate transform; see [**109**] and [**48**, Section 31.5]. Let B be a rational, invertible matrix for which the first k rows form a basis of the kernel. Applying this basis transformation to any matrix X in the relative interior of the optimal face gives a matrix of the form

$$BXB^{\mathsf{T}} = \begin{pmatrix} 0 & 0 \\ 0 & \widehat{X} \end{pmatrix},$$

where $\widehat{X}$ lies in the cone $\mathbb{S}^{n-k}_{++}$ of positive definite matrices. Let $\widehat{A}_i$ be the block of $B^{-\mathsf{T}} A_i B^{-1}$ corresponding to $\widehat{X}$, so that

$$\langle \widehat{A}_i, \widehat{X} \rangle = \langle A_i, X \rangle.$$

With

$$\widehat{\mathcal{L}} = \{\widehat{X} \in \mathbb{S}^{n-k} : \langle \widehat{A}_i, \widehat{X} \rangle = b_i \text{ for } i = 1, \ldots, m\},$$

each element in the spectrahedron $\mathbb{S}^{n-k}_{+} \cap \widehat{\mathcal{L}}$ corresponds to an optimal solution in the original semidefinite program. We can now obtain from X^* a positive definite matrix $\widehat{X}^*$ close to $\widehat{\mathcal{L}}$, and projecting into $\widehat{\mathcal{L}}$ gives a matrix that corresponds to an exact optimal solution.

To find the basis transformation matrix B, we consider two methods. In our experiments, both methods work well, and it is not clear which of these two methods is better.

The first option is to extend the basis of the kernel with linearly independent standard basis vectors. We can find these vectors by using Gram-Schmidt orthogonalization in floating-point arithmetic.

The second option is to extend the basis by using the rows of the matrix T_H from (6.2.1). Then B consists of the LLL reduced rows of T_0 and the rows of T_H, so that B is unimodular. To obtain a matrix B with small bit size it is desirable that T_H should have small entries. To achieve this we use the method of [**74**]: instead of (6.2.1), we consider the Hermite normal form of the matrix N' with an identity matrix appended to it. This decomposition amounts to

$$\begin{pmatrix} H & T_H \\ 0 & T_0 \end{pmatrix} = \begin{pmatrix} T_H \\ T_0 \end{pmatrix} \begin{pmatrix} N' & I_n \end{pmatrix},$$

where now T_0 is in row Hermite normal form and T_H is reduced with respect to T_0.

The basis transformation can be seen as a form of facial reduction (see for instance the PhD thesis [**121**]). However, typically facial reduction techniques are used to obtain a problem with a strictly feasible solution, or to reduce the size of the semidefinite program before solving, for better numerical results. Instead, we first solve the problem to high precision, and then use the numerically optimal solution to *completely* describe the optimal face.

6.4. Rounding the solution

Given an approximate solution x^* to a linear system $Ax = b$, we want to find an exact solution $\bar{x}$ close to x^*. In the rounding procedure, the linear system comes from the constraints $\langle \widehat{A}_i, \widehat{X} \rangle = b_i$ for $i = 1, \ldots, m$. For simplicity of presentation, we assume first that the rows of A are linearly independent (and will address the linearly dependent case shortly).

We start by replacing each entry of x^* by its closest rational number given some fixed denominator (say, 10^{40}).

In the approach of [**51**], the system $Ax = b$ is first converted to row echelon form, after which the exact solution $\bar{x}$ is found using back substitution, where for the non-pivot columns, the corresponding entries of x^* are used. If the conditioning of A is not too bad, then $\bar{x}$ will be close to x^*.

The problem with this approach is that computing the row-echelon form of the system $Ax = b$ has to be done in exact arithmetic and becomes too expensive for large systems. In fact, depending on how the semidefinite program is formulated (for example, using low-rank constraint matrices as we do in this thesis), forming the matrix A can already be relatively expensive.

Another approach is to compute the solution $\bar{x}$ to $Ax = b$ closest to x^* as follows. The minimum norm solution of $A\varepsilon = b - Ax^*$ is given by $\varepsilon = A^+(b - Ax^*)$, where $A^+ = A^\mathsf{T}(AA^\mathsf{T})^{-1}$ is the pseudoinverse (which has this expression since A is of full row rank). Then $\bar{x} = x^* + \varepsilon$, where $\varepsilon = A^\mathsf{T} y$ and y is the solution to

$$AA^\mathsf{T} y = b - Ax^*,$$

which can be computed efficiently using Dixon's algorithm [**50**].

This approach, however, also requires the construction of the matrix A, and in practice we find that the determinant of AA^T can have large bit size, which leads

to a solution $\bar{x}$ of large bit size. For example, with this approach the universal optimality proofs in Section 8.4 each have size around 11GB, which is about a factor 500 larger than with the following method.

We first build an invertible matrix S by randomly selecting r linearly independent columns of A, where $A \in \mathbb{Q}^{r\times s}$. To do so efficiently, we first select a submatrix containing ℓ random columns of A with $r \leq \ell \leq s$, turn this submatrix into an integral matrix by clearing the denominators of each row, and then perform row reduction modulo some large prime number p. If the reduced matrix has r pivots, these pivots yield r linearly independent columns in A, and otherwise we retry with a different set of initial columns (for the problems considered in this thesis we used $\ell = 10r$ and never needed to retry).

If the assumption that the rows of A are linearly independent does not hold, it will be impossible to find S with r linearly independent columns. In that case, we must remove linearly dependent constraints before we can construct S. To find the constraints that need to be removed, we use column reduction modulo a prime.

This overall approach works best when s is much larger than r, which is the case in our applications, because it is now no longer necessary to construct the matrix A: we can compute ε by solving

$$S\varepsilon = b - Ax^*$$

using Dixon's method, and Ax^* can often be computed efficiently without having to construct A (in our applications we use the low-rank structure to do this).

Although this approach is very fast, the conditioning of S tends to be bad, which leads to $\bar{x}$ being too far from x^*, in which case some of the strictly positive eigenvalues may become negative. We therefore interpolate between the two approaches described above, where the first approach leads to nearby solutions with large bit size, and the second to far away solutions of small bit size.

To interpolate we extend S by some randomly chosen columns of A (say, $r/10$ many columns). We can then compute $\varepsilon = S^+(b - Ax^*)$ as $\varepsilon = S^\mathsf{T} y$, where y is the solution to $SS^\mathsf{T} y = b - Ax^*$. This approach works well in practice.

6.5. Rounding to algebraic numbers

Although the overall approach does not change when rounding over an algebraic number field F of degree $d > 1$, some steps need small modifications. We give only a brief description of these modifications since the most important change is already considered in [**51**], and for the proofs in this thesis we need only $d = 1$.

Note first that, since the reduced row-echelon form can be computed over any field, we can still find the kernel vectors by recognizing the entries of the reduced row-echelon form M' of M as elements of F using the LLL algorithm.

However, our reduction approach in Section 6.2 works only over $\mathbb{Q}$ due to the use of the Hermite normal form. We therefore apply the reduction method to the matrix $M'' \in \mathbb{Q}^{dk\times dn}$ obtained from $M' \in F^{k\times n}$ by using a basis of F over $\mathbb{Q}$. The rows of M'' are linearly independent over $\mathbb{Q}$ if and only if the rows of M' are linearly independent over F. After applying the reduction, we convert back to a matrix in $F^{dk\times n}$ and find a subset of k linearly independent vectors over F using floating point arithmetic.

To find a solution to the system $Ax = b$ close to x^*, with A, b, and x defined over F, we use the same approach as [**51**]. Let g be a generator of F, and consider

the expansions $A = \sum_i A_i g^i$, $b = \sum_i b_i g^i$, and $x = \sum_i x_i g^i$. Then

$$\sum_{i,j} A_i x_j g^{i+j} = \sum_l b_l g^l,$$

which defines a system of equations $A'x' = b'$ with $A' \in \mathbb{Q}^{dr \times ds}$, and where x' and b' are the concatenations of x_i and b_i for $i = 0, \ldots, d-1$. A numerical approximation of x' can be found by adding the constraints $x^* = \sum_i x_i g^i$ and solving the system using floating-point arithmetic, after which Section 6.4 can be applied.

6.6. Verification

To verify that the rounded solution indeed is a feasible solution to the semidefinite program, we check that the affine constraints hold, in exact arithmetic, and that the block-diagonal matrix is positive semidefinite. To check positive semidefiniteness, the rounding procedure writes each diagonal block in the form BXB^T, where B is a rectangular exact matrix, and X is positive definite. We check that X is indeed positive definite by computing the Cholesky decomposition in rigorous ball arithmetic.

6.7. Finding the algebraic number field

Although for the problems in this thesis it was always easy to guess over which field the optimal face can be defined, this is in general not the case. Therefore we propose the following heuristic to find the field.

Let M' be the reduced row-echelon form of the matrix which has the numerical kernel vectors as rows. The entries of M' are approximations of elements of the field F that we want to find. For each entry we can use the LLL algorithm to find the minimal polynomial of the algebraic number it approximates.

In the heuristic we first find the minimal polynomials for a selection of the entries of M', and let p be the minimal polynomial of highest degree and lowest bit size. Then we iterate through the same selection of entries and use the LLL algorithm to check whether each entry lies in the field with minimal polynomial p. If we find an entry that does not lie in the field, we replace p by the minimal polynomial of a linear combination (typically just the sum) of an approximate root of p and that entry. In practice, this method often gives the correct field F, even when considering only a small selection of entries of M'.

6.8. General applicability of the rounding procedure

Although our rounding procedure has been developed to obtain rigorous bounds in discrete geometry, we expect this approach to be useful more generally. To illustrate its use, we consider two examples from the literature where sums-of-squares characterizations are used, so that the semidefinite program does not have a unique optimal solution and the naive rounding approach mentioned in the introduction of this chapter cannot be used. Moreover, the optimal faces in these examples cannot be defined over $\mathbb{Q}$, so that we need Section 6.7.

Consider the examples from [**119, 58**] and [**118**] of finding the global minimum of the polynomial

$$u^4 + v^4 + t^4 - 4uvt + u + v + t$$

and the minimum of Caprasse's polynomial

$$-ut^3 + 4vt^2w + 4utw^2 + 2vw^3 + 4ut + 4t^2 - 10vw - 10w^2 + 2$$

over the domain $[-1/2, 1/2]^4$. Using semidefinite programming and the rounding procedure, the correct number fields (of degree 6 and 2, respectively) as well as exact minimizers are found in seconds. Interestingly, the minimal polynomials of the number field generators found by the procedure are much simpler than those of the minima in these optimization problems (smaller by a factor 1000 in terms of bit size).

6.9. Implementation

We have implemented the rounding procedure in Julia [**8**] using the computer algebra system Nemo [**56**]. The code is available as part of the open-source semidefinite programming solver `ClusteredLowRankSolver.jl`, which is available on Github[†] and can be installed as a Julia package; a snapshot of the git repository is also available at [**37**]. The three-point bound for spherical codes is included as an example. The documentation can also be found there, including a tutorial.

[†]`https://github.com/nanleij/ClusteredLowRankSolver.jl`

Part II

Applications

CHAPTER 7

Spherical codes

7.1. Introduction

In this chapter, we consider spherical code problems. A *spherical code* is a finite set $C \subseteq S^{n-1}$. The *minimum distance* of a spherical code is

$$d_{\min}(C) = \min_{\substack{x,y \in C \\ x \neq y}} d(x,y),$$

where $d(x,y) = \sqrt{2 - 2\langle x, y \rangle}$ is the chordal distance.

Given a dimension n and a number of points N, the spherical code problem asks for a spherical code with the largest $d_{\min}(C)$ among all spherical codes C of size N. Equivalently, we may consider the maximum inner product between two distinct points in the code, since they are related by a monotone function. In terms of the inner products, the problem asks for a spherical code which minimizes

$$\max_{\substack{x,y \in C \\ x \neq y}} \langle x, y \rangle$$

among spherical codes C of size N.

In general, little is known about the optimal configurations. In dimension 1 and 2, the optimal solutions are trivial. For $N < 2n$, the optimal solutions consist of two families: the optimal spherical code for $N \leq n+1$ is given by a regular simplex around the origin with maximum inner product $-1/(N-1)$, and for $n+2 \leq N \leq 2n$ an optimal configuration is given by a subset of the vertices of the regular cross polytope (consisting of n orthogonal pairs of antipodal points). In dimension 3, geometric proofs are known for the optimal solutions with $N \leq 14$ and $N = 24$.

The other known cases use linear or semidefinite programming bounds in the optimality proofs. These are bounds on the 'transposed' problem: Given a dimension n and the maximum inner product $\cos\theta$, what is the maximum number of points N such that there is a spherical code $C \subseteq S^{n-1}$ of N points with $\langle x, y \rangle \leq \cos\theta$ for all distinct $x, y \in C$? We will refer to this problem as the spherical cap packing problem.

Linear and semidefinite programming bounds on this problem are typically derived by taking into account constraints on the distribution of k-tuples, and are therefore called k-point bounds. Except for the case $(n, N, \cos\theta) = (4, 10, 1/6)$, which requires a 3-point bound, optimality was proven using the Delsarte-Goethals-Seidel linear programming bound [**47**], a 2-point bound. This bound is equivalent to the first level of the Lasserre hierarchy [**92**]. These optimal spherical codes are either the vertices of a regular polytope whose faces are simplices, or configurations derived from the E_8 root lattice or the Leech lattice [**34**], which give the only known optimal sphere packings in dimension $n > 3$ [**140, 35**].

TABLE 7.1.1. The spherical codes we prove are optimal. Each consists of N points in $\mathbb{R}^n$, with minimal angle θ. The first three are spectral embeddings from the Gewirtz, Hoffman-Singleton and M_{22} strongly regular graphs (see, e.g., [**53**]), and the next five are derived from Kerdock binary codes [**27, 77, 81**]. The last line lists two configurations with the same minimal angle, one of which was found by Mackay [**104**], while the other can be found in the data set [**32**]. See Section 7.5 for constructions of these codes.

n	N	$\cos\theta$	Other cosines
20	56	1/15	$-2/5$
21	50	1/21	$-3/7$
21	77	1/12	$-3/8$
4	24	1/2	$-1, -1/2, 0$
16	288	1/4	-1, $-1/4$, 0
64	4224	1/8	-1, $-1/8$, 0
256	66048	1/16	-1, $-1/16$, 0
1024	1050624	1/32	-1, $-1/32$, 0
4	12	1/4	$-3/4, -1/2, 0$ or $-5/8, -3/4, -1/3, 1/8$

In this chapter, we use three and four-point semidefinite programming bounds to derive new optimality proofs for spherical codes. In this way, we significantly increase the number of known optimal spherical codes. In particular, we prove that the codes in Table 7.1.1 are optimal. Additionally, we prove optimality for several hypothetical codes coming from possible parameters for triangle-free strongly regular graphs, see Table 7.1.2. These graphs may very well not exist, but if they do, they give an optimal spherical code. Besides these codes, we give a new proof that the code with 9 points on S^2 is optimal, using the three-point bound. This code was already known to be optimal by geometric arguments. We conjecture the same is possible for the optimal code with 24 points on S^2. The bound is numerically sharp, but we were not able to round the numerically sharp solution to an exact sharp solution in that case.

Of the codes in Table 7.1.1, the last line is especially interesting. For many problems, there are 0, 1 or an infinite number of solutions. For the spherical code problem in dimension 4 with 12 points, it turns out there are exactly two solutions, up to isometry. This seems to be the only example with multiple optimal configurations which contain different inner products. In [**26**], a family of configurations is given of which some are non-unique optimal spherical codes. However, all these configurations have the set of inner products between distinct points.

A special case of the spherical cap packing problem is the kissing number problem: What is the maximum number of non-overlapping unit spheres $k(n)$ that can simultaneously touch a central unit sphere? For a kissing configuration, the set of points where the outer spheres touch the central sphere is a spherical code with maximum inner product at most $\cos(\pi/3) = \frac{1}{2}$. Famously, the kissing number in dimension 3 was subject to a discussion between Gregory and Newton in 1694. Eventually it was resolved in 1953 by Schütte and Van der Waerden to be 12

TABLE 7.1.2. Hypothetical spherical codes that are optimal if they exist. Each consists of N points in $\mathbb{R}^n$, with angles θ and ψ between distinct points, and corresponds to a strongly regular graph with the given parameters.

n	N	$\cos\theta$	$\cos\psi$	Parameters
55	176	$1/25$	$-7/25$	$(176, 25, 0, 4)$
55	210	$1/22$	$-3/11$	$(210, 33, 0, 6)$
56	162	$1/28$	$-2/7$	$(162, 21, 0, 3)$
56	266	$1/20$	$-4/15$	$(266, 45, 0, 9)$
115	392	$3/115$	$-5/23$	$(392, 46, 0, 6)$
120	352	$1/45$	$-2/9$	$(352, 36, 0, 4)$
143	352	$1/65$	$-3/13$	$(352, 26, 0, 2)$
1520	3250	$1/456$	$-8/57$	$(3250, 57, 0, 1)$

[**129**]. One possible optimal configuration comes from the densest sphere packing in dimension 3, the so-called 'cannonball' configuration. Interestingly, this is a rigid configuration, but it is not the optimal solution to the spherical code problem with $N = 12$: the minimum distance of the icosahedron is larger, so it gives a non-rigid optimal configuration for the kissing number problem. In dimension 4, the conjectured optimum was the D_4 root system (equivalently, the vertices of the 24-cell), which is also rigid. Musin [**114**] proved that indeed $k(4) < 25$, so the D_4 root system is an optimal kissing configuration. In Section 7.6.2, we prove that the D_4 root system is the unique optimal spherical code of size 24, up to isometry. In particular, this implies that it is the *unique* optimal kissing configuration.

7.2. The three-point bound*

Most of the best upper bounds for kissing numbers are obtained using the three-point bound by Bachoc and Vallentin [**2**]. The three-point bound depends on the ambient dimension n, the minimal angle θ of the spherical code, and a parameter d we call the *degree* of the bound (higher degrees yield improved bounds, at increased computational expense). For each k from 0 to d, let $Y_k^n(u, v, t)$ be the $(d-k+1) \times (d-k+1)$ matrix with entries

$$Y_k^n(u,v,t)_{ij} = u^i v^j (1-u^2)^{k/2} (1-v^2)^{k/2} P_k^{n-1}\left(\frac{t-uv}{\sqrt{(1-u^2)(1-v^2)}}\right)$$

for $0 \le i, j \le d-k$, where P_k^n is the Gegenbauer polynomial of degree k with parameter $n/2-1$, normalized so that $P_k^n(1) = 1$; note that it is convenient to index these entries starting with zero. Set

$$S_k^n = \frac{1}{6} \sum_{\sigma \in S_3} \sigma Y_k^n,$$

where the permutation group S_3 permutes the three arguments of Y_k^n.

*This section is based on Section 3 of the publication "*H. Cohn, D. de Laat and N. Leijenhorst,* Optimality of spherical codes via exact semidefinite programming bounds*, 2024, arXiv:2403.16874*".

The three-point bound involves optimizing over the choice of auxiliary functions F and f, defined by

$$F(u,v,t) = \sum_{k=0}^{d} \langle F_k, S_k^n(u,v,t)\rangle \text{ and } f(u) = \sum_{k=0}^{2d} f_k P_k^n(u),$$

where $F_0, \dots, F_d$ are symmetric matrices with F_k being a $(d-k+1) \times (d-k+1)$ matrix, and $f_k \in \mathbb{R}$.

The Bachoc-Vallentin three-point bound (with some variables removed) then reads

$$\begin{aligned} &\text{minimize} && 1 + f(1) + \langle F_0, J\rangle && \\ &\text{subject to} && f(u) + 3F(u,u,1) \le -1 && \text{for } u \in [-1, \cos\theta], \\ &&& F(u,v,t) \le 0 && \text{for } (u,v,t) \in \Delta, \\ &&& F_0, \dots, F_d \succeq 0, && \\ &&& f_0, \dots, f_{2d} \ge 0, && \end{aligned} \tag{7.2.1}$$

where J is the $(d+1) \times (d+1)$ all-ones matrix,

$$\Delta = \{(u,v,t) : -1 \le u, v, t \le \cos\theta \text{ and } 1 + 2uvt - u^2 - v^2 - t^2 \ge 0\},$$

and $M \succeq 0$ means M is a positive-semidefinite matrix.

We can see as follows that the objective function $1 + f(1) + \langle F_0, J\rangle$ is an upper bound for spherical code size. Let C be a spherical code in S^{n-1} with minimal angle at least θ, and fix a point $e \in S^{n-1}$. If $F_k \succeq 0$, then it follows from the addition formula for spherical harmonics that

$$(x,y) \mapsto \langle F_k, Y_k^n(\langle x,e\rangle, \langle y,e\rangle, \langle x,y\rangle)\rangle$$

is a positive semidefinite kernel. This implies that if F is feasible (i.e., it satisfies the constraints in the optimization problem), then

$$\sum_{x,y,z \in C} F(\langle x,z\rangle, \langle y,z\rangle, \langle x,y\rangle) \ge 0.$$

Similarly, $(x,y) \mapsto P_k^n(\langle x,y\rangle)$ is a positive definite kernel by the addition formula, so that for $a_k \ge 0$ we have

$$\sum_{x,y \in C} a_k P_k^n(\langle x,y\rangle) \ge 0.$$

Summing the two inequalities, separating terms and using the affine constraints gives

$$\begin{aligned} 0 &\le \sum_{x,y \in C} f(\langle x,y\rangle) + \sum_{x,y,z \in C} F(\langle x,z\rangle, \langle y,z\rangle, \langle x,y\rangle) \\ &= |C|(f(1) + F(1,1,1)) + \sum_{\substack{x,y \in C \\ x \ne y}} (f(\langle x,y\rangle) + 3F(\langle x,y\rangle, \langle x,y\rangle, 1)) \\ &\quad + \sum_{\substack{x,y,z \in C \\ x,\, y,\, z \text{ distinct}}} F(\langle x,z\rangle, \langle y,z\rangle, \langle x,y\rangle) \\ &\le |C|(f(1) + F(1,1,1)) - |C|(|C| - 1), \end{aligned}$$

from which it follows that

$$|C| \leq 1 + f(1) + F(1,1,1) = 1 + f(1) + \langle F_0, J \rangle, \tag{7.2.2}$$

because $S_k^n(1,1,1) = 0$ for $k > 0$ and $S_0^n(1,1,1) = J$.

This argument shows that (7.2.1) gives an upper bound for the maximal cardinality of a spherical code with minimal angle at least θ. If this bound matches the code size, then the inequalities in the above derivation must hold with equality, which implies that $f(\langle x, y \rangle) + 3F(\langle x, y \rangle, \langle x, y \rangle, 1) = -1$ for all distinct $x, y \in C$. In other words, the zeros of the polynomial $f(u) + 3F(u, u, 1) + 1$ are the only inner products other than 1 that can appear in a maximal code. This condition is a special case of complementary slackness in semidefinite programming.

So far, we have shown that every code C with minimal angle at least θ must satisfy $|C| \leq 1 + f(1) + \langle F_0, J \rangle$, but not yet that every code achieving this bound must be an optimal spherical code. To see why this must be the case, note first that the polynomial $f(u) + 3F(u, u, 1) + 1$ cannot vanish identically, since if it did, then (7.2.2) would imply that

$$|C| \leq 1 + f(1) + F(1,1,1) = 1 + (-1 - 3F(1,1,1)) + F(1,1,1) = -2F(1,1,1) \leq 0.$$

Therefore complementary slackness shows that there are only finitely many possible inner products between distinct points in C, namely the roots of $f(u)+3F(u,u,1)+1$. If there were a code C with $|C| = 1 + f(1) + \langle F_0, J \rangle$ and minimal angle strictly greater than θ, then every sufficiently small perturbation of C would be a code with minimal angle at least θ, which would contradict the finite list of possible inner products. Thus, codes achieving equality in the three-point bound must be optimal codes.

To model the optimization problem (7.2.1) as a semidefinite program, we use sum-of-squares polynomials and symmetry reduction, following Chapter 4. For the two point constraint, we have the semialgebraic set $\Delta_2 = \{u : p(u) \geq 0\}$, where $p(u) = (1+u)(\cos\theta - u)$. For the three-point constraint, the semialgebraic set reads

$$\Delta_3 = \{(u, v, t) : p(u) \geq 0, p(v) \geq 0, p(t) \geq 0, 1 + 2uvt - u^2 - v^2 - t^2 \geq 0\};$$

these polynomials certify that the Gram matrix of 3 points with inner products u, v and t is positive semidefinite. Using the approach in Section 4.3, we can find an S_3-invariant description of Δ_3, and since $F(u, v, t)$ is S_3-invariant, this allows us to use (4.3.1) with S_3-invariant sums-of-squares polynomials. Following Example 3.4, we obtain a block-diagonalized description for the sum-of-squares polynomials, and using sampling (see Section 4.2) gives a semidefinite program with low-rank structure, which can be solved using the solver of Chapter 5.

7.3. The Lasserre hierarchy

In this section we specify the Lasserre hierarchy for spherical codes with a fixed minimal distance. This is a recap and continuation of the running example of Part I, Example 2.1.

Take $V = S^{n-1}$, with distinct $x, y \in V$ adjacent when $\langle x, y \rangle > \cos\theta$. This is a topological packing graph, so Chapter 2 applies. We wish to find the maximum size of an independent set in this graph, so the objective function is given by $\lambda(\mathcal{I}_{=1})$, where $\mathcal{I}_{=k}$ is the set of all independent sets of size equal to k. There are no additional

constraints, which gives the primal problem

$$\begin{aligned} \text{maximize} \quad & \lambda(\mathcal{I}_{=1}), \\ \text{subject to} \quad & A_k^*(\lambda) \in \mathcal{M}(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}, \\ & \lambda \in \mathcal{M}(\mathcal{I}_{2k})_{\geq 0}, \\ & \lambda(\mathcal{I}_0) = 1 \end{aligned}$$

where, as explained in Chapter 2, $\mathcal{I}_k$ is the disjoint union of $\mathcal{I}_{=i}$ for $i = 0, \ldots, k$, and A_k^* is the adjoint of the operator $A_k : C(\mathcal{I}_k \times \mathcal{I}_k) \to C(\mathcal{I}_{2k})$ defined by

$$A_k K(S) = \sum_{\substack{J_1, J_2 \in \mathcal{I}_k \\ J_1 \cup J_2 = S}} K(J_1, J_2).$$

Dualizing, and slightly reformulating the dual as in equation (2.4.3), then gives the problem

$$\begin{aligned} \text{minimize} \quad & K(\emptyset, \emptyset), \\ \text{subject to} \quad & A_k K(S) \leq \chi_{\mathcal{I}=1}(S) \quad S \in \mathcal{I}_{2k} \setminus \mathcal{I}_0, \\ & K \in C(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}. \end{aligned} \tag{7.3.1}$$

We denote problem (7.3.1) for dimension n and maximum inner product $\cos\theta$ by $\mathrm{las}_k(n, \cos\theta)$.

The problem is invariant under $O(n)$, so we may assume that K is $O(n)$-invariant. For $k = 2$, such kernels are described in Section 3.4. The general form of such kernels is given by

$$K(J_1, J_2) = \sum_{|\lambda| \leq d_1} \langle K_\lambda, Z_\lambda(J_1, J_2) \rangle$$

where the entries of Z_λ are polynomials, and we restrict to the rows and columns of Z_λ with entries whose degree is bounded by d_2. To compute the zonal matrices Z_λ for $k = 2$, we can use the method described in Section 3.4 up to $d_1 = 14$; for higher d_1 we require the faster method introduced in [**88**].

The entries of $Z_\lambda(J_1, J_2)$ are polynomials in the inner products between vectors in $J_1 \cup J_2$, so on the sets $\mathcal{I}_{=i}$ we obtain polynomial inequality constraints. Using Chapter 4 and Example 3.4, we relax the constraints to semidefinite programming constraints; for this we need to choose a maximum degree δ of the sum-of-squares polynomials. The choices for the computations with the second level of the Lasserre hierarchy are recorded in Table 7.3.1. Note that d_2 and δ are the degrees of the polynomials occurring in the bound, in contrast to the three-point bound where the polynomial degree equals $2d$.

In this thesis we take $d_1 \leq d_2 \leq \delta$, with δ even. As a frame of reference, solving a semidefinite program with these parameters and 256 bits precision takes about 4 hours for $\delta = 12$ and 1 day for $\delta = 14$, using a computer with 8 cores and 256Gb of RAM. The computations with $\delta = 16$ typically take 2 weeks, which we only use for exceptional cases.

For the uniqueness proofs, we first specify Theorem 2.10 to the case of the Lasserre hierarchy for spherical codes.

COROLLARY 7.1 (Complementary slackness). *Let K be a feasible solution to $\mathrm{las}_k(n, \cos\theta)$, and let $C \subseteq S^{n-1}$ be a code with minimum angle at least θ. If $K(\emptyset, \emptyset) = |C|$, then*

TABLE 7.3.1. The degrees d_1, d_2 and δ we use to obtain upper bounds using the second step of the Lasserre hierarchy.

n	lb	ub	$\cos\theta$	d_1	d_2	δ	sharp
4	24	24	1/2	14	16	16	sharp
4	12	12	1/4	16	16	16	sharp
6	72	77.85	1/2	14	16	16	not sharp

- *For all $S \subseteq C$ with $1 \leq |S| \leq 2k$,*
$$A_k K(S) = -\chi_{\mathcal{I}_{=1}}(S).$$
- *For all λ,*
$$\langle K_\lambda, \sum_{\substack{J_1, J_2 \subseteq C \\ |J_1|, |J_2| \leq k}} Z_\lambda(J_1, J_2) \rangle = 0$$

Note that the second condition can also be written as
$$\sum_{S \subseteq C, |S| \leq 2k} A_k K^\lambda(S) = 0$$
for K^λ defined by
$$(J_1, J_2) \mapsto \langle K_\lambda, Z_\lambda(J_1, J_2) \rangle.$$

7.4. Improved kissing number bounds[†]

Initial computations with the three-point bound for the kissing number problem were performed by Bachoc and Vallentin using `CSDP` [**12**], but since this is a machine precision solver it was not possible to go beyond $d = 10$ [**2**]. Mittelman and Vallentin [**108**] then used the high precision solvers `SDPA-QD` and `SDPA-GMP` to perform computations up to $d = 14$. Later Machado and Oliviera [**103**] applied symmetry reduction and used `SDPA-GMP` to compute bounds up to degree $d = 16$.

Because of the low-rank structure of the problem and the specialized solver of Chapter 5, we can perform computations up to $d = 20$ within a reasonable time frame. We estimate that the approach from [**103**] using `SDPA-GMP` would have been slower by a factor 40 for $d = 20$. In Table 7.4.1 we show the kissing number bounds for $d = 16, \ldots, 20$ for dimensions up to 24. Dimension 2, 8, and 24 are omitted since the linear programming bound is sharp in these dimensions. After rounding down to the nearest integer, this improves the best known upper bounds in dimensions 11 through 23.

Rigorous verification of these bounds can be done using standard interval-arithmetic techniques (see, e.g., [**91, 103**]). Alternatively, the rounding procedure of Chapter 6 would work after recomputing the bound for a fixed (slightly larger) rational objective. We did not perform this verification procedure since our main goal here is to show that our approach enables us to significantly increase the degree

[†]Parts of this section are adapted from the publications "*N. Leijenhorst and D. de Laat,* Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput.* ***16*** *(2024), no. 3, 503-534, doi:10.1007/s12532-024-00264-w*" and "*D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer,* Optimality and uniqueness of the D_4 root system*, 2024, arXiv:2404.18794*".

n	lower bound	d	upper bound	n	lower bound	d	upper bound
3	12	16	12.368580	14	1932	16	3177.7812
		17	12.364503			17	3176.4354
		18	12.360782			18	3175.3519
		19	12.357869			19	3174.7746
		20	12.353979			20	3174.1890
4	24	16	24.056877	15	2564	16	4858.1937
		17	24.053495			17	4856.4186
		18	24.051431			18	4855.1064
		19	24.048769			19	4854.3872
		20	24.047205			20	4853.7561
5	40	16	44.981014	16	4320	16	7332.7695
		17	44.976437			17	7329.8545
		18	44.973846			18	7325.5713
		19	44.971353			19	7322.5461
		20	44.970252			20	7320.1068
6	72	16	78.187644	17	5730	16	11014.169
		17	78.173268			17	11004.299
		18	78.163358			18	10994.873
		19	78.151981			19	10984.895
		20	78.143569			20	10978.622
7	126	16	134.26988	18	7654	16	16469.091
		17	134.21522			17	16445.457
		18	134.17305			18	16431.764
		19	134.13115			19	16418.296
		20	134.10709			20	16406.358
9	306	16	363.67296	19	11692	16	24575.872
		17	363.59590			17	24516.534
		18	363.50742			18	24463.542
		19	363.41738			19	24443.476
		20	363.34567			20	24417.472
10	510	16	553.82278	20	19448	16	36402.676
		17	553.57125			17	36296.753
		18	553.38179			18	36250.908
		19	553.21188			19	36218.806
		20	553.05527			20	36195.348
11	592	16	869.23401	21	29768	16	53878.723
		17	868.82650			17	53724.682
		18	868.45366			18	53647.201
		19	868.15131			19	53567.621
		20	868.01070			20	53524.085
12	840	16	1356.5778	22	49896	16	81376.460
		17	1356.1536			17	81085.186
		18	1355.8837			18	80962.164
		19	1355.4776			19	80860.092
		20	1355.2976			20	80810.158
13	1154	16	2066.3465	23	93150	16	123328.40
		17	2065.5348			17	122796.10
		18	2064.9493			18	122657.49
		19	2064.4859			19	122481.07
		20	2064.0029			20	122351.67

TABLE 7.4.1. Three-point bounds for the kissing number problem in dimensions 3-23. Dimension 8 is omitted since there the linear programming bound is sharp. New records after rounding down to the nearest integer are underlined. Lower bounds are taken from [**40**].

of the polynomials used, thereby obtaining better bounds. Note that the bounds we

report for $d = 16$ are slightly different from the bounds reported in [**103**] since their verification procedure increases the bounds by a configurable parameter $\varepsilon > 0$.

We also compute the second step of the Lasserre hierarchy for the kissing number problem in dimensions $4 - 7, 10, 12$ and 16. This gives a sharp bound in dimension 4 (see Section 7.6.2), and improves the bound in dimension 6 from $k(6) \leq 78$ to $k(6) \leq 77$. Since this is, relative to the bound, a large improvement, we do verify this bound. For these bounds we use the parameters $d_1 = 14$ and $d_2 = \delta = 16$.

In dimension 6, the bound is not sharp, so that the optimal objective value and optimal solution potentially require high algebraic degree or bit size. This means the rounding procedure from Chapter 6 may not be able to find an exact optimal solution here. Therefore, we solve the problem as a feasibility problem, where we add the constraint that the objective $K(\emptyset, \emptyset)$ is equal to 77.85. Since this is strictly larger than the numerically computed optimal objective, the solver will return a strictly feasible solution (a feasible solution where all matrix variables are positive definite), from which it is easy to extract an exact feasible solution. This gives a rigorous proof of $k(6) \leq 77$.

We verify the proof using Section 6.6. As part of the verification procedure, the zonal matrices Z_λ need to be constructed; the method described in Section 3.4 takes less than two days on a modern computer. The much faster method of [**88**] gives the same zonal matrices up to a positive constant factor depending only on n and λ. The remainder of the verification procedure takes less than two hours.

The script and data files to perform this verification procedure are available at [**86**]. There we also make available the implementation we used for generating the proofs. Our scripts are written in Julia [**8**] and use the Nemo computer algebra system [**56**].

7.5. Constructions of spherical codes[‡]

In Section 7.6, we prove optimality and uniqueness of certain spherical codes. In this section, we give constructions for these codes.

7.5.1. Spectral embeddings of triangle-free strongly regular graphs. Recall that a *strongly regular graph with parameters* (n, k, λ, μ) is an n-vertex graph, not a complete graph or its complement, such that every vertex has degree k, every pair of adjacent vertices has λ common neighbors, and every pair of distinct, non-adjacent vertices has μ common neighbors. One can check that such a graph is connected if and only if $\mu > 0$, and we will always assume this is the case.

A graph is *triangle-free* if there do not exist three mutually adjacent vertices; for a strongly regular graph, this condition amounts to $\lambda = 0$. The only known connected, triangle-free strongly regular graphs are one infinite family and seven exceptional cases, namely

(1) the complete bipartite graph $K_{n,n}$, which has parameters $(2n, n, 0, n)$,
(2) the 5-cycle, which has parameters $(5, 2, 0, 1)$,
(3) the Petersen graph, which has parameters $(10, 3, 0, 1)$,
(4) the Clebsch graph, which has parameters $(16, 5, 0, 2)$,

[‡]This section is based on Section 1 of the publication "*H. Cohn, D. de Laat and N. Leijenhorst,* Optimality of spherical codes via exact semidefinite programming bounds*, 2024, arXiv:2403.16874*" and Section 3.2 of "*D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer,* Traceless projection of tensors with applications to hierarchies in discrete geometry*, In preparation*".

(5) the Hoffman-Singleton graph, which has parameters $(50, 7, 0, 1)$,
(6) the Gewirtz graph, which has parameters $(56, 10, 0, 2)$,
(7) the M_{22} graph, which has parameters $(77, 16, 0, 4)$, and
(8) the Higman-Sims[§] graph, which has parameters $(100, 22, 0, 6)$.

See [**23**] for constructions of these graphs, as well as more information and references.

Each of these graphs corresponds to a spherical code via *spectral embedding*. Specifically, let A be the graph's adjacency matrix, which is indexed by vertices. Orthogonally projecting the basis vectors indexed by these vertices into an eigenspace of A yields points on a sphere, which we can rescale to be the unit sphere.

The complete bipartite graph $K_{n,n}$ is a special case, because it is bipartite. The eigenvalues of its adjacency matrix are 0 (with multiplicity $2n-2$) and $\pm n$, and projecting into the eigenspace with eigenvalue 0 gives two orthogonal $(n-1)$-dimensional regular simplices in $\mathbb{R}^{2n-2}$, which is an optimal spherical code, but not unique for $n > 2$.

For each of the remaining cases, we orthogonally project into the eigenspace with the smallest eigenvalue. The resulting spherical code is a two-distance set (i.e., there are only two distances between distinct points), with the larger distance corresponding to adjacency in the graph. Each of these codes is either already known to be optimal (see, e.g., [**38**, Table 1.1]) or proved to be optimal in Section 7.6 (Table 7.1.1).

7.5.2. Kerdock spherical codes. Kerdock codes [**77**] are a family of binary error-correcting codes in $\{0,1\}^{2^{2k}}$ with 2^{4k} codewords and minimal Hamming distance $2^{2k-1} - 2^{k-1}$, which can be constructed using $\mathbb{Z}/4\mathbb{Z}$-linear codes [**65**]. The case $k = 1$ is trivial, while $k = 2$ is the Nordstrom-Robinson code [**117**], which is known to be not only optimal [**117**] but also unique [**134**]. For $k > 2$ it has not been known whether the Kerdock codes are optimal.

For each k, one can use the Kerdock code to construct a spherical code with $2^{4k} + 2^{2k+1}$ points in 2^{2k} dimensions and maximal inner product $1/2^k$ (see [**27, 101, 81**]). Specifically, this spherical code contains the standard orthonormal basis and its negatives as well as the 2^{4k} points $((-1)^{c_1}, \dots, (-1)^{c_{2^{2k}}})/2^k$ with $(c_1, \dots, c_{2^{2k}})$ in the Kerdock code. We call these configurations the *Kerdock spherical codes*. They are known to be optimal among antipodal spherical codes [**101**], or equivalently as point configurations in real projective space, but have not previously been known to be optimal among all spherical codes.

7.5.3. Spherical codes of size 12 in dimension 4. In this section we give the Gram matrices of two codes of size 12 in dimension 4 with maximum inner product $1/4$. In Section 7.6.3 we will prove that these are optimal spherical codes, and that in fact these are the only two optimal spherical codes with these parameters, up to isometry.

Let C_1 and C_2 be the codes with respective Gram matrices

$$\begin{bmatrix} S_{1/4} & E & F \\ E & S_{-1/3} & E \\ F & E & S_{1/4} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} G & E & E \\ E & G & E \\ E & E & G \end{bmatrix},$$

[§]Before it was rediscovered by Higman and Sims [**71**] in the process of constructing the Higman-Sims group, this graph was discovered by Mesner [**106**], as pointed out by Klin and Woldar [**79**].

where $S_\alpha = (1-\alpha)I + \alpha J$ is the Gram matrix of the simplex with pairwise inner product α, and where

$$E = \begin{bmatrix} 1/4 & -3/4 & 1/4 & 1/4 \\ -3/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & -3/4 \\ 1/4 & 1/4 & -3/4 & 1/4 \end{bmatrix}, \; F = \begin{bmatrix} 1/8 & -5/8 & -5/8 & -5/8 \\ -5/8 & 1/8 & -5/8 & -5/8 \\ -5/8 & -5/8 & 1/8 & -5/8 \\ -5/8 & -5/8 & -5/8 & 1/8 \end{bmatrix},$$

$$G = \begin{bmatrix} 1 & 0 & -1/2 & -1/2 \\ 0 & 1 & -1/2 & -1/2 \\ -1/2 & -1/2 & 1 & 0 \\ -1/2 & -1/2 & 0 & 1 \end{bmatrix}.$$

One can check that the matrices are of rank 4, which means they are indeed Gram matrices of 12-point codes in S^3. The code C_2 was first found in [**104**] and the code C_1 was first found in [**32**].

7.6. Optimality and uniqueness of spherical codes[¶]

7.6.1. New sharp three-point bounds. We compute the three-point bound for the parameters listed in Table 7.6.1, where we also list the degrees and the timing information. For the bounds in this section, we set $f = 0$, since this auxiliary function does not contribute in these cases. In each case, $|C| = 1 + \langle F_0, J\rangle$, and therefore C must be optimal. Furthermore, in each case $3F(u,u,1)+1$ has no roots in $[-1, \cos\theta]$ other than the inner products that occur in C; this observation will play a key role in proving uniqueness, because it shows that no other inner products can occur in any optimal code.

THEOREM 7.2. *The spherical codes with parameters* $(n, N, \cos\theta)$ *in the set*

$$\{(16, 288, 1/4), (64, 4224, 1/8), (20, 56, 1/15), (21, 50, 1/21), (21, 77, 1/12)\}$$

are unique up to isomorphism.

The proofs of uniqueness can be split into two types: one for antipodal (kerdock) codes, and one for strongly regular graphs.

PROOF FOR ANTIPODAL CODES. There are two codes with -1 as possible inner product in the list, with $(n, N) \in \{(16, 288), (64, 4224)\}$. Suppose C is a spherical code with one of these parameters and the maximal inner product given by Table 7.6.1. Let $T = \{-1, -\alpha, 0, \alpha\}$ be the list of inner products possible in the code by complementary slackness from the three-point bound. Then $T \cup \{1\} = -(T \cup \{1\})$, and hence C must be antipodal: Suppose $-x \notin C$ for some $x \in C$. Then $C \cup \{-x\}$ has the same list of inner products as C, and contains an extra point, which contradicts the three-point bound. Therefore, the N points lie on $N/2$ lines through the origin.

By Proposition 3.12 from [**25**], the $N/2$ lines can be partitioned in $N/(2n)$ sets of n orthogonal lines. Let one of these sets define the coordinates, then $2n$ lie on

[¶]This section is based on Section 4 of the publication "*H. Cohn, D. de Laat and N. Leijenhorst,* Optimality of spherical codes via exact semidefinite programming bounds*, 2024, arXiv:2403.16874*", Section 5.1 of "*D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer,* Optimality and uniqueness of the D_4 root system*, 2024, arXiv:2404.18794*", and Section 3.2 of "*D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer,* Traceless projection of tensors with applications to hierarchies in discrete geometry*, In preparation*".

TABLE 7.6.1. The degree d we use to obtain sharp three-point bounds, the number of digits of accuracy to which we solve the semidefinite programs, and the time in seconds it took to solve the semidefinite programs, round the solutions, and verify the proofs. For the last two cases, the three-point bound was already known to be sharp.

n	N	$\cos\theta$	d	Digits	Solve	Round	Check
16	288	1/4	8	40	205	4	2
64	4224	1/8	9	40	484	7	6
256	66048	1/16	10	40	1075	15	17
1024	1050624	1/32	18	100	219131	6147	3135
20	56	1/15	5	40	9	0.2	0.1
21	50	1/21	5	40	9	0.2	0.06
21	77	1/12	5	40	9	0.3	0.06
55	176	1/25	5	40	9	0.2	0.08
55	210	1/22	5	40	10	0.2	0.07
56	162	1/28	5	40	10	0.3	0.07
56	266	1/20	5	40	10	0.3	0.07
115	392	3/115	5	40	10	0.2	0.08
120	352	1/45	5	40	10	0.3	0.08
143	352	1/65	6	40	34	0.9	0.6
1520	3250	1/456	9	60	586	10	17
3	9	1/3	10	40	958	23	24
4	10	1/6	4	40	4	0.1	0.03
3	8	$1/(\sqrt{8}+1)$	7	60	112	210	100

the coordinate axes, and the other points have coordinates $\pm\alpha$. The signs then define a binary code $\tilde{C}$: we have $c \in \tilde{C}$ if and only if $\alpha((-1)^{c_1}, \dots (-1)^{c_n}) \in C$. The inner product between two such points is given by $1 - 2\alpha^2 d(c, c')$, where $d(c, c')$ is the Hamming distance $d(c, c') = |\{i : c_i \neq c'_i\}|$. This gives a binary code with $|\tilde{C}| = N - 2n$ and minimal Hamming distance $(1 - \alpha)/(2\alpha^2)$, and a certain distance distribution. Theorem 10.2.1 in [**134**] and Theorem 1 in [**122**] show that there are unique binary codes with these properties, up to permutation and translation, which translates into uniqueness of C up to isometry. □

PROOF FOR TRIANGLE-FREE STRONGLY REGULAR GRAPHS. This applies to the codes in dimension 21 and 22, which are spherical embeddings from the Gewirtz, Hoffman-Singleton and M_{22} strongly regular graphs.

By complementary slackness, an optimal code C of size N can have at most 2 different inner products, say $\alpha < \beta$. It therefore defines a graph with vertex set C and an edge between two vertices c, c' if $\langle c, c' \rangle = \alpha$. It suffices to show that the code must be a spherical 2-design: by Theorem 7.4 in [**47**], any spherical 2-design which is a 2-distance set yields a strongly regular graph uniquely defined by the dimension, size and distances of the design. Since the strongly regular graphs corresponding

to these spherical codes are unique up to isomorphism by [**59, 21, 73, 22**], this implies that the spherical codes are also unique, up to isometry.

Let $k = |\{(x, y) \in C^2 : \langle x, y\rangle = \alpha\}|/N$. The characterization of spherical designs in terms of Gegenbauer polynomials P_k^n of [**47**, Theorem 5.5] show that

$$kP_i^n(\alpha) + (N - k + 1)P_i^n(\beta) + P_i^n(1) \geq 0$$

for $i = 1, 2$, and C is a spherical 2-design if we have equality. In our case, N, n, α and β are determined, and one can check that the only way that k can satisfy both inequalities is when they are equalities. Therefore C is a 2-design, as desired. □

These arguments also show that any spherical code with the hypothetical parameters in Table 7.6.1 must be a spectral embedding of a strongly regular graph. However, it is unknown whether such graphs would be unique if they exist.

7.6.2. The D_4 root system. In this section, we prove that the D_4 root system is the unique optimal kissing configuration in dimension four and is an optimal spherical code.

For this we first compute a numerically optimal solution to $\mathrm{las}_2(4, 1/2)$. To get a sharp bound, we use $d_1 = 14$ and $d_2 = \delta = 16$ for the truncation of the inverse Fourier transform and the sums-of-squares degrees. The resulting semidefinite program is large, and to solve it the use of the semidefinite programming solver from Chapter 5 is essential. We compute the optimal solution to 40 digits of precision using 256-bit floating-point arithmetic. This takes about two weeks on 8 cores of a modern computer equipped with 128GB of working memory.

The next step is to round the numerical solution to an exact optimal solution using Chapter 6. Although a semidefinite program defined over the rationals does not necessarily admit a rational optimal solution (see, e.g., [**116**]), this is the case here, and the rounding procedure finds a rational optimal solution within 4 hours. This gives an exact feasible solution K with objective value $K(\emptyset, \emptyset) = 24$. We use the same verification procedure as for the kissing number in dimension 6, which is available at [**86**] together with the exact solution.

Using Sturm sequences we verify that the polynomial corresponding to $A_2K|_{\mathcal{I}_{=2}}$ has roots -1, $\pm 1/2$, and 0 in the interval $[-1, 1/2]$. That is, by complementary slackness, for distinct $x, y \in S^3$ with $\langle x, y\rangle \leq 1/2$, $A_2K(\{x, y\}) = 0$ if and only if $\langle x, y\rangle \in \{-1, \pm 1/2, 0\}$ (see Lemma 7.3). In the remainder of this section, we use this fact to show the D_4 root system is an optimal spherical code and is the unique optimal kissing configuration up to isometry.

LEMMA 7.3. *If $C \subseteq S^3$ is a subset of size 24 with minimal angle at least $\pi/3$, then*

$$\langle x, y\rangle \in \{-1, -1/2, 0, 1/2\}$$

for all distinct $x, y \in C$.

PROOF. By Corollary 7.1, we have for every $x \neq y \in C$ that

$$A_2(K)(\{x, y\}) = 0.$$

As mentioned above, for distinct $x, y \in S^3$, we have $A_2K(\{x, y\}) = 0$ if and only if $\langle x, y\rangle \in \{-1, \pm 1/2, 0\}$.

□

This shows the D_4 root system corresponds to an optimal spherical code: among the 24-point subsets of S^3, the minimal distance between distinct points is as large as possible.

THEOREM 7.4. *The D_4 root system is an optimal spherical code.*

PROOF. If there were a spherical code C of cardinality 24 with smallest angle strictly larger than $\pi/3$, then any small enough perturbation of C would correspond to a kissing configuration of size 24, which contradicts with Lemma 7.3. □

Note that for cases where the Bachoc-Vallentin three-point bound is sharp, optimality of the corresponding spherical code follows immediately, because in that case sharpness directly implies that there are only finitely many possible inner products; see Section 7.2. We do not know whether the same is always true for a truncation of the Lasserre hierarchy, since it is not clear whether the polynomial p_2 can be identically zero when the bound is sharp.

THEOREM 7.5. *The D_4 root system is the unique optimal kissing configuration in $\mathbb{R}^4$ up to isometry.*

PROOF. Let $C \subseteq S^3$ be an optimal kissing configuration in $\mathbb{R}^4$. We first verify that C is a root system.

(1) The vectors in C must span $\mathbb{R}^4$, since otherwise C would give a kissing configuration in $\mathbb{R}^3$ of size 24.
(2) Since C is a subset of the unit sphere, the only scalar multiples of $\alpha \in C$ can be α and $-\alpha$.
(3) Let $\alpha, \beta \in C$ and consider the reflection $\beta' = \beta - 2\langle \alpha, \beta \rangle \alpha$ of β through the hyperplane orthogonal to α. By Lemma 7.3, it follows that
$$\langle \beta', \gamma \rangle \in \{\pm 1, \pm 1/2, 0\}$$
for every $\gamma \in C$. So, β' must be in C by optimality of C.
(4) By Lemma 7.3, for $\alpha, \beta \in C$, the value $2\langle \alpha, \beta \rangle$ is an integer. In other words, the reflection of β through the hyperplane orthogonal to α is obtained by subtracting an integer multiple of α from β.

Hence, the set C is a root system in $\mathbb{R}^4$.

The irreducible root systems have been classified, and the only irreducible root systems where all vectors have the same length are A_j, D_j, E_6, E_7, and E_8; see, for instance, [**133**, Table 4.1]. Since all roots in C have the same length, it must be a direct sum of these irreducible root systems. In other words,

$$C = \bigoplus_{i=1}^{k} \Phi_i$$

for some k and root systems $\Phi_1, \ldots, \Phi_k$, where each Φ_k is isomorphic to A_j, D_j, E_6, E_7, or E_8.

Let us assume that D_4 does not occur in the decomposition. By considering the dimensions, the summands must be isomorphic to A_j with $1 \leq j \leq 4$ and D_j with $1 \leq j \leq 3$. We denote by r the total number of roots occurring in C, by r_i the number of roots of Φ_i, and by d_i the rank of Φ_i. For A_j we have $r_j/d_j = j + 1$ and for D_j we have $r_j/d_j = 2(j-1)$. Hence, we have $r_i/d_i < 6$ for the root systems

which occur in the decomposition. Furthermore, since the span of C is $\mathbb{R}^4$, we have $\sum_{i=1} d_i = 4$. We then have

$$r = \sum_{i=1}^{k} r_i = \sum_{i=1}^{k} \frac{r_i}{d_i} d_i < 6 \sum_{i=1} d_i = 24.$$

Since the number of roots in C is equal to 24, this gives a contradiction. Hence, C is D_4 up to orthogonal transformations. □

Note that an alternative proof can be given along the lines of Theorem 7.2: the code must be antipodal, and the corresponding set of lines is by [**25**, Proposition 3.12] a union of 3 orthonormal bases. Taking one of the bases to define the coordinates, the other points must have $\pm 1/2$ in every coordinate. The only possibility is to have every such point, so the resulting configuration is unique.

7.6.3. Two 12-point codes in dimension 4. In this section we prove that there are exactly two optimal spherical codes of size 12 in dimension 4, namely the codes C_1 and C_2 constructed in Section 7.5.3. The proof relies on an exact solution to $\mathrm{las}_2(4, 1/4)$, and this solution and the Julia code to verify the statements we make in this section about the solution are available at [**87**].

The *k-point distance distribution* of a subset $C \subseteq S^{n-1}$ assigns to each subset $S \subseteq S^{n-1}$ with $2 \leq |S| \leq k$ the size of the set $\{R \subseteq C : R \in O(n)S\}$. The elliptope is defined by

$$\mathcal{E}_n = \{X \in \mathcal{S}_+^n : X_{1,1} = \cdots = X_{n,n} = 1\},$$

where $\mathcal{S}_+^n$ is the cone of positive semidefinite matrices. We will think of the 4-point distance distribution of a subset C of S^3 as the function that assigns to each matrix in $\mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4$ the number of times this matrix appears as a principal submatrix of the Gram matrix of C up to simultaneous permutations of the rows and columns.

LEMMA 7.6. *Let $C \subseteq S^3$ be a subset of size 12. If the 4-point distance distributions of C and C_1 agree, then C is equal to C_1 up to isometry.*

PROOF. From the distance distribution we see there is a subset S of 4 points in C with pairwise inner products $-1/3$. We may assume this 3-simplex lies in the hyperplane $x_1 = 0$. By the distance distribution there are 12 pairs of points in C with inner product $-1/3$, which means any such pair is a subset of S.

By the distance distribution there are 8 subsets of C of size 4 such that 3 points have inner product $-1/3$ with each other, and the fourth point has inner product $1/4$ with the other three points in the subset. Given a subset of size 3 of S, there are exactly 2 points in S^3 that have inner product $1/4$ with those three points: one at each side of the hyperplane $x_1 = 0$. Since there are 4 different subsets of size 3 of S, this implies all points in $C \setminus S$ are of this form, and this fully describes C up to isometry. So C is equal to C_1 up to isometry. □

LEMMA 7.7. *Let $C \subseteq S^3$ be a subset of size 12. If the 4-point distance distributions of C and C_2 agree, then C is equal to C_2 up to isometry.*

PROOF. From the distance distribution we see that there are 3 subsets of C with Gram matrix G. To see that these subsets are disjoint we observe the following. The distance distribution shows that there are 6 pairs of orthogonal points, and no triples of orthogonal points. Together this implies that the 12 points can be partitioned into 6 pairs of orthogonal points. Therefore, the subsets of 4 points

either have an overlap of 2 orthogonal points, or no overlap. Since the only triple of inner products with $-1/2$ occurring at least twice is $(-1/2, -1/2, 0)$, there are no three pairs J_1, J_2, and J_3 of orthogonal points such that the inner products between the points in J_1 and J_2 as well as the inner products between the points in J_1 and J_3 are $-1/2$.

From the distance distribution we see that there are 6 subsets of 4 points with Gram matrix

$$\begin{bmatrix} 1 & 0 & 1/4 & 1/4 \\ 0 & 1 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1 & 0 \\ 1/4 & 1/4 & 0 & 1 \end{bmatrix}.$$

and 6 subsets of 4 points with Gram matrix

$$\begin{bmatrix} 1 & 0 & -3/4 & 1/4 \\ 0 & 1 & 1/4 & -3/4 \\ -3/4 & 1/4 & 1 & 0 \\ 1/4 & -3/4 & 0 & 1 \end{bmatrix}.$$

This implies the Gram matrix of C must be of the form

$$\begin{bmatrix} I & -\frac{1}{2}J & A_1 & A_2 & A_3 & A_4 \\ -\frac{1}{2}J & I & A_5 & A_6 & A_7 & A_8 \\ A_1 & A_5 & I & -\frac{1}{2}J & A_9 & A_{10} \\ A_2 & A_6 & -\frac{1}{2}J & I & A_{11} & A_{12} \\ A_3 & A_7 & A_9 & A_{11} & I & -\frac{1}{2}J \\ A_4 & A_8 & A_{10} & A_{12} & -\frac{1}{2}J & I \end{bmatrix}$$

up to simultaneous permutation of the rows and columns, where

$$A_1, \ldots, A_{12} \in \left\{ \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix}, \begin{bmatrix} 1/4 & -3/4 \\ -3/4 & 1/4 \end{bmatrix}, \begin{bmatrix} -3/4 & 1/4 \\ 1/4 & -3/4 \end{bmatrix} \right\}.$$

Using a computer we see that, up to simultaneous permutation of the rows and columns, the Gram matrix of C_2 is the only matrix of this form of rank at most 4, which shows C is isometric to C_2. □

In the remainder of this section show that if $C \subseteq S^3$ is a subset of size 12 with maximal inner product $1/4$, then its 4-point distance distribution agrees with the 4-point distance distribution of C_1 or C_2. For this we use Corollary 7.1.

We use a computer to find a rational solution to $\text{las}_2(4, 1/4)$. For this we first solve the semidefinite program in 256 bit floating-point arithmetic using the solver of Chapter 5, and then we use the rounding procedure from Chapter 6 to extract an optimal solution over the rationals from this. For some matrices, we could not find the kernel vectors using the method of Section 6.2 with the given precision. Therefore, we use the kernel vectors given by Corollary 7.1, where the first condition gives kernel vectors of the sum-of-squares matrices, and the second of the matrices K_λ. Numerically computing the rank of the matrices shows that in fact this gives all kernel vectors in this case.

We first check that the kernel K defined by this solution satisfies the following properties:

(1) K is a feasible for $\text{las}_2(4, 1/4)$ with $K(\emptyset, \emptyset) = 12$.
(2) For $x, y \in S^3$ with $\langle x, y \rangle \leq 1/4$, $A_2 K(\{x, y\}) = 0$ if and only if

$$\langle x, y \rangle \in P := \{-3/4, -5/8, -1/2, -1/3, 0, 1/8, 1/4\}.$$

(3) There are no $x, y, v, w \in S^3$ with pairwise inner products in P, $\langle x, y\rangle \in \{-1/2, 0\}$, $\langle v, w\rangle \in \{-5/8, -1/3, 1/8\}$, and with $A_2K(\{x, y, u, v\}) = 0$.

Property (1) is verified by first checking that the matrices $\widehat{K}_\lambda$, as well as the matrices for the sums-of-squares characterizations, are positive semidefinite. This is done by first writing each matrix in the form BAB^T of Section 6.3, where B rectangular matrix and A is a positive definite matrix, both over the rationals. Then we check the matrices are indeed positive definite by computing the Cholesky factorizations in ball arithmetic. Finally, we check in exact arithmetic that the linear conditions enforcing the sum-of-squares characterizations are satisfied. For Property (2) we use that $A_2K(\{x, y\})$ is a univariate polynomial in $\langle x, y\rangle$. We check this polynomial is zero at the appropriate inner products, and we use Sturm sequences to show these are the only zeros in the interval $[-1, 1/4]$. Property (3) is verified by enumerating all possibilities (by Property (2) there are only finitely many of them). Here we do use the symmetries of S_4 to do this efficiently. Such an enumeration also yields a small subset of $\mathcal{E}_3 \cup \mathcal{E}_4$ on which the distance distribution is potentially nonzero.

Now let $C \subseteq S^3$ be a subset of size 12 with maximal inner product $1/4$. By Corollary 7.1 and Property (1) it follows that $A_2K(S) = 0$ for all $S \subseteq C$ with $2 \leq |S| \leq 4$. By Properties (2) and (3) it then follows that either

$$\langle x, y\rangle \in \{-3/4, -1/2, 0, 1/4\}$$

for all distinct $x, y \in C$ or

$$\langle x, y\rangle \in \{-3/4, -5/8, -1/3, 1/8, 1/4\}$$

for all distinct $x, y \in C$.

Denote the 4-point distance distribution of C by z. By Corollary 7.1 and Property (1) we have

$$\sum_{\substack{S \subseteq C \\ |S| \leq 4}} A_2K_\lambda(S) = 0$$

for each $|\lambda| \leq d$, which translates into a number of linear constraints on the distance distribution z. We have the additional linear constraints

$$(12 - i)z(A) = \sum_{B \in \mathcal{E}_{i+1}} c(A, B)z(B)$$

for $A \in \mathcal{E}_i$ and $i = 2, 3$, where $c(A, B)$ counts how often the matrix A appears as a principal submatrix of B up to simultaneous permutations of the rows and columns.

We know that the inner products 0 and $-1/3$ cannot both appear in the code C. If we add the linear constraint

$$z\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0,$$

then by using row reduction in exact arithmetic we see there is a unique 4-point distance distribution, which agrees with the distribution of C_1. If instead we add the linear constraint

$$z\left(\begin{bmatrix} 1 & -1/3 \\ -1/3 & 1 \end{bmatrix}\right) = 0,$$

then there is a unique 4-point distance distribution which agrees with the distribution of C_2. Together with Lemma 7.6 and 7.7 the following result then follows.

THEOREM 7.8. *Suppose $C \subseteq S^3$ is a code of size 12 with pairwise inner products at most $1/4$. Then C is equal to C_1 or C_2 up to isometry.*

Notice that the above theorem implies that C_1 and C_2 are optimal spherical codes. A posteriori, it also implies C_1 and C_2 are defined, up to isometry, by their 2-point distance distributions.

CHAPTER 8

Energy minimization on the sphere

8.1. Introduction

In this chapter we consider the problem of energy minimization on a sphere. Let $C \subseteq S^{n-1}$ be a spherical code, and let $g : [0, 4) \to \mathbb{R}$ be a potential function. The *energy* of C with respect to g is

$$\hat{E}_g(C) = \frac{1}{2} \sum_{\substack{x,y \in C \\ x \neq y}} g(d(x, y)^2)$$

where $d(x, y) = \sqrt{2 - 2\langle x, y \rangle}$ is the chordal distance.

The energy minimization problem asks for a spherical code on S^{n-1} of size N which minimizes $\hat{E}_g(C)$ over all spherical codes C.

An important type of potential function is Riesz-s energy, given by

$$R_s(r) = \begin{cases} r^{-s/2} & s > 0, \\ -r^{-s/2} & s < 0. \end{cases}$$

For $n = 3$ and $s = 1$, the corresponding energy minimization problem is known as the Thompson problem, which minimizes the Coulomb potential. A natural generalization to higher dimensions is the minimization of harmonic energy: the case $s = n - 2$.

Yudin introduced a linear programming bound, an analogue of the Delsarte-Goethals-Seidel linear programming bound for spherical codes, and used this to prove that the regular simplex ($N = n + 1$) and the cross polytopes ($N = 2n$) are minimizers of harmonic energy [**144**]. The bound is equivalent to the first level of the Lasserre hierarchy for energy minimization.

Of special interest are configurations which are optimal for the class of *completely monotonic* potential functions; these configurations are called *universally optimal*. A function g is completely monotonic if it is infinitely differentiable, and $(-1)^k g^{(k)} \geq 0$ for all $k \geq 0$. The Riesz-s energy kernels are examples of completely monotonic functions. Universally optimal configurations are automatically optimal spherical codes: minimizing the energy of R_s as $s \to \infty$ is essentially maximizing the minimum distance.

Since our techniques mainly work with polynomials in terms of the inner products between points on the sphere, we instead optimize the energy

$$E_f(C) = \frac{1}{2} \sum_{\substack{x,y \in C \\ x \neq y}} f(\langle x, y \rangle)$$

where $f(u) = g(2-2u)$. Then g is completely monotonic if and only if f is *absolutely monotonic*: f is infinitely differentiable and $f^{(k)} \geq 0$ for all $k \geq 0$.

On the sphere, Cohn and Kumar use Yudin's bound to show in [**34**] that all sharp configurations are universally optimal. Recall that a spherical code $C \subseteq S^{n-1}$ is a *spherical k-design* if for every polynomial up to degree k, the average of the polynomial over the sphere is the same as the average over the code. A spherical code $C \subseteq S^{n-1}$ is *sharp* if there are m inner products between distinct points of C and it is a $(2m-1)$-design. Other than sharp configurations, the only known universal optima are in projective space [**42, 36**].

In Section 8.2 and 8.3, we introduce the three-point bound and the Lasserre hierarchy for energy minimization. We use the three-point bound in Section 8.4 to show universal optimality of the Nordstrom-Robinson spherical code, and perform in Section 8.5 computations on the second level of the Lasserre hierarchy for harmonic energy. This leads to the conjecture that the second level of the Lasserre hierarchy is sharp for harmonic energy for several families of configurations: $n+2$ and $2n-1$ points in dimension n. For $2n+2$ points in dimension n the same may hold, but it is less evident from the numerical results. We conclude in Section 8.6 with a theorem giving conditions for two codes to be universally optimal *together*: for every completely monotonic potential function, one of the two codes is optimal.

8.2. The three-point bound

The three-point bound for energy minimization was first defined in [**42**]. In contrast to the three-point bound for packing, the three-point bound for energy minimization is not clearly a relaxation of the second level of the Lasserre hierarchy for this problem. Interestingly, it only uses a 3-point constraint, whereas k-point bounds typically use l-point constraints for $k_0 \leq l \leq k$ with $k_0 \in \{0,1,2\}$.

The bound utilizes the same matrices

$$Y_k^n(u,v,t)_{ij} = u^i v^j \sqrt{1-u^2}^k \sqrt{1-v^2}^k P_k^{n-1}\left(\frac{t-uv}{\sqrt{1-u^2}\sqrt{1-v^2}}\right)$$

as the three-point bound for the spherical code problem. Instead of the matrices $S_k^n = 1/6 \sum_{\sigma \in S_3} \sigma Y_k^n$, we now define the matrices

$$T_k^n(u,v,t) = (N-2)S_k^n(u,v,t) + S_k^n(u,u,1) + S_k^n(v,v,1) + S_k^n(t,t,1).$$

Define the function $H(u,v,t)$ by

$$H(u,v,t) = c + \sum_{k=0}^{d} \langle H_k, T_k^n(u,v,t)\rangle$$

with $H_k \succeq 0$ and $c \in \mathbb{R}$. Then the problem

$$\begin{aligned} &\text{maximize} && \frac{N}{2}((N-1)c - \langle H_0, J\rangle) \\ &\text{subject to} && H(u,v,t) \leq \frac{1}{3}(f(u)+f(v)+f(t)) && \text{for } (u,v,t) \in \Delta \\ & && H_k \succeq 0 && k = 0,\ldots,d \\ & && c \in \mathbb{R} \end{aligned} \tag{8.2.1}$$

gives a lower bound on the energy of any configuration of N points on S^{n-1} for the potential function f. Here $\Delta = \{(u,v,t) : -1 \leq u,v,t < 1, 1+2uvt-u^2-v^2-t^2\}$. If f is continuous at 1, we may extend Δ to also include $1 \in \{u,v,t\}$. In that case the points in the configuration do not have to be distinct for the proof to work.

PROOF THAT (8.2.1) GIVES A LOWER BOUND ON THE ENERGY. Let C be a configuration of N points, and let $(c, H_1, \dots, H_d)$ be a feasible solution to the problem. Then

$$\begin{aligned} 0 &\leq \sum_{x,y,z \in C} \sum_{k=0}^{d} \langle H_k, S_k^n(u,v,t) \rangle \\ &= N\langle H_0, J \rangle + \frac{1}{N-2} \sum_{\substack{x,y,z \in C \\ \text{distinct}}} (H(\langle x,z \rangle, \langle y,z \rangle, \langle x,y \rangle) - c) \\ &\leq N\langle H_0, J \rangle + \sum_{x \neq y \in C} f(\langle x,y \rangle) - N(N-1)c, \end{aligned}$$

and hence $E_f(C) \geq \frac{N}{2}((N-1)c - \langle H_0, J \rangle)$. □

8.3. The Lasserre hierarchy for energy minimization

The Lasserre hierarchy for energy minimization was introduced by de Laat [**85**], who showed that it was numerically sharp for 5 points on S^2 for the Riesz-s energy with $s = 1, \dots, 7$, using polynomials of degree at most 8. With the harmonic analysis in Chapter 3 and the solver from Chapter 5, we are able to perform computations in any dimension using a much higher polynomial degree.

Let $G = (V, E)$ be a graph, and f a function on the non-edges. The $0-1$ problem is then given by

$$\begin{aligned} \text{minimize} \quad & \sum_{\{i,j\} \notin E} f(\{i,j\}) x_i x_j \\ \text{subject to} \quad & \sum_{i \in V} x_i = N. \end{aligned}$$

To generalize this, we take the graph with vertex set $V = S^{n-1}$, where distinct $x, y \in V$ are adjacent whenever $\langle x, y \rangle > \varepsilon$ for some $\varepsilon < 1$. Recall that $\mathcal{I}_{=k}$ denotes the independent sets in this graph of size equal to k, and that $\mathcal{I}_k$ is the disjoint union of $\mathcal{I}_{=i}$ for $i = 0, \dots, k$. Define the function

$$F(S) = \begin{cases} f(\langle x,y \rangle) & S = \{x,y\}, x \neq y, \\ 0 & \text{otherwise}, \end{cases}$$

where $f : [-1, 1) \to \mathbb{R}$ is the potential function as in the introduction of this chapter. Note that F is continuous, because $\mathcal{I}_{=2}$ is not connected to $\mathcal{I}_k \setminus \mathcal{I}_{=2}$. We reformulate the cardinality constraints using Example 2.6, which results in the problem

$$\begin{aligned} \text{minimize} \quad & \lambda(F) \\ \text{subject to} \quad & A_k^*(\lambda) \in \mathcal{M}(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}, \\ & \lambda(\mathcal{I}_{=i}) = \binom{N}{i}, \qquad i = 0, \dots, 2k. \end{aligned}$$

The dual problem is then given by

$$\begin{aligned} \text{maximize} \quad & \sum_{i=0}^{2k} a_i \binom{N}{i} \\ \text{subject to} \quad & a_i + A_k K(S) \leq F(S) \quad S \in \mathcal{I}_{=i}, i = 0, \ldots, 2k \\ & K \in C(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0}, \\ & a_0, \ldots, a_{2k} \in \mathbb{R} \end{aligned}$$

Similar to the Lasserre hierarchy for spherical codes, the problem is $O(n)$-invariant, and by averaging over $O(n)$ we may assume that the kernel K is $O(n)$-invariant. This allows us to use the description of Section 3.4 for the K. The i-th constraint has S_i-invariance which acts on the polynomials in the inner products by permuting the rows and columns of the Gram matrix of S. For the constraints i with $i \neq 2$, we can use Chapter 4 and Example 3.4 to give an efficient semidefinite programming relaxation through sum-of-squares polynomials.

If f is a polynomial, the same strategy works for constraint $i = 2$. Otherwise it is possible to get lower bounds on E_f using a polynomial lower approximation of f. However, if f is a Riesz-s potential for some positive integer s, it is possible to reformulate the constraint. In this case we multiply all terms in the constraint by $(2 - 2t)^{s/2}$, and if s is odd, change variables to $w = \sqrt{2 - 2t}$. This gives a polynomial inequality in the variable w (or t if s is even), which can be reformulated to semidefinite programming constraints using sums-of-squares polynomials, as in Chapter 4.

Up to now, we have not considered a specific value for the maximum inner product ε. Suppose there is an optimal configuration with maximum inner product U. Then we may take $\varepsilon = U$. However, since the hierarchy contains cardinality constraints, it has finite convergence for every $\varepsilon \in [U, 1]$. In practice, taking $\varepsilon \to 1$ does not significantly worsen the bounds. All bounds in this chapter take $\varepsilon \to 1$.

For certain potential functions f, it is possible to explicitly find a bound on the maximum inner product. For example, let f be a strictly increasing function with $f(u) \to \infty$ when $u \to 1$ (e.g., a Riesz-s potential function), and consider a (possibly suboptimal) configuration $C \subseteq S^{n-1}$. Then we can find a U such that $f(u) \geq E_f(C)$ for all $u \in [U, 1)$. In particular, this implies that any optimal configuration has maximum inner product at most U.

8.4. Universal optimality of the Nordstrom-Robinson spherical code*

The cone of absolutely monotonic functions has infinitely many extreme rays, which makes it difficult to prove universal optimality. Instead of proving optimality for a set of extreme rays, we follow the approach of [**42**]: we replace the cone of absolutely monotonic functions by a slightly larger but finitely generated cone.

First we recall some facts about Hermite interpolation, see [**34**, Section 2.1]. Given a nonempty, finite multiset $T \subseteq \mathbb{R}$, the Hermite interpolate $H_T(f)$ of a function f is the unique polynomial of degree at most $|T| - 1$ that agrees with f to order $\text{mult}_T(t)$ at t for all $t \in T$:

$$f^{(i)}(t) = H_T(f)^{(i)}(t)$$

*Part of this section is based on Section 4 of the publication "*H. Cohn, D. de Laat and N. Leijenhorst,* Optimality of spherical codes via exact semidefinite programming bounds*, 2024, arXiv:2403.16874*".

for all $t \in T$ and $i < \mathrm{mult}_T(t)$.

LEMMA 8.1 ([**42**, Lemma 9]). *Let T be a finite, nonempty multiset of $I = [a, b]$ such that each point in the interior of I has even multiplicity. For each absolutely monotonic function $f : I \to \mathbb{R}$ we have*

$$H_T(f)(t) \leq f(t)$$

for all $t \in I$.

LEMMA 8.2 ([**42**, Lemma 10]). *Let $T = \{t_1, \ldots, t_M\}$ be a nonempty multiset of an interval I, with t_i written according to multiplicity. If $f : I \to \mathbb{R}$ is absolutely monotonic, then $H_T(f)$ is a conic combination of the partial products*

$$t \mapsto \prod_{i=1}^{m}(t - t_i)$$

for $m = 0, \ldots, M - 1$.

Recall that a conic combination is a linear combination with nonnegative coefficients.

COROLLARY 8.3 ([**42**, Corollary 11] for the sphere). *Let C be a finite subset of S^{n-1}, and let $T = \{t_1, \ldots, t_M\}$ be a finite subset of $[-1, 1)$, written with multiplicities, such that all elements other than -1 have even multiplicity and*

$$\{\langle x, y\rangle : x \neq y \in C\} \subseteq T.$$

If C minimizes E_{f_m} for the potential functions

$$f_m(t) = \prod_{i=1}^{m}(t - t_i)$$

with $m = 0, \ldots, M - 1$, then C is universally optimal in S^{n-1}.

PROOF. Let f be an absolutely monotonic function. By Lemma 8.1 we have $f(t) \geq H_T(f)(t)$ for all $t \in [-1, 1)$, and by Lemma 8.2

$$H_T(f) = \sum_m \alpha_m f_m$$

for some $\alpha_m \geq 0$. Let D be any configuration with $|D| = |C|$, then

$$E_f(D) \geq E_{H_T(f)}(D) = \sum_m \alpha_m E_{f_m}(D) \geq \sum_m \alpha_m E_{f_m}(C) = E_f(C)$$

since C minimizes E_{f_m} and all inner products of C occur in T. □

To apply this result, one might need to use a multiplicity significantly higher than 2 for some of the inner products. In the following, we require $\mathrm{mult}_T(-1) = 11$ to prove universal optimality using the three-point bound.

THEOREM 8.4. *For each absolutely monotonic potential function $f : [-1, 1) \to \mathbb{R}$, the Nordstrom-Robinson spherical code C minimizes the energy E_f over all subsets $C \subseteq S^{15}$ with $|C| = 288$, and is the unique minimizer up to isometry unless f is a polynomial of degree at most 5.*

Uniqueness fails when f is a polynomial of degree 4 or less, since there exist 4-designs non-isomorphic to C. It is unclear whether C is the unique spherical 5-design of size 288 in dimension 16, but the existence of a different 5-design is the only way uniqueness can fail to hold for polynomials of degree 5.

PROOF. We use Corollary 8.3 in combination with the three-point bound for the set $T = \{(-1)^{11}, (-\frac{1}{4})^2, 0^2, (\frac{1}{4})^2\}$, where the superscripts indicate the multiplicities. Since C is a 5-design, the only cases we have to check are f_m with $m > 5$. Instead of using f_m, we prove optimality of C for the polynomials $f_m(t) - p(t)$ where

$$p(t) = (1+t)^{11}(t+1/4)^2t^2(t-1/4)^2/10000.$$

For this, we use the rounding procedure of Chapter 6 to obtain an exact optimal solution. The data set [**37**] includes the exact matrices as well as code for verifying feasibility of the solution.

Since $p(t) \geq 0$ on $[-1, 1)$ and $p(t) = 0$ if and only if $t \in T$, this ensures that E_{f_m} is minimized by codes with the same inner products as C: we have $E_{f_m - p(t)}(D) \geq E_{f_m - p(t)}(C)$, and if $\langle x, y\rangle \notin T$ for some $x \neq y \in D$ we have $E_{f_m}(D) > E_{f_m}(C)$, since $p(t) \geq 0$ on $[-1, 1]$ and $p(t) = 0$ only for $t \in T$. Since the Nordstrom-Robinson code is the unique code of size 288 in S^{15} with these inner products up to isometry, by Theorem 7.2, this proves uniqueness for the basis polynomials except for polynomials up to degree 5.

To prove uniqueness, let f be an absolutely monotonic function with hermite interpolant h. Suppose h is a polynomial of degree less than 16. We prove that $f = h$, so that the only way uniqueness fails is when f is a polynomial of degree at most 5. Note that $f - h$ has 17 roots, counted with multiplicity, so that Rolle's theorem implies that $(f-h)^{(16)}(t) = 0$ for some $t \in (-1, 1)$. Since h is a polynomial of degree less than 16, we have $f^{(16)}(t) = 0$, and absolute monotonicity implies that $f^{(16)}$ vanishes on $(-1, t)$. By Theorem 3a of Chapter IV of [**141**], f is analytic, and therefore $f^{(16)} = 0$ everywhere. In particular, f is a polynomial of degree at most 16, and $f = h$ because $f - h$ has at least 17 roots.

□

8.5. Families of configurations for harmonic energy†

In [**4**], Ballinger, Blekherman, Cohn, Giansiracusa, Kelly, and Schürmann find many configurations and families of configurations that they conjecture to be harmonic energy minimizers, other than the known cases $N = n + 1$ and $N = 2n$ by Yudin. In this section, we compute the second level of this hierarchy in higher dimensions and using higher degrees. We use this to find various numerically sharp cases for harmonic energy.

8.5.1. The case of $n + 2$ points in $\mathbb{R}^n$. Consider the configuration of $n + 2$ points in S^{n-1} consisting of two antipodal points and a regular simplex in the plane orthogonal to the line through these points. In Table 8.5.1 we show for which smallest even degrees we find a sharp harmonic energy lower bound with the second level of the Lasserre hierarchy. Based on this data we make the following conjecture:

†This section is based on Section 3.4 of the publication "*D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer,* Traceless projection of tensors with applications to hierarchies in discrete geometry*, In preparation*".

CONJECTURE 8.5. *For each $n \geq 3$, the second level of the Lasserre hierarchy gives a sharp lower bound for the harmonic energy of $n+2$ points in S^{n-1}.*

n	3	4-6	7-18	19-37	38
d	6	8	10	12	> 12

TABLE 8.5.1. Smallest even degree d for which the second level of the Lasserre hierarchy is numerically sharp for harmonic energy of $n+2$ points in $\mathbb{R}^n$.

8.5.2. The case of $2n-1$ points in $\mathbb{R}^n$. In Table 8.5.2 we show for which smallest even degrees we find a numerically sharp harmonic energy lower bound for $2n-1$ points in $\mathbb{R}^n$. In each case the configuration consists of the north pole and a cross polytope on a circle below the equator. Based on this data we make the following conjecture:

CONJECTURE 8.6. *For each $n \geq 3$, the second level of the Lasserre hierarchy gives a sharp lower bound for the harmonic energy of $2n-1$ points in S^{n-1}.*

n	3	4	5-10	11-30	31-59	60
d	6	8	10	12	14	> 14

TABLE 8.5.2. Smallest even degree d for which the second level of the Lasserre hierarchy is numerically sharp for harmonic energy of $2n-1$ points in $\mathbb{R}^n$.

8.5.3. The case of $2n+2$ points in $\mathbb{R}^n$. The diplo-simplex is the union of a simplex with its antipodal simplex [**46**]. In [**4**] it is shown that the diplo-simplex is suboptimal for harmonic energy with $3 \leq n \leq 5$ but appears to be optimal for $n \geq 6$. Our calculations show that the second level of the Lasserre hierarchy is numerically sharp for the problem of minimizing harmonic energy in at least dimensions $3 \leq n \leq 12$. In Table 8.5.3 we list the smallest even degree d for which the bound is numerically sharp in dimension n.

CONJECTURE 8.7. *For each $n \geq 3$, the second level of the Lasserre hierarchy gives a sharp lower bound for the harmonic energy of $2n+2$ points in S^{n-1}.*

n	3	4	5	6	7	8	9	10	11	12	13
d	10	12	14	14	12	12	12	12	12	14	≥ 14

TABLE 8.5.3. Smallest even degree d for which the second level of the Lasserre hierarchy is numerically sharp for harmonic energy of $2n+2$ points in $\mathbb{R}^n$.

8.5.4. Other numerically sharp bounds. In Table 8.5.4 we list other cases where the second level of the Lasserre hierarchy is numerically sharp. All computations have been performed with degree $d = 14$ or less.

n	N
$n \leq 9$	$n+3$
$n \leq 7$	$n+4$
$n = 4, 8$	$2n+4$
$n = 3$	$9, 10$

TABLE 8.5.4. Other cases where the second level of the Lasserre hierarchy is numerically sharp with $n \geq 3$. For $(n, N) = (4, 12)$, the bound requires degree 14 polynomials; in the other cases degree 12 suffices, although it may not necessarily be the lowest degree for which the bound is sharp.

8.6. Two-code universal optimality

Cohn and Woo [**42**] conjecture that for 10 points in S^3, every absolutely monotonic potential function is optimized by the so-called Peterson code (the unique optimal spherical code with maximal inner product 1/6, see [**3**]), or the code consisting of two regular pentagons in orthogonal planes. They could prove that for the potential functions $(1+t)^k$, the Peterson code is optimal for $k \geq 7$, and the pentagon code is optimal for $3 \leq k \leq 6$, using the three-point bound. These potentials span the cone of absolutely monotonic functions, and one approach to proving the conjecture would be to prove that both codes are optimal for all functions of the form

$$(1+t)^j + \alpha_{j,k}(1+t)^k$$

with $3 \leq j \leq 6$ and $k \geq 7$, where $\alpha_{j,k} > 0$ is chosen such that both codes have the same energy.

Here we record a result that could potentially be used in such a case, and can be seen as a simple generalization of Corollary 8.3 for two codes.

COROLLARY 8.8. *Let $C_1, C_2 \subseteq S^{n-1}$ be two codes of the same size. Let $T = \{t_1, \ldots, t_M\}$ be a finite subset of $[-1, 1)$, written with multiplicities, such that all elements other than -1 have even multiplicity, and*

$$\{\langle x, y\rangle : x, y \in C_i, i = 1, 2, x \neq y\} \subseteq T.$$

Define

$$f_m(t) = \prod_{i=1}^{m}(t - t_i)$$

for $m = 0, \ldots, M-1$. Suppose C_i is the unique optimum for E_{f_m} with $m \in I_i$, and both codes are optimal for $m \notin I_1 \cup I_2$. Suppose furthermore that both codes are optimal for the functions

$$g_{i,j}(t) = f_i(t) + \alpha_{i,j} f_j(t)$$

with $i \in I_1$ and $j \in I_2$, where $\alpha_{i,j} > 0$ is chosen such that both codes have the same energy. Then for every absolutely monotonic potential function, either C_1 or C_2 is optimal.

PROOF. Let f be an absolutely monotonic function. By Lemma 8.1 we have $f(t) \geq H_T(f)(t)$ for all $t \in [-1, 1)$, and by Lemma 8.2,

$$H_T(f) = \sum_m \beta_m f_m$$

for some $\beta_m \geq 0$. Since the $g_{i,j}$ are linear combinations of the f_m, we can redistribute the coefficients and write

$$H_T(f) = \sum_{i \in I_1, j \in I_2} \lambda_{i,j} g_{i,j} + \sum_{m \in I_1} \lambda_m f_m + \sum_{m \in I_2} \lambda_m f_m + \sum_{m \notin I_1 \cup I_2} \beta_m f_m$$

where $\lambda_{i,j}, \lambda_m$ and β_m are nonnegative, such that $\sum_j \lambda_{m,j} + \lambda_m = \beta_m$ for $m \in I_1$ and similarly $\sum_i \lambda_{i,m} \alpha_{i,m} + \lambda_m = \beta_m$ for $m \in I_2$. In particular, we can choose $\lambda_{i,j}$ such that either $\lambda_m = 0$ for all $m \in I_1$ or $\lambda_m = 0$ for all $m \in I_2$: Suppose $\lambda_i > 0$ for some $i \in I_1$ and $\lambda_j > 0$ for some $j \in I_2$, then we can increase $\lambda_{i,j}$ by $\min\{\lambda_i, \lambda_j/\alpha_{i,j}\}$ to make at least one of λ_i and λ_j zero.

If $\lambda_m = 0$ for all $m \in I_1$, clearly C_2 is optimal for $H_T(f)$, and if $\lambda_m = 0$ for all $m \in I_2$, C_1 is optimal. The codes have equal energy if $\lambda_m = 0$ for $m \in I_1 \cup I_2$.

Then, for a code D of the same size, we have

$$\begin{aligned} E_f(D) &\geq E_{H_T(f)}(D) \\ &\geq E_{H_T(f)}(C_i) \\ &= E_f(C_i) \end{aligned}$$

for some $i \in \{1, 2\}$ such that C_i is optimal for $H_T(f)$. The last equality holds because all inner products of C_i occur in T by assumption. □

Unfortunately, it seems that for the earlier mentioned case $(n, N) = (4, 10)$ the maximum degree of the polynomials f_i and $g_{i,j}$ needs to be relatively high, if it is possible to prove two-code universal optimality using Corollary 8.8 for this case. For every set T of size at most 18, containing the inner products of the two codes with even multiplicity, we could find a polynomial f_i and a code C such that C had strictly lower energy for the polynomial f_i than both the pentagon and the Peterson code, using a basic gradient descent algorithm. We used a tolerance of $\varepsilon = 10^{-5}$ to account for possible floating point errors, which are typically of much smaller size.

CHAPTER 9

Polarization with a threshold on the sphere*

9.1. Introduction

Let $f : [-1, 1) \to \mathbb{R}$ be a piecewise continuous potential function, and fix a number $E \in \mathbb{R}$, which we call the *threshold*. In this chapter, we consider the problem of minimizing the number $N \in \mathbb{N}$ such that there is a configuration $C \subseteq S^{n-1}$ of size N with the property that

$$P_f(C, y) = \sum_{x \in C} f(\langle x, y \rangle) \geq E$$

for all $y \in S^{n-1}$. The quantity $P_f(C, y)$ is called the polarization of y with respect to C and f. Intuitively, the polarization problem asks for the minimum number of light sources such that the darkest point has at least a certain brightness.

The polarization problem is typically stated for continuous functions f. In our statement, we include piecewise continuous functions so that the problem can be specialized to the sphere covering problem. For this, take $f(u) = \chi_{[r,1]}(u)$ for some radius r. For fixed y, the corresponding function $x \mapsto f(\langle x, y \rangle)$ is then the indicator function of the spherical cap

$$B(y, r) = \{x \in S^{n-1} : \langle x, y \rangle \geq r\}.$$

This was considered by Riener, Rolfes, and Vallentin in [**124**].

In both cases, we assume there is an optimal configuration C such that $\langle x, y \rangle \leq \varepsilon$ for all distinct $x, y \in C$, for some given $\varepsilon \in (-1, 1)$. In general, this maximum inner product is unknown, and the cases where it is known are typically solved instances.

In this chapter, we perform computations on semidefinite programming bounds for this problem. We introduce the Lasserre hierarchy for the packing-polarization problem with threshold in Section 9.2. For the special case of covering the sphere, this is equivalent to the symmetry-reduced version of the hierarchy introduced in [**124**], which is given in [**125**]. In Section 9.4, we show the importance of the maximum inner product ε: With $\varepsilon = 1$, the hierarchy collapses to the first level of the hierarchy, which we show to be the volume bound in Section 9.3.

From the formulation of the hierarchy, it is not directly clear that the constraints are polynomial inequality constraints. In Section 9.5 we show that this is the case for polynomial $O(n)$-invariant kernels, and perform explicit computations on the integrals over the sphere present in the bounds to compute the polynomial inequality constraints. This allows us to write the problem as a semidefinite program using the techniques of Chapters 3 and 4. In Section 9.6 we give an adaptation to $(2k - 1)$-point bounds, and a short explanation of how the bounds can be adapted

*Parts of this chapter are based on the paper "*N. Leijenhorst, A. Spomer, F. Vallentin and M. C. Zimmerman,* Semidefinite approaches to covering and polarization on the sphere*, In preparation*".

to $(2k+t)$-point bounds in general. In Section 9.7 we compute both the three-point bound and the second step of the Lasserre hierarchy for a number of instances, and show that the bounds are numerically sharp in several cases. We also consider the relationship between the maximum inner product ε and the bound for various polynomial degrees.

9.1.1. Polarization on the sphere with a threshold. For fixed N and $f : [-1, 1) \to \mathbb{R}$, the polarization problem asks for a spherical code of size N that maximizes

$$\min_{y \in S^{n-1}} P_f(C, y)$$

over all spherical codes of the same size. This is a bilevel optimization problem, and as such, much more difficult than the problems considered in this thesis.

Little is known about exact solutions in the most general setting: only the cases $N \leq n + 1$ are solved [**16**]. Some results are known in slightly more specific cases. For example, in [**15**] Borodachov shows that the set of vertices of the cross-polytope is optimal among all antipodal configurations of size $2n$ for potential functions with a positive and convex second derivative. In [**20**] upper and lower bounds for optimizing the polarization over all spherical t-designs of the same size are given.

As with the spherical code problem, we give bounds on the 'transposed' problem. Given a polarization threshold E, what is the minimum number N such that there is a configuration $C \subseteq S^{n-1}$ of N points with

$$P_f(C, y) \geq E$$

for all $y \in S^{n-1}$. One basic observation is that the bound requires $E > 0$, since for $E \leq 0$ zero points suffice.

In general, a simple bound on the number of points can be given as follows. Suppose the polarization equals the threshold everywhere, then the integral of the polarization over the sphere equals $E\omega_n(S^{n-1})$. Each point in the configuration adds a polarization $\int_{S^{n-1}} f(\langle x, y\rangle) d\omega_n(y)$, so one needs at least

$$N \geq \frac{E\omega_n(S^{n-1})}{\int_{S^{n-1}} f(\langle x, y\rangle) d\omega_n(y)}.$$

This is the polarization version of the volume bound for covering.

9.1.2. Sphere covering. A little more is known for the problem of covering the sphere. It was first considered by Fejes-Toth [**55**], who gave a geometric bound for dimension 3. The bound is sharp for $N = 4, 6$ and 12 points, corresponding to the regular simplex, the cross-polytope and the icosahedron. In all other dimensions the best known general bound is the volume bound. Similarly to the spherical code problem, only few cases are known to be optimal. Examples of known optimal configurations are the regular simplices in every dimension, and the cross-polytope in dimensions 3 and 4; see, e.g., [**18**, Section 3.3.2] and the references therein. The cross-polytope is conjectured to be optimal in every dimension.

Recently, Riener, Rolfes and Vallentin introduced in [**124**] the Lasserre hierarchy for the 'transposed' problem: what is the minimum number of caps needed to cover the sphere, given a covering radius r? As in Chapter 2, the hierarchy optimizes over measures on independent sets of cardinality at most $2k$, and to set up the problem, one needs a maximum inner product ε. The points in an optimal configuration have such a maximum inner product, but it is unknown in general. Without knowing a

bound on the maximum inner product, the hierarchy gives a bound on a packing-covering problem: what is the minimum number of points you need to cover the sphere with a covering radius r, while giving a packing with packing radius ε?

9.2. The Lasserre hierarchy

In this section we apply the framework of Chapter 2 to the polarization problem with threshold. For packing-covering, this gives a hierarchy equivalent to the symmetry-reduced hierarchy in [**125**].

We minimize the number of points, so the objective function is given by $-\lambda(\mathcal{I}_{=1})$; the minus sign is required because the general primal problem (2.3.1) is a maximization problem. There are no equality constraints, and for every point on the sphere there is an inequality constraint. In terms of a finite graph with binary variables, the constraint reads

$$\sum_{i:\{i,j\} \text{ is an edge}} f(\{i,j\})x_i - E \geq 0$$

for all $j \in V$. This gives for our infinite graph the set of inequality constraints G defined by

$$G = \{g(\cdot, y) : y \in S^{n-1}\}$$

where

$$g(Q,y) = \begin{cases} -E & \text{if } Q = \emptyset, \\ f(\langle x, y\rangle) & \text{if } Q = \{x\}, \\ 0 & \text{otherwise.} \end{cases}$$

Recall that by a configuration is valid if

$$\sum_{Q \subseteq C} g(Q,y) \geq 0$$

for all $y \in S^{n-1}$. This gives for a spherical code C that

$$\sum_{Q \subseteq C} g(Q,y) = -E + \sum_{x \in C} f(\langle x, y\rangle)$$

so C is valid if and only if $P_f(C,y) \geq E$ for all $y \in S^{n-1}$, so validity exactly corresponds to feasibility of a code. Furthermore, if f is continuous or $f(u) = \chi_{[r,1]}(u)$, an invalid configuration is strictly invalid, so Theorem 2.8 implies that the hierarchy converges in a finite number of steps. For a valid configuration C, define the measure

$$\lambda = \chi_C = \sum_{S \subseteq C} \delta_S,$$

so that

$$\begin{aligned} \langle (A_k^G)^* \lambda, K\rangle &= \int \sum_{\substack{S \subseteq C \\ |S| \leq 2k}} A_k^y K(S) d\omega_n(y) \\ &= \int \sum_{\substack{Q \subseteq C \\ |Q| \leq 1}} g(Q,y) \sum_{\substack{J_1, J_2 \subseteq C \\ |J_1|, |J_2| \leq k-1}} K(J_1, J_2, y) d\omega_n(y) \geq 0 \end{aligned}$$

for every slice-positive kernel $K \in C(\mathcal{I}_{k-1} \times \mathcal{I}_{k-1} \times S^{n-1})$. Hence, this measure, when restricted to $\mathcal{I}_{2k}$ is feasible for the primal problem with objective value $|C|$, so

every step of the hierarchy gives a lower bound on the size of an optimal configuration, and strong duality holds.

The dual of the hierarchy is given by

$$\begin{aligned} &\text{maximize} && a_0 \\ &\text{subject to} && -A_k K - \int_{S^{n-1}} A_k^y K_y' d\omega_n(y) \\ &&& \quad - a_0 \chi_{\mathcal{I}_0} \geq -\chi_{\mathcal{I}_{=1}} && \text{on } \mathcal{I}_{2k} \\ &&& K' \in C(\mathcal{I}_{k-1} \times \mathcal{I}_{k-1} \times S^{n-1}) \text{ is slice-positive} \\ &&& K \in C(\mathcal{I}_k \times \mathcal{I}_k)_{\succeq 0} \end{aligned} \tag{9.2.1}$$

where

$$A_k^y K_y'(S) = \sum_{\substack{x \in S^{n-1}, J_1, J_2 \in \mathcal{I}_{k-1} \\ J_1 \cup J_2 \cup \{x\} = S}} f(\langle x, y \rangle) K'(J_1, J_2, y) - \sum_{\substack{J_1, J_2 \in \mathcal{I}_{k-1} \\ J_1 \cup J_2 = S}} EK(J_1, J_2, y).$$

and we changed a_0 to $-a_0$ to obtain a maximization problem instead of a minimization problem. We denote both the problem and the optimal value by $\mathrm{las}_k(f, E)$. Any feasible solution will give a lower bound on the optimal size of a configuration.

To compute the second level of this hierarchy, we again set

$$K(J_1, J_2) = \sum_{|\lambda| \leq d_1} \langle K_\lambda, Z_\lambda(J_1, J_2) \rangle$$

where the entries of Z_λ are polynomials with degree at most d_2; see Chapter 3. For K', we use the three-point function of Bachoc and Vallentin, and extend it to a slice-positive kernel in $C(\mathcal{I}_1 \times \mathcal{I}_1 \times S^{n-1})$. That is, we have

$$K'(J_1, J_2, y) = \sum_{k=0}^{d_3} \langle K_k', Y_k^n(J_1, J_2, y) \rangle$$

where for $k > 0$, $Y_k^n(\{x\}, \{z\}, y) = Y_k^n(x, z, y)$ and 0 otherwise, and $Y_0^n(J_1, J_2, y) = v(J_1, y) v(J_2, y)^\mathsf{T}$ with

$$v(J, y)_i = \begin{cases} 1 & \text{if } J = \emptyset \text{ and } i = -1 \\ \langle x, y \rangle^i & \text{if } J = \{x\} \text{ and } i \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

In Section 9.5, we give an explicit expression to compute the integral

$$\int_{S^{n-1}} A_k^y K_y'(S) d\omega_n(y),$$

which results in a polynomial in the inner products between vectors in S whenever f is integrable over S^{n-1}. In particular, this implies that we can use Chapter 4 to write the problem as a semidefinite program.

9.3. The first level of the hierarchy

In this section we show that the first level of the hierarchy gives the volume bound for the polarization problem with threshold.

THEOREM 9.1. *Suppose $E > 0$ and the integral of f over the sphere is at most E. Then $\mathrm{las}_1(f, E)$ equals*

$$v = \frac{E\omega_n(S^{n-1})}{\int_{S^{n-1}} f(\langle x, y\rangle) d\omega_n(y)}.$$

That is, the first level gives the volume bound.

PROOF. Let C_1 be a configuration with

$$|C_1| = \lceil v \rceil,$$

and C_2 a configuration with

$$|C_2| = \lfloor v \rfloor.$$

Let $a_1, a_2 \geq 0$ be constants such that $a_1|C_1| + a_2|C_2| = v$ and $a_1 + a_2 = 1$, and consider the measure

$$\lambda = \int_{O(n)} a_1 \chi_{\gamma C_1} + a_2 \chi_{\gamma C_2} d\gamma$$

where the Haar measure is used for the integration over $O(n)$. That is, we take an average of a convex combination of the indicator measures of C_1 and C_2 such that the objective function gives

$$\lambda(\mathcal{I}_{=1}) = a_1|C_1| + a_2|C_2| = v.$$

Then λ is feasible: for any positive definite kernel $K \in C(\mathcal{I}_1 \times \mathcal{I}_1)$,

$$\langle A_1^* \lambda, K\rangle = \sum_{i=1}^{2} a_i \int_{O(n)} \sum_{\substack{J_1, J_2 \subseteq \gamma C_i \\ |J_1|, |J_2| \leq 1}} K(J_1, J_2) d\gamma \geq 0$$

by positivity of K, and for any nonnegative function $K' \in C(\mathcal{I}_0 \times \mathcal{I}_0 \times S^{n-1})$ we have

$$\begin{aligned}
&\langle (A_0^G)^* \lambda, K'\rangle \\
&= \int_{S^{n-1}} \int_{O(n)} \sum_{i=1}^{2} a_i \sum_{x \in \gamma C_i} f(x, y) d\gamma K(\emptyset, \emptyset, y) - E K(\emptyset, \emptyset, y) d\omega_n(y) \\
&= \int_{S^{n-1}} \left(\sum_{i=1}^{2} a_i |C_i| \omega_n(S^{n-1})^{-1} \int_{S^{n-1}} f(x, y) d\omega_n(x) - E \right) K(\emptyset, \emptyset, y) d\omega_n(y) \\
&= 0
\end{aligned}$$

because $\sum_i a_i|C_i| = v$. This gives v as upper bound on the optimal objective function value.

For the other equality, let $K \in C(\mathcal{I}_1 \times \mathcal{I}_1)$ be the zero kernel, take $K' \in C(\mathcal{I}_0 \times \mathcal{I}_0 \times S^{n-1})$ constant and equal to $v/(E\omega_n(S^{n-1}))$, and let $a_0 = v$. This is a feasible solution to the dual, since the constraints read

$$-\int_{S^{n-1}} \sum_{\substack{x \in S^{n-1} \\ \{x\} = S}} f(x, y) K'(\emptyset, \emptyset, y) d\omega_n(y) = -1$$

for all $S \in \mathcal{I}_{=1}$, and

$$\int_{S^{n-1}} E K'(\emptyset, \emptyset, y) d\omega_n(y) - a_0 = 0$$

for $S = \emptyset$. For $S = \{x, y\}$ the constraint vanishes, so all constraints are satisfied. The objective is given by v, so the optimal objective function value of the first level of the hierarchy equals v, the volume bound. □

9.4. Low minimum distances

The bound $\text{las}_k(f, E)$ has as additional parameter an upper bound on the maximum inner product ε between the points in the configuration. In this section, we show that for $\varepsilon = 1$, polynomial relaxations of the bound reduce to the first level of the hierarchy: the volume bound. Without bound on the minimal distance, the convergence proof breaks because of independent sets of infinite size. In particular, this means that to improve on the volume bound with the second level of the hierarchy, we need a lower bound on the minimum distance, equivalently, an upper bound on the maximum inner product, between points in an optimal configuration.

We denote by $\text{las}_{k,\varepsilon}(f, E)_d$ problem (9.2.1) after restricting to polynomials of finite degree d, with constraints on inner products in $[-1, \varepsilon]$.

PROPOSITION 9.2. *Let $k > 1$ and $d \geq 0$. If $f \geq 0$ on $[-1, 1]$, then*

$$\text{las}_{k,1}(f, E)_d = \text{las}_{k-1,1}(f, E)_d.$$

PROOF. Let (K, K') be a feasible solution to the problem. Note that, since K and K' are polynomials in the inner products in $[-1, 1)$, the constraints also hold for the inner product 1.

Consider the constraint

$$A_k K(S) = \sum_{\substack{J_1, J_2 \in \mathcal{I}_k \\ J_1 \cup J_2 = S}} K(J_1, J_2) \leq 0$$

with $S = \{x_1, \ldots, x_{2k}\} \subseteq S^{n-1}$. Since $K(J_1, J_2)$ is a polynomial in the inner products between vectors in $J_1 \cup J_2$, the limit of $x_i \to x_j$ exists. This implies we may take S to be a multiset. Given a tuple $(\lambda_1, \ldots, \lambda_k)$ of nonnegative integers and points $x_1, \ldots, x_k \in S^{n-1}$, we denote by x^λ the multiset with elements x_i with multiplicity λ_i. For an integer l and $x \in S^{n-1}$, we denote by $\{x^l\}$ the multiset with the element x with multiplicity l. With multiplicities, the constraint is given by

$$A_k K(x^\lambda) = \sum_{\substack{\mu_1, \mu_2 \in \mathbb{N}^k \\ \mu_1 + \mu_2 = \lambda}} \alpha_{\mu_1, \mu_2} K(x^{\mu_1}, x^{\mu_2}) \leq 0,$$

where

$$\alpha_{\mu_1, \mu_2} = \prod_i \binom{\lambda_i}{(\mu_1)_i}$$

counts in how many ways x^λ can be distributed over x^{μ_1} and x^{μ_2} when all elements of the multiset would be labeled.

Consider the colexicographical order on the set P of partitions of k

$$P = \{(\lambda_1, \ldots, \lambda_k) \in \mathbb{Z}^k : \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0, \sum_i \lambda_i = k\}.$$

That is, for $\lambda, \mu \in P$ we have $\lambda < \mu$ if $\lambda_i < \mu_i$ for the last i where λ and μ differ. For example, $(k, 0, \ldots, 0)$ is the minimal partition, and $(1, \ldots, 1)$ is the maximal partition of t.

We apply induction on the partitions of k with this total order. As base case, take $\lambda = (k, 0, \ldots, 0)$. For this partition we get

$$A_k K(x^{2\lambda}) = \alpha_{\lambda,\lambda} K(x^\lambda, x^\lambda) = 0,$$

where $\alpha_{\lambda,\lambda} > 0$ is the product of binomials, since the left-hand side is nonpositive by the constraint and nonnegative by positivity of K. By the positivity of K, this implies that $K(\{x^k\}, J) = 0$ for all $x \in S^{n-1}$ and $J \in \mathcal{I}_k$.

Now suppose that $K(x^\lambda, J) = 0$ for any partition $\lambda < \mu$, with $x_i \in S^{n-1}$ and $J \in \mathcal{I}_k$. Then

$$A_k K(x^{2\mu}) = \sum_{\substack{\mu_1,\mu_2 \in \mathbb{N}^k \\ \mu_1+\mu_2=2\mu}} \alpha_{\mu_1,\mu_2} K(x^{\mu_1}, x^{\mu_2}) = \alpha_{\mu,\mu} K(x^\mu, x^\mu) = 0,$$

where the sum is over tuples of nonnegative integers of length k. The last equality holds since either $\mu_1 < \mu < \mu_2, \mu_2 < \mu < \mu_1$, or $\mu_1 = \mu_2 = \mu$, and by assumption $K(x^{\mu_1}, x^{\mu_2}) = 0$ in the first two cases. As above, the constraint and positivity of K imply that $K(x^\mu, J) = 0$ for all $x_i \in S^{n-1}$ and $J \in \mathcal{I}_k$.

Hence for any solution (K, K'), $K(J_1, J_2) = 0$ whenever J_1 or J_2 is of size k. In particular, this is always the case for the constraints on $\mathcal{I}_{2k}$ and $\mathcal{I}_{2k-1}$. Therefore the constraint on $\mathcal{I}_{2k-1}$ reduces to

$$\int_{S^{n-1}} \sum_{\substack{x \in S^{n-1}, J_1, J_2 \in \mathcal{I}_k \\ J_1 \cup J_2 \cup \{x\} = S}} f(x, y) \alpha_{J_1,J_2} K'(J_1, J_2, y) d\omega_n(y) \leq 0$$

for $S \in \mathcal{I}_{2k-1}$. Suppose $S = \{x^{2k-1}\}$. Then for every y, we have the integrant $f(x, y)\alpha_{k-1,k-1} K'(\{x^{k-1}\}, \{x^{k-1}\}, y)$, which is nonnegative due to positivity of K and nonnegativity of f, but the integral over y is nonpositive, so $K'(\{x^{k-1}\}, \{x^{k-1}\}, y) = 0$.

Let μ be any partition of $k-1$, and suppose $K(x^\lambda, J) = 0$ whenever $\lambda < \mu$. Take $S = x^{2\mu+(1,0,\ldots,0)}$. Then the constraint reads

$$\begin{aligned} &\int \sum_{\substack{x \in B(y,r), \mu_1, \mu_2 \in \mathbb{N}^k \\ \mu_1+\mu_2+\chi_x = 2\mu+(1,0,\ldots,0)}} f(x, y) \alpha_{\mu_1,\mu_2} K'(x^{\mu_1}, x^{\mu_2}, y) dy \\ &= \int f(x_1, y) \alpha_{\mu,\mu} K'(x^\mu, x^\mu, y) dy \leq 0 \end{aligned}$$

since other terms have multiplicities $\mu_i < \mu$ for some i. By positivity, we then obtain $K'(x^\mu, J, y) = 0$ for all multisets J of size $k-1$.

Hence the kernels K and K' are identically 0 on arguments from $\mathcal{I}_{=k}$ and $\mathcal{I}_{=k-1}$, respectively, so the solutions of $\mathrm{las}_{k,1}(f, E)_d$ exactly correspond to the solutions of $\mathrm{las}_{k-1,1}(f, E)_d$. This implies that the objective function values are equal. □

COROLLARY 9.3. *The optimal objective function value of* $\mathrm{las}_{k,1}(f, E)_d$ *equals*

$$\frac{E\omega_n(S^{n-1})}{\int_{S^{n-1}} f(\langle x, y \rangle) d\omega_n(y)}.$$

PROOF. Iterating Proposition 9.2 gives that the optimal value equals the optimal value of the first level of the hierarchy, which equals $\frac{E\omega_n(S^{n-1})}{\int_{S^{n-1}} f(\langle x,y \rangle) d\omega_n(y)}$ by Theorem 9.1 □

Intuitively, it is reasonable to ask for a minimum distance r for the covering problem: points in the covering should lie outside the caps around other points. This does not hold for all optimal coverings, but it is very well possible that on the sphere there always exists an optimal covering with this property. In dimension 3, all putatively optimal coverings in [**66**] satisfy this property.

CONJECTURE 9.4. *Given a dimension n and a covering radius r, there exists an optimal covering $C \subseteq S^{n-1}$ with covering radius r such that $d(x,y) \geq r$ for all distinct $x, y \in C$.*

9.5. Explicit integration on the sphere

In this section we explicitly compute the integrals

$$\int_{S^{n-1}} f(\langle w,z\rangle)\langle x,w\rangle^{i_1}\langle y,w\rangle^{i_2} d\omega_n(w)$$

and

$$\int_{S^{n-1}} \langle x,w\rangle^{i_1}\langle y,w\rangle^{i_2} d\omega_n(w)$$

in terms of the inner products between x, y and z. We will show that this gives a polynomial in these inner products, which allows us to use sum-of-squares polynomials and semidefinite programming to compute a relaxation of the second level of the hierarchy.

9.5.1. Integrating inner products over the sphere. Consider the integral

$$\frac{1}{\omega_n}\int_{S^{n-1}} \langle x,w\rangle^{i_1}\langle y,w\rangle^{i_2} d\omega_n(w)$$

for $n \geq 3$.

We write $w = ux + \sqrt{1-u^2}\xi$ and $y = tx + \sqrt{1-t^2}\zeta$, where ξ and ζ are unit vectors orthogonal to x. The measures $d\omega_n$ and $d\omega_{n-1}$ are then related by

$$d\omega_n(w) = (1-u^2)^{(n-3)/2} du d\omega_{n-1}(\xi),$$

which gives

$$\begin{aligned}
&\frac{1}{\omega_n}\int_{S^{n-2}}\int_{-1}^{1} u^{i_1}(ut+\sqrt{1-u^2}\sqrt{1-t^2}\langle\xi,\zeta\rangle)^{i_2}(1-u^2)^{(n-3)/2} du d\omega_{n-1}(\xi)\\
&\quad= \frac{1}{\omega_n}\sum_{k=0}^{i_2}\binom{i_2}{k} t^{i_2-k}(1-t^2)^{k/2}\\
&\qquad\int_{-1}^{1} u^{i_1+i_2-k}(1-u^2)^{(n+k-3)/2} du \int_{S^{n-2}}\langle\xi,\zeta\rangle^k d\omega_{n-1}(\xi)\\
&\quad= \frac{1}{\omega_n}\sum_{k=0}^{i_2}\binom{i_2}{k} t^{i_2-k}(1-t^2)^{k/2} I_{i_1+i_2-k,(n+k-3)/2}\omega_{n-2}\int_{-1}^{1} u^k(1-u^2)^{(n-4)/2} du\\
&\quad= \frac{\omega_{n-2}}{\omega_n}\sum_{k=0}^{i_2}\binom{i_2}{k} t^{i_2-k}(1-t^2)^{k/2} I_{i_1+i_2-k,(n+k-3)/2} I_{k,(n-4)/2},
\end{aligned}$$

where I_{ij} denotes the integral

$$\int_{-1}^{1} u^i(1-u^2)^j du.$$

This integral equals

$$\begin{aligned}(1+(-1)^i)\int_0^1 u^i(1-u^2)^j du \\ &= \frac{1}{2}(1+(-1)^i)\int_0^1 u^{i/2-1/2}(1-u)^j du \\ &= \frac{1}{2}(1+(-1)^i)B(i/2+1/2, j+1),\end{aligned}$$

where $B(x,y)$ denotes the Beta function.

Note that terms with k or i_1+i_2-k odd become zero, so that the expression is 0 whenever i_1+i_2 is odd, and a polynomial in t otherwise. Since the original expression is invariant under interchanging i_1 and i_2, we may replace i_1 by $\max\{i_1,i_2\}$ and i_2 by $\min\{i_1,i_2\}$ to reduce the number of terms present in the sum. In particular this implies that the maximum power of t equals $\min\{i_1,i_2\}$, and that only powers with an even difference with $\min\{i_1,i_2\}$ occur.

When $n=2$, which we will need in Section 9.5.2 when $n=3$, the integrals slightly change. Without loss of generality, we may assume that $x=e_1$ and $y=te_1+\sqrt{1-t^2}e_2$. We can write the integral in polar coordinates, which gives

$$\begin{aligned}\int_0^{2\pi} \cos(\theta)^{i_1}(\cos(\theta)t+\sin(\theta)\sqrt{1-t^2})^{i_2}d\theta \\ &= \sum_{k=0}^{i_2}\binom{i_2}{k}t^{i_2-k}(1-t^2)^{k/2}\int_0^{2\pi}\cos(\theta)^{i_1+i_2-k}\sin(\theta)^k d\theta \\ &= \sum_{k=0}^{i_2}\binom{i_2}{k}t^{i_2-k}(1-t^2)^{k/2}(1+(-1)^{i_1+i_2})\int_0^{\pi}\cos(\theta)^{i_1+i_2-k}\sin(\theta)^k d\theta.\end{aligned}$$

By changing variables to $u=\cos(\theta)$, we can calculate the last integral in terms of the Beta function. This gives

$$\int_{-1}^1 u^{i_1+i_2-k}(1-u^2)^{(k-1)/2}du = I_{i_1+i_2-k,(k-1)/2},$$

so that the complete result is

$$\sum_{k=0}^{i_2}\binom{i_2}{k}t^{i_2-k}(1-t^2)^{k/2}(1+(-1)^{i_1+i_2})I_{i_1+i_2-k,(k-1)/2}.$$

9.5.2. Integrating inner products over a spherical cap. Consider the integral

$$\int_{S^{n-1}} f(\langle w,z\rangle)\langle x,w\rangle^{i_1}\langle y,w\rangle^{i_2}d\omega_n(w).$$

In the following, we write $u=\langle x,z\rangle, v=\langle y,z\rangle$ and $t=\langle x,y\rangle$.

We first decompose the variables in parts parallel and orthogonal to z. This gives

$$\int_{S^{n-2}} \int_{-1}^{1} f(u)(u_1 u + \sqrt{1-u_1^2}\sqrt{1-u^2}\langle \zeta_1, \xi\rangle)^{i_1}(u_1 v + \sqrt{1-u_1^2}\sqrt{1-v^2}\langle \zeta_2, \xi\rangle)^{i_2}$$
$$(1-u_1^2)^{(n-3)/2} du_1 d\omega_{n-1}(\xi)$$
$$= \sum_{j_1=0}^{i_1}\sum_{j_2=0}^{i_2} \binom{i_1}{j_1}\binom{i_2}{j_2} u^{j_1} v^{j_2} (1-u^2)^{(i_1-j_1)/2}(1-v^2)^{(i_2-j_2)/2}$$
$$\int_{-1}^{1} f(u) u_1^{j_1+j_2}(1-u_1^2)^{(n+i_1+i_2-j_1-j_2-3)/2} du_1 \int_{S^{n-2}} \langle \zeta_1, \xi\rangle^{i_1-j_1}\langle \zeta_2, \xi\rangle^{i_2-j_2} d\omega_{n-1}(\xi),$$

where ξ is the part of w orthogonal to z, and ζ_1, ζ_2 are the parts of x and y orthogonal to z, all normalized to be unit vectors.

Recall from Section 9.5.1 that the last integral is a polynomial in

$$\langle \zeta_1, \zeta_2\rangle = \frac{t-uv}{\sqrt{1-u^2}\sqrt{1-v^2}}$$

if $i_1 + i_2 - j_1 - j_2$ is even, and 0 otherwise. Furthermore, the maximum degree of this polynomial is given by $\min\{i_1 - j_1, i_2 - j_2\}$ and every monomial with nonzero coefficient has a power which differs from this by an even number. In particular, this implies that the terms $(1-u^2)^{1/2}$ and $(1-v^2)^{1/2}$ in the denominator of $\langle \zeta_1, \zeta_2\rangle$ are canceled by similar terms in the summant, where such terms with even powers remain. Hence this is a polynomial in the inner products between x, y and z.

It remains to calculate the integral

$$\int_{-1}^{1} f(u) u^i (1-u^2)^j du.$$

We distinguish several cases. For the covering problem, we have $f(u) = \chi_{[r,1]}(u)$, with $r > 0$[†]. We can compute this in terms of the incomplete Beta function:

$$\begin{aligned}\int_r^1 u^i(1-u^2)^j du &= \int_0^1 u^i(1-u^2)^j du - \int_0^r u^i(1-u^2)^j du \\ &= \frac{1}{2}B(i/2+1/2, j+1) - \frac{1}{2}\int_0^{r^2} u^{i/2-1/2}(1-u)^j du \\ &= \frac{1}{2}B(i/2+1/2, j+1) - \frac{1}{2}B(i/2+1/2, j+1; r^2),\end{aligned}$$

where

$$B(a, b; x) = \int_0^x u^{a-1}(1-u)^{b-1} du$$

is the incomplete Beta function.

For the general polarization problem, it is interesting to consider polynomials and Riesz-s potentials for f. Polynomials increase i in the integral I_{ij} by the powers of the monomials, and will be considered in Section 9.7.

[†]For $-1 < r \leq 0$ the optimal number of caps required for a covering is 2 in any dimension.

9.6. Adaption to a $(2k-1)$-point bound

In Section 7.4, our computations on the second level of the Lasserre hierarchy for packing problems do not necessarily give better bounds than our computations on the three-point bound. This is mainly due to our inability to compute the second level of the hierarchy with polynomial degree higher than 16 (which takes about two weeks, while computing the three-point bound with polynomials of degree at most 40 takes about a day). In this section we show how the k-th level of the Lasserre hierarchy can be adapted to a $(2k-1)$-point bound. In Section 9.7, we give results for both the second level of the hierarchy and the corresponding 3-point bound.

Let us first give a simple proof that the k-th level of the hierarchy gives a lower bound on the number of points of an optimal configuration.

PROOF THAT $\mathrm{las}_k(f,E)$ GIVES A LOWER BOUND ON $|C|$. Let (K,K',a_0) be a feasible solution to $\mathrm{las}_k(f,E)$, and C a feasible configuration. Then

$$\sum_{\substack{S\subseteq C\\ |S|\leq 2k}} A_k(K)(S) = \sum_{\substack{J_1,J_2\subseteq C\\ |J_1|,|J_2|\leq k}} K(J_1,J_2) \geq 0,$$

because K is positive definite. Furthermore, we have for every point $y\in S^{n-1}$ that

$$\begin{aligned}
&\sum_{\substack{S\subseteq C\\ |S|\leq 2k-1}} A_k^y(K'_y)(S)\\
&= \sum_{x\in C} \sum_{\substack{J_1,J_2\subseteq C\\ |J_1|,|J_2|\leq k-1}} f(\langle x,y\rangle)K'(J_1,J_2,y) - \sum_{\substack{J_1,J_2\subseteq C\\ |J_1|,|J_2|\leq k-1}} EK(J_1,J_2,y)\\
&= \sum_{\substack{J_1,J_2\subseteq C\\ |J_1|,|J_2|\leq k-1}} K(J_1,J_2,y)\Big(\sum_{x\in C} f(\langle x,y\rangle) - E\Big) \geq 0,
\end{aligned}$$

because K' is slice-positive and

$$\sum_{x\in C} f(\langle x,y\rangle) \geq E$$

for every feasible configuration C.

Summing the constraints over all sets $S\subseteq C$ of size at most $2k$ thus gives

$$-|C| \leq -\sum_{\substack{S\subseteq C\\ |S|\leq 2k}} \left(A_k(K)(S) - \int_{S^{n-1}} A_k^y(K'_y)(S)d\omega_n(y) - a_0\chi_{\mathcal{I}_0}(S)\right) \leq -a_0,$$

and hence $a_0\leq |C|$. □

Note that in the proof, the only property we required of K is that summing the constraints over all sets $S\subseteq C$ of size at most $2k$ is nonnegative. Define the operator $B_k : C(\mathcal{I}_{k-1}\times\mathcal{I}_{k-1}\times\mathcal{I}_1)\to C(\mathcal{I}_{2k-1})$ by

$$B_k(F)(S) = \sum_{\substack{J_1,J_2\in\mathcal{I}_{k-1}, Q\in\mathcal{I}_1\\ J_1\cup J_2\cup Q=S}} F(J_1,J_2,Q).$$

Then for any slice-positive function $F \in C(\mathcal{I}_{k-1} \times \mathcal{I}_{k-1} \times \mathcal{I}_1)$, we have

$$\sum_{\substack{S \subseteq C \\ |S| \leq 2k-1}} B_k(F)(S) = \sum_{\substack{J_1, J_2, Q \subseteq C \\ |J_1|, |J_2| \leq k-1, |Q| \leq 1}} F(J_1, J_2, Q) \geq 0,$$

so replacing $A_k(K)$ by $B_k(F)$ still gives a valid proof.

This can easily be generalized to slice-positive functions in $C(\mathcal{I}_a \times \mathcal{I}_a \times \mathcal{I}_b)$ with $2a + b \leq 2k$, which, for $a = 1$, gives bounds reminiscent of the k-point bounds for packing problems [**89**].

The adaptation for packing problems converges when $a \to \infty$ by Theorem 2.8, and when $b \to \infty$ by [**6**], since the bound for (a, b) gives at least as good a bound as the cases $(a, 0)$ and $(1, b)$. More generally, the corresponding $(2a + b)$-point bounds will converge when $a \to \infty$ by Theorem 2.8, and proving that the bounds converge when $b \to \infty$ can most likely be done in a similar vein as for packing problems in [**6**].

9.7. Computations

To compute the second level of the hierarchy, we use the kernels of Section 3.4 and the extension of the Bachoc-Vallentin kernel to $\mathcal{I}_1$ to parametrize the problem (see Section 9.2). By Section 9.5, this results in polynomials in the inner products between points in the sets S on which the constraints should hold, and therefore we can again use Chapter 4 to write the problem as a semidefinite program.

In contrast to earlier chapters, the problem is not necessarily defined over an algebraic field. The integrals present in the problem can result in polynomials with powers of π in the coefficients. Instead, we use Arb [**76**] through the computer algebra system Nemo.jl [**56**] to compute the semidefinite program, and in particular the (incomplete) beta function, in high-precision ball-arithmetic. This also implies that we cannot use the rounding procedure of Chapter 6 to obtain an exact solution for the cases where the bounds are numerically sharp.

For covering problems, we denote by $\text{las}_{k,r,\varepsilon}$ the k-th level of the hierarchy for $f(u) = \chi_{[r,1]}(u)$ and $E = 1$, where ε is the maximum inner product allowed between distinct points in an independent set. That is, r is the covering radius expressed as an inner product. Similarly, we denote by $\Delta_{3,r,\varepsilon}$ the 3-point bound defined in Section 9.6 by taking $k = 2$. The polynomial degrees used for the computation will be mentioned in the captions of the tables and figures. For polarization problems, we mention the function and threshold explicitly.

The code used to compute the bounds in this section is available at [**97**].

9.7.1. Packing-covering in low dimensions. Hardin, Sloane and Smith computed in 1994 a number of putatively optimal coverings in dimension 3 [**66**]. Furthermore, it is known that the simplex is optimal in every dimension [**14**], and conjectured that the cross-polytope is optimal in every dimension (this has been proven only in dimension 3 [**55**] and 4 [**14**]).

In the computations in this section we always give both the result obtained by taking the minimum distance of the putatively optimal configuration as the bound on the minimum distance, and by taking the covering radius of the configuration. Typically, the configurations will be optimal for their covering radius, so that using their minimum distance as bound is allowed. In general, however, such an optimal

configuration is not known; the computations show how much the bound typically improves by using a better bound on the minimum distance.

In Table 9.7.1, we show the results for the configurations in dimension 3 of at most 12 points, together with the geometric bound of Fejes-Toth [**55**]. The bounds for the covering radii corresponding to $N = 4, 6$ and 12 points are numerically sharp when using the best possible bound on the minimum distance, but not when using a worse bound.

N	$\text{las}_{2,r,\varepsilon}$	$\text{las}_{2,r,r}$	$\Delta_{3,r,\varepsilon}$	$\Delta_{3,r,r}$	Fejes-Toth bound
4	3.9999	3.8648	3.9999	3.9186	4
5	4.4901	4.3034	4.8099	4.6218	4.6979
6	5.9999	5.7244	5.9999	5.8606	6
7	6.0292	5.9904	6.5112	6.3929	6.7808
8	6.74	6.6446	7.2053	7.136	7.5219
9	7.6334	7.2906	8.2108	7.9375	8.2051
10	8.8468	8.4219	9.2428	9.0739	9.5199
11	8.8348	8.6657	9.3997	9.3659	9.8982
12	11.999	10.297	11.999	11.341	12

TABLE 9.7.1. Lower bounds on the minimum number of points needed to cover the sphere with caps $B(x, r)$. The second step of the Lasserre hierarchy is calculated using polynomial degree 12, and the three-point bound is calculated with polynomial degree 32 (which corresponds to $d = 16$ in the construction of the three-point function).

In Table 9.7.2, we focus on the simplex in low dimensions. The simplex is known to be optimal in all dimensions, and for the spherical code and the energy minimization problems, the first level of the hierarchy is already sharp using polynomials of low degree. As can be seen here, this is not the case for the covering problem, even when using the optimal minimum distance.

The cross-polytope is conjectured to be optimal in every dimension, and recently it was shown to be optimal among antipodal configurations [**15**]. The second level of the Lasserre hierarchy numerically reproduces the results that it is optimal in dimension 3 and 4 when using the optimal minimum distance, but not beyond dimension 4, see Table 9.7.3. However, as with all bounds in this chapter, these bounds are valid if the bound on the maximum inner product is correct.

9.7.2. Sharp configurations. In this section we show results for several sharp configurations. Recall that a configuration is called sharp if it is a $(2t - 1)$-design with t distinct inner products between distinct points in the configuration [**34**].

In [**20**], upper and lower bounds are given for the best max-min and min-max polarization among spherical k-designs. For max-min polarization, which we also consider in this chapter, they used these bounds to give an alternative proof of optimality of the simplex, and the optimality of the cross-polytope among 2-designs under the condition that it is the optimal covering with $N = 2n$ points among all 2-designs.

n	$\text{las}_{2,r,\varepsilon}$	$\text{las}_{2,r,r}$	$\Delta_{3,r,\varepsilon}$	$\Delta_{3,r,r}$
3	3.9999	3.8648	3.9999	3.9186
4	4.5512	4.0707	4.4333	4.2137
5	4.1997	4.0068	4.4114	4.1677
6	4.0872	3.8642	4.3301	4.0292
7	3.9635	3.704	4.2297	3.8743
8	3.837	3.5576	4.1237	3.7323
9	3.714	3.4343	4.0171	3.6106
10	3.5995	3.3309	3.9114	3.5061

TABLE 9.7.2. Lower bounds on the minimum number of points needed to cover the sphere S^{n-1} with caps $B(x, 1/n)$, with $\varepsilon = -1/n$. For the second step of the Lasserre hierarchy we use polynomial degree 12, and for the three-point bound we use polynomial degree 32.

n	$\text{las}_{2,r,\varepsilon}$	$\text{las}_{2,r,r}$	$\Delta_{3,r,\varepsilon}$	$\Delta_{3,r,r}$
3	5.9999	5.7244	5.9999	5.8606
4	7.9999	6.4533	7.9872	6.899
5	7.2325	6.8552	7.7518	7.3547
6	7.7243	7.132	8.2825	7.6347
7	8.0521	7.2642	8.7457	7.7262
8	8.2825	7.3195	9.0416	7.7286
9	8.5289	7.3335	9.2657	7.6991
10	8.7691	7.3253	9.4589	7.6558

TABLE 9.7.3. Lower bounds on the minimum number of points needed to cover the sphere S^{n-1} with caps $B(x, 1/\sqrt{n})$, with $\varepsilon = 0$. For the second step of the Lasserre hierarchy we use polynomial degree 12, and for the three-point bound we use polynomial degree 32.

The volume bound is sharp for polynomials up to degree k for k-designs. In Table 9.7.4, we give bounds on the minimum number of points needed to reach the threshold given by the polarization of a number of sharp configurations, for polynomials of the form $(1+u)^k$. These polynomials form a basis of the cone of absolutely monotonic functions, which follows from [**141**, Theorem 9b, p.154].

We also computed the bound with the same parameters for the D_4 and E_6 root systems, which are not sharp configurations, and in those cases the bound was only sharp up to the design strength.

9.7.3. Dependence on the maximum inner product. As proven in Section 9.4, the bound on the minimum distance, or the maximum inner product, is fundamental to obtain nontrivial bounds through the Lasserre hierarchy for polarization. As shown in the previous sections, there are cases where the bounds are sharp when using the best possible bound on the minimum distance. In this section, we show the dependence of the bounds on ε for low degrees.

Configuration	n	N	design strength	k
Icosahedron	3	12	5	23
Schläfli	6	27	4	8
Simplex	4	5	2	6
	5	6	2	5
	6	7	2	4
	7	8	2	4
	8	9	2	4
	9	10	2	4
	10	11	2	3
Cross-polytope	4	8	3	11
	5	10	3	10
	6	12	3	9
	7	14	3	9
	8	16	3	8
	9	18	3	8
	10	20	3	8

TABLE 9.7.4. Several configuration together with the maximum k for which the bound $\text{las}_2(f, E)$ is numerically sharp for the potential function $f(u) = (1+u)^k$, with degree 10 polynomials. We use the minimum distance of the configuration as ε, and the polarization of the configuration as threshold E. The minima of the configurations are described in [**17**].

Figure 9.7.1 shows the three-point bound for covering as a function of the maximum inner product ε allowed for the putatively optimal configuration of 5 points in dimension 3, for degree $4, 5$ and 6 (from bottom to top). As can be seen, for high ε, the function is flat, and gives the volume bound. For fixed ε, the bound increases with the degree. For a given degree d, the bound decreases slowly for ε close to the minimum ε possible, and then decreases rapidly to the volume bound; the ε where the bound starts to decrease faster increases with the degree.

In Figure 9.7.2, we show a similar plot but for the second step of the Lasserre hierarchy, for the simplex in dimension 4. Note that the polynomial degrees used are much lower than for the three-point bound, and it is possible the graph for higher degree will have a similar shape as the graph of the three-point bound for ε not close to the best maximum inner product. However, for ε close to $-1/4$ the second level of the Lasserre hierarchy with degree 8 is already better than the three-point bounds with degree 12.

In Figure 9.7.3, a similar plot is shown for the second level of the Lasserre hierarchy for polarization, for the function $(1+u)^8$. The threshold is given by the minimum polarization of the simplex in dimension 4. The shape of the bound is very comparable to the shape of the Lasserre hierarchy for covering, for the same configuration. A key difference is that the ratio between the second level of the hierarchy with $\varepsilon = -1/4$ and the volume bound is much larger for this function compared to the covering problem.

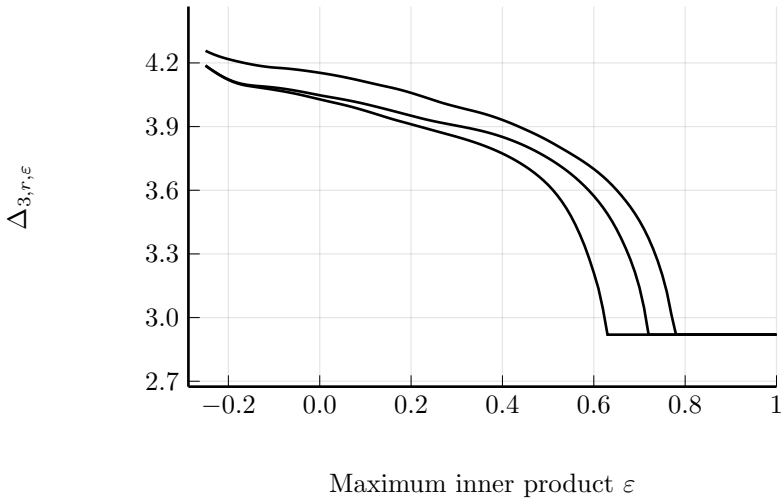


FIGURE 9.7.1. The three-point bound for covering in dimension 4 with covering radius $r = 1/4$, corresponding to the simplex, for varying minimum distance. From bottom to top, the lines correspond with $d = 4, 5, 6$, or equivalently, polynomial degree $8, 10$ and 12.

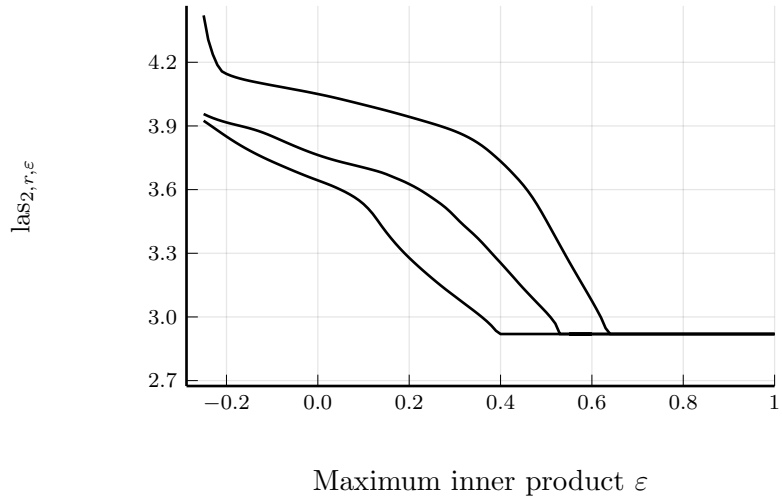


FIGURE 9.7.2. The second level of the Lasserre hierarchy for covering in dimension 4 with $r = 1/4$, for varying minimum distance. The optimal configuration is the regular simplex. From bottom to top, the lines correspond with polynomial degrees $4, 6$ and 8.

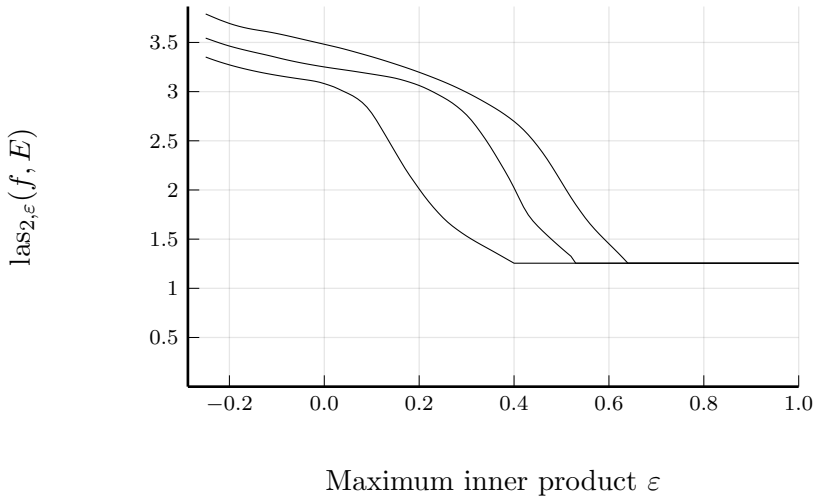


FIGURE 9.7.3. The second level of the Lasserre hierarchy for polarization with a threshold, where the potential function is $(1+u)^8$, where we use the polarization of the simplex in dimension 4 as threshold. From bottom to top, the lines correspond with polynomial degrees $4, 6$ and 8.

CHAPTER 10

n-point correlations in analytic number theory*

10.1. Introduction

In this chapter, we give universal bounds on the fraction of nontrivial zeros having given multiplicity for L-functions attached to a cuspidal automorphic representation of $\mathrm{GL}_m/\mathbb{Q}$. For this, we apply the higher-level correlation asymptotic of Hejhal [**69**] and Rudnick and Sarnak [**127**] in conjunction with semidefinite programming bounds.

Let $m \geq 1$ and let π be an irreducible cuspidal automorphic representation of $\mathrm{GL}_m/\mathbb{Q}$. Let $L(s,\pi)$ be its attached L-function (see [**75**] for theoretical background). Such functions generalize the classical Dirichlet L-functions (the case $m = 1$). Let $\rho_j = 1/2 + i\gamma_j$, with $j \in \mathbb{Z}$, be an enumeration of the nontrivial zeros of $L(s,\pi)$ repeated according to multiplicity. The Generalized Riemann Hypothesis (GRH) for $L(s,\pi)$ states that all its nontrivial zeros are aligned in the line $\operatorname{Re} s = 1/2$; that is, $\gamma_j \in \mathbb{R}$ for all j. Assuming GRH, we can enumerate the ordinates of the zeros in a nondecreasing fashion

$$\ldots \leq \gamma_{-2} \leq \gamma_{-1} < 0 \leq \gamma_1 \leq \gamma_2 \leq \ldots$$

and we have

$$N(T) := \sum_{\substack{j \geq 1 \\ \gamma_j \leq T}} 1 \sim \frac{m}{2\pi} T \log T.$$

Hejhal [**69**] and later on Rudnick and Sarnak [**126, 127**] investigated the n-level correlation distribution of the zeros of $L(s,\pi)$ guided by the relationship with Gaussian ensemble models in random matrix theory, pointed out by the breakthrough work of Montgomery [**111**]. This area has a long history, and we recommend [**127**], and the references therein, for more information.

Let $a_\pi(n)$ be the coefficients such that

$$L(s,\pi) = \sum_{n \geq 1} \frac{a_\pi(n)}{n^s}.$$

The results of Rudnick and Sarnack hold under the hypotheses that GRH holds for $L(s,\pi)$ and that

$$\sum_{p \text{ prime}} \frac{|a_\pi(p^k) \log p|^2}{p^k} < \infty$$

*This chapter is based on the publication "*F. Gonçalves, D. de Laat and N. Leijenhorst,* Multiplicity of nontrivial zeros of primitive L-functions via higher-level correlations*, Math. Comp. (2024), arXiv:2303.01095, doi:10.1090/mcom/4005*". The main difference is that Section 10.4 is generalized to $m > 1$, which allows us to prove the case $(n,m) = (3,2)$ of Theorem 10.1 without the additional assumption required for the case $(n,m) = (4,1)$.

for all $k \geq 2$. This second condition holds whenever $m \leq 3$ [**127**, Proposition 2.4]. They then show that

$$\frac{1}{N(T)} \sum_{\substack{j_1,\dots,j_n \geq 1 \text{ distinct} \\ \gamma_{j_1},\dots,\gamma_{j_n} \leq T}} f\left(\tfrac{m \log T}{2\pi}\gamma_{j_1}, \dots, \tfrac{m \log T}{2\pi}\gamma_{j_n}\right) \to \int_{\mathbb{R}^n} f(x) W_n(x) \delta\left(\frac{x_1 + \cdots + x_n}{n}\right) dx_1 \cdots dx_n \tag{10.1.1}$$

as $T \to \infty$, where

$$W_n(x) = \det\left[\frac{\sin(\pi(x_i - x_j))}{\pi(x_i - x_j)}\right]_{i,j=1,\dots,n}$$

is Dyson's limiting n-level correlation density for the Gaussian Unitary Ensemble (GUE) model, and f is any *admissible* test function satisfying

- $f \in C^\infty(\mathbb{R}^n)$, f is symmetric, and $f(x + t(1, \dots, 1)) = f(x)$ for $t \in \mathbb{R}$;
- $\overline{\text{supp}}(\widehat{f}) \subseteq \{x \in \mathbb{R}^n : |x_1| + \ldots + |x_n| < 2/m\}$ (in the distributional sense);
- $f(x) \to 0$ rapidly as $|x| \to \infty$ in the hyperplane $\sum_j x_j = 0$.

Let m_ρ denote the multiplicity of ρ, and define

$$N^*(T) := \sum_{\substack{j \geq 1 \\ \gamma_j \leq T}} m_{\rho_j}.$$

Montgomery's pair correlation approach has been used to obtain bounds of the form

$$N^*(T) \leq (c + o(1))N(T), \tag{10.1.2}$$

where for $m = 1$ the current smallest known value for c assuming the Riemann hypothesis is 1.3208 [**30, 31, 111, 112**]. Let $Z_n(T)$ count the number of nontrivial zeros of $L(s, \pi)$ with multiplicity at most $n - 1$ up to height T:

$$Z_n(T) = \sum_{\substack{j \geq 1 \\ \gamma_j \leq T,\, m_{\rho_j} \leq n-1}} 1.$$

Then $Z_2(T) \geq 2N(T) - N^*(T)$, so (10.1.2) can for instance be used to give a lower bound on the fraction of simple zeros, although the current best known bound is obtained using different techniques [**24, 30, 31, 45, 60**].

The goal of this chapter is to use the n-level correlations by Hejhal, Rudnick, and Sarnak (that is, Equation 10.1.1) to obtain bounds analogous to (10.1.2), and use this to obtain universal bounds on the fraction of nontrivial zeros of $L(s, \pi)$ having multiplicity at most $n - 1$.

THEOREM 10.1. *Let $L(s, \pi)$ be the L-function attached to an irreducible cuspidal automorphic representation π of $\mathrm{GL}_m/\mathbb{Q}$. Assuming GRH for $L(s, \pi)$ we have*

$$\liminf_{T \to \infty} \frac{Z_n(T)}{N(T)} \geq \begin{cases} 0.9614 & \text{if } n = 3 \text{ and } m = 1, \\ 0.2997 & \text{if } n = 3 \text{ and } m = 2, \\ 0.9787 & \text{if } n = 4 \text{ and } m = 1, \end{cases}$$

where the result for $(n, m) = (4, 1)$ holds under the additional assumption that certain series of rational functions are summed correctly using Maple; see Section 10.6.

As far as we know, the bounds in Theorem 10.1 are all new. Using a different method, it has been shown that at least 95.5% of the nontrivial zeros of the zeta function are simple or double zeros [**44**]. Our $(n, m) = (3, 1)$ case improves on this result. The case $m = 2$ is of special interest, since for this there are no previous effective bounds in the literature. Proving a positive proportion of zeros are simple in the case $(n, m) = (2, 2)$ is a notoriously difficult problem [**54**], and no positive lower bounds for $Z_2(T)/N(T)$ are known (in general), but there has been a great deal of interesting work on simple zeros for GL_2 L-functions (via Ω-results, for instance) [**9, 10, 11, 29, 43, 107, 135**]. Our methods do not seem to give positive lower bounds when $m \geq n$.

Montgomery's approach to finding a good value for c in (10.1.2) leads to an optimization problem that is similar to the one-dimensional version of the Cohn-Elkies bound [**33**] for the sphere packing problem. However, the higher-order correlation approach in this chapter is very different from the higher-order correlation approach for the sphere packing problem recently introduced in [**39**], although both approaches essentially restrict the support of the function or the Fourier transform to a compact set.

In Section 10.2 we set up the optimization problem to compute the above bounds, and we show how the optimal solution connects to reproducing kernels, inspired by the approach from [**28**].

We use a computational approach to obtain rigorous bounds. In Section 10.3 we use the symmetry and the rank-1 structure to find the bounds by solving linear systems. We then give two approaches for finding good solutions to the optimization problem. In Section 10.4 we use a parametrization using polynomials to find a good value when $n = 3$ and explain why this approach becomes problematic for $n > 3$. In Section 10.6 we give an alternative parametrization based on lattice shifts and use this to compute bounds for all cases in Theorem 10.1. Since the bounds using the polynomials are slightly better, the additional hypothesis in Theorem only applies to $(n, m) = (4, 1)$.

10.2. Derivation of the optimization problem

Let

$$M_n(T) = \sum_{\substack{j_1,\dots,j_n \geq 1 \text{ distinct} \\ \gamma_{j_1} = \dots = \gamma_{j_n} \leq T}} 1 = \sum_{\substack{j \geq 1 \\ \gamma_j \leq T}} (m_{\rho_j} - 1) \cdots (m_{\rho_j} - n + 1).$$

Then $M_1(T) = N(T)$ and $M_2(T) = N^*(T) - N(T)$. For a nonnegative admissible test function f, we have that

$$\frac{M_n(T)}{N(T)} f(1, \dots, 1) \leq \frac{1}{N(T)} \sum_{\substack{j_1,\dots,j_n \geq 1 \text{ distinct} \\ \gamma_{j_1},\dots,\gamma_{j_n} \leq T}} f(\tfrac{m \log T}{2\pi}\gamma_{j_1}, \dots, \tfrac{m \log T}{2\pi}\gamma_{j_n}),$$

where we use that

$$f(1, \dots, 1) = f(\tfrac{m \log T}{2\pi}\gamma_{j_1}, \dots, \tfrac{m \log T}{2\pi}\gamma_{j_n})$$

whenever $\gamma_{j_1} = \dots = \gamma_{j_n}$. Dividing by $f(1, \dots, 1)$ and applying (10.1.1) therefore gives

$$\limsup_{T \to \infty} \frac{M_n(T)}{N(T)} \leq c_{n,m},$$

where

$$c_{n,m} := \inf \frac{1}{f(1,\dots,1)} \int_{\mathbb{R}^n} f(x) W_n(x) \delta\Big(\frac{x_1+\ldots+x_n}{n}\Big)\, dx$$

and where the infimum is taken over nonnegative admissible test functions f.

We have

$$\begin{aligned}
&\int_{\mathbb{R}^n} f(x) W_n(x) \delta\bigg(\frac{x_1+\ldots+x_n}{n}\bigg)\, dx \\
&\quad = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} f(x+(x_n,\dots,x_n,0)) W_n(x+(x_n,\dots,x_n,0)) \\
&\qquad \cdot \delta\bigg(\frac{x_1+\ldots+x_{n-1}}{n}+x_n\bigg)\, dx_n dx_1 \cdots dx_{n-1} \\
&\quad = \int_{\mathbb{R}^{n-1}} f(x_1,\dots,x_{n-1},0) W_n(x_1,\dots,x_{n-1},0)\, dx_1 \cdots dx_{n-1}.
\end{aligned}$$

With $g(x_1,\dots,x_{n-1}) = f(x_1,\dots,x_{n-1},0)$, we see that $\overline{\operatorname{supp}}(\widehat{g})$ is now contained in the interior of $\frac{2}{m} H_{n-1}$, where

$$H_{n-1} = \{x \in \mathbb{R}^{n-1} : |x_1| + \ldots + |x_{n-1}| + |x_1+\ldots+x_{n-1}| \le 1\}.$$

Note that $H_{n-1} \subseteq [-1/2, 1/2]^{n-1}$. Now let

$$\nu_n(g) := \int_{\mathbb{R}^{n-1}} g(x) W_n(x,0)\, dx$$

for any $g \in L^1(\mathbb{R}^{n-1})$ satisfying

(i) $\overline{\operatorname{supp}}(\widehat{g}) \subseteq \frac{2}{m} H_{n-1}$;
(ii) g is nonnegative;
(iii) $g(0) = 1$.

It is now obvious that $c_{n,m} \ge \inf_g \nu_n(g)$, where the infimum is taken over functions g satisfying (i), (ii) and (iii). We show that $c_{n,m} = \inf_g \nu_n(g)$. Let $g \in L^1(\mathbb{R}^{n-1})$ satisfy the conditions above and assume further that g is invariant under the group G generated by

$$F_i g(x) = g(x - x_i(1,\dots,1) - x_i e_i),$$

for $i = 1,\dots,n-1$. Note that $F_i^2 = \mathrm{Id}$ and $F_i F_j F_i$ simply flips x_i by x_j. In particular, G is a finite group. Let $\varphi_\epsilon \in C^\infty(\mathbb{R}^{n-1})$ be a radial approximate identity as $\epsilon \to 0$, such that $\operatorname{supp}(\varphi_\epsilon) \subset \epsilon H_{n-1}$, $\int \varphi_\epsilon = 1$ and $\varphi_\epsilon \ge 0$. Let $g_\epsilon = g\widehat{\varphi}_\epsilon$, and observe that since $\widehat{g}_\epsilon = \widehat{g} * \varphi_\epsilon$, we then have that $g_\epsilon \in \mathcal{S}(\mathbb{R}^{n-1})$ and

$$\operatorname{supp}(\widehat{g}_\epsilon) \subset \bigg(\frac{2}{m} + \epsilon\bigg) H_{n-1}.$$

In particular, the function $f_\epsilon(x_1,\dots,x_n) = g_\epsilon(a(x_1-x_n),\dots,a(x_{n-1}-x_n))$, with $a = 1/(1+m\epsilon/2)$, is now an admissible test function for (10.1.1) and we obtain

$$c_{n,m} \le \int_{\mathbb{R}^n} f_\epsilon(x) W_n(x) \delta\Big(\frac{x_1+\ldots+x_n}{n}\Big)\, dx = \nu_n(g_\epsilon(a\cdot)) \le \nu_n(g(a\cdot)).$$

Taking $\epsilon \to 0$ we conclude that $c_{n,m} \le \inf_g \nu_n(g)$ over functions g satisfying (i), (ii), (iii) and invariant under G. Assume now g is not invariant under G. Since $W_n(x_1,\dots,x_{n-1},0)$ is G-invariant, is it clear that $\nu_n(g) = \nu_n(\widetilde{g})$, where

$$\widetilde{g}(x) = \frac{1}{|G|} \sum_{\rho \in G} \rho g(x),$$

and so $\nu_n(g) \geq c_{n,m}$.

Let $\mathrm{PW}(\Omega)$ be the Paley-Wiener space of functions $g \in L^2(\mathbb{R}^{n-1})$ whose Fourier transform is supported in Ω. Given a finite-dimensional subspace F of $\mathrm{PW}(\frac{1}{m}H_{n-1})$ and a basis $\{g_i\}$ of F, we can optimize over functions of the form

$$g(x) = \sum_{i,i'} X_{i,i'} g_i(x) g_{i'}(x), \tag{10.2.1}$$

where X is a positive semidefinite matrix. These are integrable functions satisfying conditions (i) and (ii). Since $\nu_n(g)$ and $g(0)$ are linear in the entries of X, this reduces the optimization problem to the semidefinite program

$$\begin{aligned} &\text{minimize} && \nu_n(g) \\ &\text{subject to} && g(0) = 1 \\ &&& X \succeq 0 \end{aligned} \tag{10.2.2}$$

Here $X \succeq 0$ means X is a positive semidefinite matrix of appropriate size. Since we optimize a linear functional over positive semidefinite matrices with affine constraints, this is a semidefinite program, which can be solved efficiently in practice.

We have that

$$\sum_{\substack{j \geq 1 \\ \gamma_j \leq T,\, m_{\rho_j} \geq n}} 1 \leq \frac{M_n(T)}{(n-1)!},$$

and therefore

$$\frac{Z_n(T)}{N(T)} \geq 1 - \frac{M_n(T)}{N(T)(n-1)!}.$$

Thus,

$$\liminf_{T \to \infty} \frac{Z_n(T)}{N(T)} \geq 1 - \frac{c_{n,m}}{(n-1)!},$$

which shows we can use the problem in (10.2.2) to obtain a lower bound on the fraction of zeros having multiplicity at most $n-1$.

As an alternative to $M_n(T)$ we could also use the parameters

$$N_k(T) = \sum_{\substack{j \geq 1 \\ \gamma_j \leq T}} m_{\rho_j}^{k-1},$$

which satisfy $N(T) = N_1(T)$ and $N^*(T) = N_2(T)$. We have

$$N_n(T) = \sum_{k=1}^{n} S(n,k) M_k(T), \tag{10.2.3}$$

where $S(n,k)$ are the Stirling numbers of the second kind. Since $S(n,k) \geq 0$ for all n and k, we can use our bounds on $c_{n,m}$, computed Section 10.4 and 10.6, to obtain

$$\limsup_{T \to \infty} \frac{N_n(T)}{N(T)} \leq \begin{cases} 2.0597 & \text{if } n = 3 \text{ and } m = 1, \\ 5.8984 & \text{if } n = 3 \text{ and } m = 2, \\ 3.8834 & \text{if } n = 4 \text{ and } m = 1, \end{cases} \tag{10.2.4}$$

provided the hypothesis of Theorem 10.1 hold. We also adapted the problem in (10.2.2) to bound $N_n(T)$ directly, but we were not able to improve over the results in (10.2.4) in this way.

We end this section with a possible formal solution to the problem of computing $c_{n,m}$. Note that the following result is true for H_{n-1} replaced by any compact set with a nonempty interior, with essentially the same proof.

THEOREM 10.2. *For $n, m > 0$ there exists a kernel $K\colon \mathbb{C}^{n-1} \times \mathbb{C}^{n-1} \to \mathbb{C}$ such that*

$$g(w) = \int_{\mathbb{R}^{n-1}} g(v)\overline{K(w,v)}\, d\nu_n(v) \tag{10.2.5}$$

for all $g \in \mathrm{PW}(\frac{1}{m}H_{n-1})$ and $w \in \mathbb{C}^{n-1}$. We have

$$c_{n,m} \le \frac{1}{K(0,0)},$$

and equality is attained if every nonnegative $g \in L^1(\mathbb{R}^n)$ with $\operatorname{supp}\widehat{g} \subseteq \frac{2}{m}H_{n-1}$ is a sum of squares in the sense that there are $g_{j,N} \in \mathrm{PW}(\frac{1}{m}H_{n-1})$ such that

$$g(x) = \lim_{N\to\infty} \sum_{j=1}^{N} |g_{j,N}(x)|^2, \tag{10.2.6}$$

with almost everywhere convergence.

PROOF. We have that $0 \le W_n(x) \le 1$ for all x, and $W_n(x) = 0$ if and only if $x_i = x_j$ for some $i \ne j$. Using induction it can be shown that for every $\epsilon > 0$ there exists $c(\epsilon) > 0$ such that $W_n(x) \ge c(\epsilon)$ whenever $|x_i - x_j| \ge \epsilon$ for all $i \ne j$. Define

$$S_\epsilon = \{x \in \mathbb{R}^{n-1} : |x_i| > \epsilon \text{ and } |x_i - x_j| > \epsilon \text{ for all distinct } i, j = 1, \dots, n-1\}.$$

Then $\nu_n(|g|^2) \le \|g\|_2^2$ and

$$\nu_n(|g|^2) \ge c(\epsilon) \int_{S_\epsilon} |g(x)|^2\, dx.$$

Observe that the set S_ϵ is relatively dense, that is, there is a cube Q (e.g., $Q = [-3\epsilon, 3\epsilon]^{n-1}$) and $\gamma > 0$ such that

$$\inf_{x\in\mathbb{R}^{b-1}} \mathrm{Vol}(S_\epsilon \cap (Q + x)) \ge \gamma \mathrm{Vol}(Q).$$

We could then apply the Logvinenko-Sereda Theorem [**68**, p. 112], which says there exists $b(\epsilon) > 0$ such that

$$\int_{S_\epsilon} |g(x)|^2\, dx \ge b(\epsilon)\|g\|_2^2,$$

proving the norms $\nu_n(|g|^2)$ and $\|g\|_2^2$ are equivalent. For completeness, we give a direct proof of the above inequality in our particular case.

We apply the Hardy-Littlewood-Sobolev theorem of fractional integration (see, e.g., [**136**, Section V.1.2]), which implies that there exists a constant $C > 0$ such that

$$\||h(\cdot)| \cdot |^{-1/4}\|_2 \le C\|\widehat{h}\|_{4/3}$$

for all $h \in L^2(\mathbb{R})$. If in addition $\overline{\operatorname{supp}}(\widehat{h}) \subseteq [-a, a]$, then using Hölder's inequality we obtain

$$\||h(\cdot)| \cdot |^{-1/4}\|_2 \le C\|\widehat{h}\|_{4/3} \le (2a)^{1/4}C\|\widehat{h}\|_2 = (2a)^{1/4}C\|h\|_2.$$

Then, with $h(x_i) = \widetilde{g}(x + x_j e_i)$ and $a = 1/2$, we get

$$\int_{\{x:|x_i-x_j|<\epsilon\}} |g(x)|^2\,dx \le \epsilon^{1/2} \int_{\{x:|x_i-x_j|<\epsilon\}} \left|\frac{g(x)}{|x_i - x_j|^{1/4}}\right|^2 dx \le C\epsilon^{1/2}\|g\|_2^2.$$

A similar procedure shows that $\int_{\{x:|x_i|<\epsilon\}} |g(x)|^2\,dx \le C\epsilon^{1/2}\|g\|_2^2$. Hence,

$$\begin{aligned}\int_{S_\epsilon} |g(x)|^2\,dx &\ge \|g\|_2^2 - \sum_{i=1}^{n-1} \int_{\{x:|x_i|<\epsilon\}} |g(x)|^2\,dx - \sum_{1\le i<j\le n} \int_{\{x:|x_i-x_j|<\epsilon\}} |g(x)|^2\,dx \\ &= \left(1 - \binom{n}{2} C\epsilon^{1/2}\right)\|g\|_2^2.\end{aligned}$$

By choosing $\epsilon > 0$ small enough the constant $1 - \binom{n}{2}C\epsilon^{1/2}$ is positive, which shows ν_n defines a norm equivalent to the L^2 norm. Hence, $\mathrm{PW}(\frac{1}{m}H_{n-1})$ is a Hilbert space with the inner product given by ν_n.

Using Fourier inversion, Cauchy's inequality, and Plancherel's theorem it can be shown that the evaluation functions $f \mapsto f(x)$ on $\mathrm{PW}(\frac{1}{m}H_{n-1})$ are continuous in the L^2 topology, so by the Riesz representation theorem there exists a kernel $K\colon \mathbb{C}^{n-1} \times \mathbb{C}^{n-1} \to \mathbb{C}$ with $K(x,\cdot) \in L^2(\mathbb{C}^{n-1})$ satisfying (10.2.5).

To finish, observe that if g satisfies conditions (i), (ii), and (iii) and condition (10.2.6), then by the monotone convergence theorem, limit (10.2.6) holds also in the $L^1(\mathbb{R}^{n-1})$-sense, and since $\mathrm{PW}(\frac{1}{m}H_{n-1}) \cap L^1(\mathbb{R}^{n-1})$ is a reproducing kernel space, it is simple to show that we actually have uniform convergence in compact sets. Then,

$$\begin{aligned}1 = g(0) = \lim_{N\to\infty} \sum_{j=1}^{N} |g_{j,N}(0)|^2 &= \lim_{N\to\infty} \sum_{j=1}^{N} |\nu_n(g_{j,N}(\cdot)K(0,\cdot))|^2 \\ &\le \lim_{N\to\infty} \sum_{j=1}^{N} \nu_n(|g_{j,N}|^2)K(0,0) = \nu_n(g)K(0,0).\end{aligned}$$

This shows $c_{n,m} \ge 1/K(0,0)$. Since $g(x) = K(0,x)^2/K(0,0)^2$ satisfies conditions (i), (ii), and (iii) and $\nu_n(g) = 1/K(0,0)$, we have $c_{n,m} \le 1/K(0,0)$, which completes the proof. □

It is worth mentioning that condition (10.2.6) holds true, and so equality $c_{n,m} = K(0,0)^{-1}$ is true, in the one-dimensional case (Hadamard factorization), which implies this condition is true in higher dimensions if H_{n-1} is replaced by a cube. Condition (10.2.6) also holds true for a ball (see [**52**]) and radial g, which is enough to show $c_{n,m} = K(0,0)^{-1}$. It is an open problem to prove condition (10.2.6) when H_{n-1} is replaced by a generic convex body. J. Vaaler, via personal communication, proposed an interesting way of obtaining (10.2.6). He conjectures the following: For any centrally symmetric convex body K there is a constant $c(K) \in (0,1)$ such that for any nonnegative $g \in \mathrm{PW}(K) \cap L^1$ there is $h \in \mathrm{PW}(K)$ such that

$$g \ge |h|^2 \quad \text{and} \quad \|h\|_2^2 \ge c(K)\|g\|_1.$$

If this conjecture is true, iterating this procedure ones finds functions $h_1, h_2, \ldots$ such that $g \ge |h_1|^2 + \ldots + |h_N|^2$ and

$$\|h_N\|_2^2 \ge c(K)(\|g\|_1 - \|h_1\|_2^2 - \ldots - \|h_{N-1}\|_2^2).$$

Then it is easy to see that we must have $g = \sum_{n\geq 1} |h_n|^2$ pointwise.

10.3. Symmetry and rank-1 structure

In this section, we show how we can use symmetry and a rank-1 structure to solve the problem in (10.2.2) efficiently.

Let G_n be the symmetry group of H_{n-1}; that is, let G_n be the group containing the matrices $A \in \mathbb{R}^{(n-1)\times(n-1)}$ with $AH_{n-1} = H_{n-1}$. Let

$$\Gamma_n = \{A^{\mathsf{T}} : A \in G_n\}.$$

If g is a function satisfying conditions (i), (ii), and (iii), then, for $\gamma \in \Gamma_n$, the function $L(\gamma)g$ defined by $L(\gamma)g(x) = g(\gamma^{-1}x)$, also satisfies these constraints and $\nu_n(g) = \nu_n(L(\gamma)g)$. It follows that the function $\overline{g}$ defined by

$$\overline{g}(x) = \frac{1}{|\Gamma_n|} \sum_{\gamma\in\Gamma_n} L(\gamma)g(x)$$

also satisfies these constraints and $\nu_n(\overline{g}) = \nu_n(g)$. Since $\overline{g}$ is Γ_n-invariant, this shows we may restrict to Γ_n-invariant functions, which can be parametrized using Section 3.2.

Suppose F is some Γ_n-invariant space of functions $\mathbb{R}^{n-1} \to \mathbb{R}$. Consider a complete set $\widehat{\Gamma}_n$ of irreducible, unitary representations of Γ_n. Let $\{g_{\pi,i,j}\}$ be a symmetry adapted system of the space F. By this we mean that $\{g_{\pi,i,j}\}$ is a basis of F such that $H_{\pi,i} := \operatorname{span}\{g_{\pi,i,j} : j = 1, \ldots, d_\pi\}$ is a Γ_n-irreducible representation with $H_{\pi,i}$ equivalent to $H_{\pi',i'}$ if and only if $\pi = \pi'$, and for each π, i and i' there exists a Γ_n-equivariant isomorphism $T_{\pi,i,i'} \colon H_{\pi,i} \to H_{\pi,i'}$ such that $g_{\pi,i',j} = Tg_{\pi,i,j}$ for all j. In other words, a symmetry adapted system is a basis of F according to a decomposition of F into Γ_n-irreducible subspaces, where the bases of equivalent irreducibles transform in the same way under the action of Γ_n.

Then, assuming the representations of Γ_n are of real type, we have the following block-diagonalization result: any function g that is a sum of squares of functions from F can be written as

$$g(x) = \sum_{\pi} \sum_{i,i'} X^{(\pi)}_{i,i'} \sum_{j} g_{\pi,i,j}(x) g_{\pi,i',j}(x),$$

for positive semidefinite matrices $X^{(\pi)}$. This follows from Schur's lemma as explained in [**58**] (see also Section 3.2 for a detailed proof).

To generate a concrete symmetry-adapted system $\{g_{\pi,i,j}\}$ for F we use an implementation of the projection algorithm as described in [**130**]. For this, first define the operators

$$p^{(\pi)}_{j,j'} = \frac{d_\pi}{|\Gamma|} \sum_{\gamma\in\Gamma} \pi(\gamma^{-1})_{j,j'} L(\gamma), \tag{10.3.1}$$

where d_π is the dimension of π and $L(\gamma)$ is the operator $L(\gamma)g(x) = g(\gamma^{-1}x)$. Then choose bases $\{g_{\pi,i,1}\}_i$ of $\operatorname{Im}(p^{(\pi)}_{1,1})$ and set $g_{\pi,i,j} = p^{(\pi)}_{j,1} g_{\pi,i,1}$. As shown in Section 3.2 this indeed gives a symmetry-adapted system of F.

LEMMA 10.3. *If π is not the trivial representation, then $g_{\pi,i,j}(0) = 0$.*

PROOF. Let $p_{j,j'}^{(\pi)}$ be the operator as defined in (10.3.1) and $g \in F$. Then,

$$(p_{j,1}^{(\pi)} p_{1,1}^{(\pi)} g)(x) = \frac{d_\pi^2}{|\Gamma|^2} \sum_{\beta,\gamma\in\Gamma} \pi(\beta^{-1})_{j,1} \pi(\gamma^{-1})_{1,1} g(\beta^{-1}\gamma^{-1}x),$$

and thus

$$(p_{j,1}^{(\pi)} p_{1,1}^{(\pi)} g)(0) = \frac{d_\pi^2}{|\Gamma|^2} \left(\sum_{\beta\in\Gamma} \pi(\beta^{-1})_{j,1} \right) \left(\sum_{\gamma\in\Gamma} \pi(\gamma^{-1})_{1,1} \right) g(0).$$

If π is not the trivial representation, then $\sum_{\gamma\in\Gamma} \pi(\gamma^{-1})_{j,j'} = 0$ by orthogonality of matrix coefficients. By linearity, the result then follows. □

This shows optimizing over functions of the form (10.2.1) is the same as optimizing over functions of the form

$$g(x) = \sum_{i,i'} X_{i,i'} g_{1,i,1}(x) g_{1,i',1}(x),$$

where X is positive semidefinite and where we denote the trivial representation by 1. That is, instead of using a Γ-invariant space of functions, we can use the much smaller space of Γ-invariant functions $g_i(x) = g_{1,i,1}(x)$. In other words, the problem in (10.2.2) reduces to the simpler semidefinite program

$$\begin{aligned} &\text{minimize} && \langle A, X\rangle \\ &\text{subject to} && \langle bb^\mathsf{T}, X\rangle = 1 \\ &&& X \succeq 0 \end{aligned} \tag{10.3.2}$$

where $A_{i,i'} = \nu_n(g_i g_{i'})$ and $b_i = g_i(0)$. Here we use the notation $\langle A, B\rangle = \operatorname{tr}(A^\mathsf{T} B)$.

A semidefinite program with m equality constraints has an optimal solution of rank r with $r(r+1)/2 \le m$ [**120**]. Since the above semidefinite program has only one equality constraint, we know that there exists an optimal solution of rank 1. In fact, the following lemma shows we can find this rank 1 solution by solving a linear system. We can view this lemma as a finite-dimensional version of Theorem 10.2.

LEMMA 10.4. *Let A be a positive semidefinite matrix and let x be a solution to the system $Ax = b$. Then $xx^\mathsf{T}/(x^\mathsf{T}b)^2$ is an optimal solution to the semidefinite program*

$$\begin{aligned} &\textit{minimize} && \langle A, X\rangle \\ &\textit{subject to} && \langle bb^\mathsf{T}, X\rangle = 1 \\ &&& X \succeq 0 \end{aligned}$$

PROOF. Let X be feasible to the semidefinite program and consider the spectral decomposition $X = \sum_i v_i v_i^\mathsf{T}$. Using Cauchy-Schwarz we have

$$1 = b^\mathsf{T} X b = \sum_i (b^\mathsf{T} v_i)^2 = \sum_i (x^\mathsf{T} A v_i)^2 \le \sum_i (x^\mathsf{T} A x)(v_i^\mathsf{T} A v_i) = x^\mathsf{T} b \langle A, X\rangle,$$

so

$$\langle A, X\rangle \ge \frac{1}{x^\mathsf{T} b} = \left\langle A, \frac{xx^\mathsf{T}}{(x^\mathsf{T}b)^2} \right\rangle.$$

□

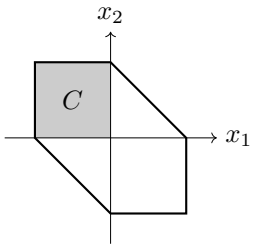


FIGURE 10.4.1. The hexagon H_2 and the square C (shaded).

10.4. Parametrization using polynomials

Here we consider the case $n = 3$. Let $\chi_{\frac{1}{m}H_2}(x)$ be the indicator function of the region $\frac{1}{m}H_2$ and define the functions $g_\alpha(x)$ for $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}_0^2$ by

$$\widehat{g}_\alpha(x) = x_1^{\alpha_1} x_2^{\alpha_2} \chi_{\frac{1}{m}H_2}(x).$$

Fix a degree d and let $F = \operatorname{span}\{g_\alpha : \alpha_1 + \alpha_2 \le d\}$. As before, we let $\{g_i\}$ be a basis for the Γ_n-invariant functions in F, and we use the parametrization

$$g(x) = \sum_{i,i'} X_{i,i'} g_i(x) g_{i'}(x).$$

Since each function g_i is a linear combination of the functions in F we have that $\widehat{g}_i(x) = p_i(x)\chi_{\frac{1}{m}H_2}(x)$ for some polynomial p_i of degree at most d; and as explained in Section 10.2.2 these polynomials p_i can be computed explicitly.

To set up the linear system for solving (5.1.1) we need to compute explicitly how the linear functionals $g(0)$ and $\nu_3(g)$ depend on the entries of the matrix X.

Set $a = \frac{1}{2m}$. The region $\frac{1}{m}H_2$ is the open hexagon with vertices $(a, 0)$, $(a, -a)$, $(0, -a)$, $(-a, 0)$, $(-a, a)$, $(0, a)$. Its symmetry group G_3 is the dihedral group of order 12 with generators

$$r = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad s = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

With $C = [-a, 0] \times [0, a]$ we have $\frac{1}{m}H_2 = C \cup r^2C \cup r^4C$; see Figure 10.4.1.

Since g_i is Γ_3-invariant, the polynomial $p_i(x)$ is G_3-invariant, so

$$\begin{aligned} g_i(0) &= \int_{\frac{1}{m}H_2} p_i(x)\,dx = \int_{\frac{1}{m}C} \big(p_i(x) + p_i(r^2x) + p_i(r^4x)\big)\,dx = 3\int_{\frac{1}{m}C} p_i(x)\,dx \\ &= 3\int_0^a \int_{-a}^0 p_i(x_1, x_2)\,dx_1 dx_2, \end{aligned}$$

which can be computed explicitly. This shows how we can write

$$g(0) = \sum_{i,i'} X_{i,i'} g_i(0) g_{i'}(0)$$

explicitly as a linear combination of the entries of the matrix X.

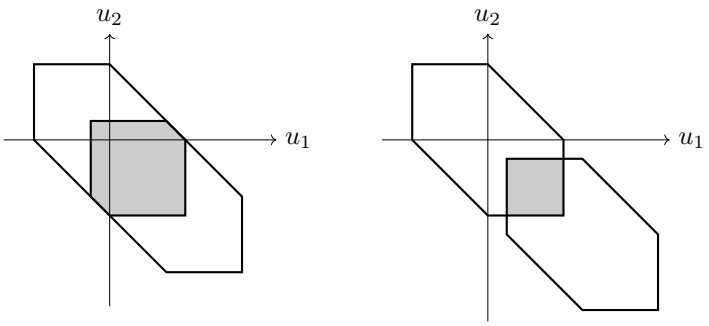


FIGURE 10.4.2. The possible shapes of the intersection of the supports of $g_i(u_1, u_2)$ and $g_{i'}(u_1 + x, u_2 - x)$ for $0 \le x \le 1/2$ (left) and $1/2 \le x \le 1$ (right).

As computed in Hejhal [**69**], we have

$$\begin{aligned}\nu_3(g) &= \int_{\mathbb{R}^2} g(x_1, x_2) W_3(x_1, x_2, 0)\, dx_1 dx_2 \\ &= 2 + \widehat{g}(0) + 6\int_0^1 \widehat{g}(x, 0)(x-1)\, dx - 12\int_0^1 \int_{-x_2}^0 \widehat{g}(x_1, x_2) x_2\, dx_1 dx_2.\end{aligned}$$

To compute this explicitly, note that

$$\widehat{g}(x) = \sum_{i,i'} X_{i,i'} (\widehat{g}_i * \widehat{g}_{i'})(x).$$

The term $\widehat{g}(0)$ can be computed similarly to how $g_i(0)$ is computed above, since

$$\widehat{g}(0) = \sum_{i,i'} X_{i,i'} \int_{H_2} p_i(x) p_{i'}(x)\, dx,$$

where we used that $p_{i'}(-x) = p_{i'}(x)$.

Since $\widehat{g}_i$ and $\widehat{g}_{i'}$ have bounded support, the convolution $(\widehat{g}_i * \widehat{g}_{i'})(x)$ is an integral over a bounded region whose shape depends on x. For the third term, symmetry gives

$$\int_0^1 \widehat{g}(x, 0)(x-1)\, dx = \int_0^1 \widehat{g}(x, -x)(x-1)\, dx.$$

Figure 10.4.2 shows the shape of the integration region used to compute $\widehat{g}(x, -x)$ for different values of x. This gives

$$\int_0^1 \widehat{g}(x, -x)(x-1)\, dx = \sum_{i,i} X_{i,i'} \int_0^{2a} \mathcal{J}_x\, p_i(u_1, u_2) p_{i'}(u_1 + x, u_2 - x)(x-1)\, du_1 du_2 dx,$$

where

$$\mathcal{J}_x = \begin{cases} \int_{-a+x}^0 \int_{-a-u_2}^{a-x} + \int_0^x \int_{-a}^{a-x} + \int_x^a \int_{-a}^{a-u_2} & 0 \le x \le a, \\ \int_{x-a}^a \int_{-a}^{a-x} & a \le x \le 2a. \end{cases}$$

Similarly, we split the convolution integral for the fourth term into the 4 regions $R_1 = \{-2a \le -x_2 \le x_1 \le -a\}$, $R_2 = \{-a \le a - x_2 \le x_1 \le 0\}$, $R_3 = \{0 \le -2x_1 \le$

TABLE 10.4.1. Upper bounds on the best possible value for $c_{3,1}$ and $c_{3,2}$ obtained using a parametrization using polynomials of degree d.

d	$c_{3,1}$	$c_{3,2}$
10	0.077516654	1.401390812
20	0.077222625	1.400561480
30	0.077206761	1.400505357
40	0.077200000	1.400480863
50	0.077198398	1.400474931
60	0.077197284	1.400470845

$x_2 \le a - x_1 \le 2a\}$ and $R_4 = \{0 \le -x_1 \le x_2 \le -2x_1 \le 2a\}$. Then the fourth term can be computed as

$$\int_0^1 \int_{-x_2}^0 \widehat{g}(x_1, x_2) x_2 \, dx_1 dx_2 = \sum_{i,i'} X_{i,i'} \int_0^{2a} \int_{-x_2}^0 \mathcal{I}_x \, p_i(u) p_{i'}(u-x) x_2 \, du_1 du_2 dx_1 dx_2,$$

where

$$\mathcal{I}_x = \begin{cases} \int_{x_2-a}^{a} \int_{-a}^{x_1+a} & (x_1, x_2) \in R_1, \\ \int_{x_2-a}^{a} \int_{-a-u_2+x_1+x_2}^{a-u_2} & (x_1, x_2) \in R_2, \\ \int_{x_1+x_2}^{a} \int_{-a}^{a-u_2} + \int_{-x_1}^{x_1+x_2} \int_{x_1+x_2-u_2-a}^{a-u_2} + \int_{x_2-a}^{-x_1} \int_{x_1+x_2-u_2-a}^{x_1+a} & (x_1, x_2) \in R_3, \\ \int_{-x_1}^{a} \int_{-a}^{a-u_2} + \int_{x_1+x_2}^{-x_1} \int_{-a}^{x_1+a} + \int_{x_2-a}^{x_1+x_2} \int_{x_1+x_2-u_2-a}^{x_1+a} & (x_1, x_2) \in R_4. \end{cases}$$

In Table 10.4.1 we show the bounds on $c_{3,1}$ and $c_{3,2}$ obtained with this parametrization for degree d. This proves the cases $(n, m) \in \{(3, 1), (3, 2)\}$ of Theorem 10.1. At [**97**][†] , we give the GAP source code to compute the polynomials p_i in exact arithmetic and a Julia file to compute the above bound rigorously by solving the linear system from Section 10.3 in ball arithmetic.

10.5. The Fourier transform of $\chi_{H_{n-1}}$

In Section 10.6, we give bounds using an alternative parametrization using phase shifts of the indicator functions of H_{n-1}. This requires, in particular, the Fourier transform of the indicator function of H_{n-1}.

To compute this, we use the Fourier transform of the indicator function of the simplex $S_n = \{x \in \mathbb{R}^n : x_j \ge 0,\ x_1 + \ldots + x_n \le 1\}$ as computed in [**13**, Eq. 18]:

$$\widehat{\chi}_{S_n}(y) = \frac{(-1)^{n+1}}{(2\pi i)^n} \sum_{j=1}^{n} \frac{1 - e^{-2\pi i y_j}}{y_j \prod_{i \ne j} (y_j - y_i)}.$$

[†]Compared to the code available with the arXiv version of the paper on which this chapter is based, the code allows for general m instead of only $m = 1$. This allows us to remove the additional assumption in Theorem 10.1 for the case $(n, m) = (3, 2)$, but does not change the result in the theorem.

The Fourier transform of the indicator function of the cross-polytope $C_n = \{x \in \mathbb{R}^n : |x_1| + \ldots + |x_n| \leq 1\}$ is given by

$$\widehat{\chi}_{C_n}(y) = \sum_{\sigma \in \{\pm 1\}^n} \widehat{\chi}_{S_n}(\sigma_1 y_1, \ldots, \sigma_n y_n).$$

If we consider each term in $\widehat{\chi}_{S_n}(y)$ individually, then it is simple to deduce that

$$\sum_{\sigma \in \{\pm 1\}^n} \frac{1 - e^{-2\pi i \sigma_j y_j}}{\sigma_j y_j \prod_{i \neq j} (\sigma_j y_j - \sigma_i y_i)} = \sum_{\sigma \in \{\pm 1\}} (2\sigma y_j)^{n-1} \frac{1 - e^{-2\pi i \sigma y_j}}{\sigma y_j \prod_{i \neq j} (y_j^2 - y_i^2)},$$

while the latter depends on the parity of n. This implies that

$$\widehat{\chi}_{C_n}(y) = \begin{cases} \pi^{-n}(-1)^{(n-1)/2} \sum_{j=1}^n \frac{y_j^{n-2} \sin(2\pi y_j)}{\prod_{i \neq j}(y_j^2 - y_i^2)} & \text{if } n \text{ is odd, and} \\ \pi^{-n}(-1)^{n/2-1} \sum_{j=1}^n \frac{2 y_j^{n-2} \sin(\pi y_j)^2}{\prod_{i \neq j}(y_j^2 - y_i^2)} & \text{if } n \text{ is even.} \end{cases}$$

Finally, notice that H_{n-1} can be identified with $C_n \cap \{x : x_1 + \ldots + x_n = 0\}$ and thus one can rigorously justify that

$$\widehat{\chi}_{H_{n-1}}(y_1, \ldots, y_{n-1}) = \lim_{T \to \infty} \int_{-T}^{T} \widehat{\chi}_{C_n}(y_1 + t, \ldots, y_{n-1} + t, t)\, dt$$

The function under the integral sign above only has simple poles at $t = -\frac{y_j + y_k}{2}$ for $k \neq j$ and $t = -y_j/2$ (for generic $y \in \mathbb{R}^{n-1}$). A routine application of the residue theorem and a grouping of terms shows that $\widehat{\chi}_{H_{n-1}}(y)$ is equal to

$$\frac{(-1)^{(n-1)/2}}{2^{n-2}\pi^{n-1}} \Bigg[- \sum_{1 \leq j < k \leq n-1} \frac{(y_j - y_k)^{n-3} \cos(\pi(y_j - y_k))}{y_j y_k \prod_{i \neq j,k}(y_j - y_i)(y_k - y_i)} + \sum_{j=1}^{n-1} \frac{y_j^{n-3} \cos(\pi y_j)}{\prod_{i \neq j}(y_j - y_i) y_i} \Bigg]$$

if n is odd, and equal to

$$\frac{(-1)^{n/2-1}}{2^{n-2}\pi^{n-1}} \Bigg[\sum_{1 \leq j < k \leq n-1} \frac{(y_j - y_k)^{n-3} \sin(\pi(y_j - y_k))}{y_j y_k \prod_{i \neq j,k}(y_j - y_i)(y_k - y_i)} + \sum_{j=1}^{n-1} \frac{y_j^{n-3} \sin(\pi y_j)}{\prod_{i \neq j}(y_j - y_i) y_i} \Bigg]$$

if n is even.

10.6. Parametrization using shifts

Now we consider the functions g_λ obtained by shifting the Fourier transform of the indicator function of $\frac{1}{m} H_{n-1}$ by $\lambda \in m\mathbb{Z}^{n-1}$:

$$g_\lambda(x) = \widehat{\chi_{\frac{1}{m} H_{n-1}}}(x - \lambda) = \frac{1}{m^{n-1}} \widehat{\chi_{H_{n-1}}}\left(\frac{1}{m}(x - \lambda)\right).$$

Then $g_\lambda(x)$ is the Fourier transform of $\chi_{\frac{1}{m} H_{n-1}}(\cdot) e^{-2\pi i \langle \lambda, \cdot \rangle}$, so $g_\lambda(\gamma x) = g_{\gamma^{-1}\lambda}(x)$ for all γ in the group Γ_n as defined in Section 10.3.

We observe that for $n = 2$, the functions g_λ for $\lambda \in \Lambda = m\mathbb{Z}^{n-1}$ form an orthogonal basis of $\mathrm{PW}(\frac{1}{m} H_{n-1})$ with respect to the Lebesgue L^2-norm. This, however, is not true for $n \geq 3$. For $n = 3$, a classical result of Venkov [**139**] shows that it is an orthogonal basis if we take Λ as the dual lattice of

$$\{(u/2 + v, u/2 - v/2) : u, v \in \frac{1}{m}\mathbb{Z}\}.$$

For $n \geq 4$, a recent result [**62, 100**] implies that no set of translations Λ exists for which $\{g_\lambda\}_{\lambda \in \Lambda}$ is an orthogonal basis of $\mathrm{PW}(\frac{1}{m} H_{n-1})$, otherwise H_{n-1} would

need to be centrally symmetric, have only centrally symmetric faces and tile the space by translations, but none of these properties hold. In any case, in practice, using $\Lambda = m\mathbb{Z}^{n-1}$ works reasonably well since $\{g_\lambda\}_{\lambda\in\Lambda}$ is an orthogonal basis of $\mathrm{PW}(\frac{1}{m}[-1/2,1/2]^{n-1})$ which contains $\mathrm{PW}(\frac{1}{m}H_{n-1})$.

Let F be a finite-dimensional, Γ_n-invariant subspace of $\mathrm{span}\{g_\lambda : \lambda \in m\mathbb{Z}^{n-1}\}$, and again use the parametrization

$$g(x) = \sum_{i,i'} X_{i,i'} g_i(x) g_{i'}(x),$$

where $\{g_i\}$ is a basis for the Γ_n-invariant functions in F.

To set up the linear system from Section 10.3 we need to compute the matrix $A_{i,i'} = \nu_n(g_i g_{i'})$ and the vector $b_i = g_i(0)$. By Lemma 10.4 the optimal function then is given by

$$g(x) = \frac{(\sum_i c_i g_i(x))^2}{(c^{\mathsf{T}} b)^2}$$

with c a solution to $Ac = b$.

In Section 10.5 we calculate the Fourier transform of $\chi_{H_{n-1}}(x)$ for general n. This function has removable singularities whenever $x_i = 0$ for some i or $x_i = x_j$ for some $i \neq j$. The function g_i is a linear combination of shifts of the Fourier transform, and thus can also have removable singularities. To compute $g_i(0)$ we consider each term individually, check whether the denominator evaluates to zero and if it does, we use repeated applications of L'Hôpital's rule.

To compute the entries $A_{i,i'} = \nu_n(g_i g_{i'})$, we need to compute integrals of the form $\int_{\mathbb{R}^{n-1}} f(x)\,dx$, where $f(x) = g_i(x) g_{i'}(x) W_n(x,0)$. To simplify the computations we change variables to $m^{-1}x$. Since the Fourier transform of $g_i(mx)$ is supported in $[-1/2,1/2]^{n-1}$ and the Fourier transform of $W_n(mx,0)$ is supported in $[-m,m]^{n-1}$, the Fourier transform of f is supported in $[-(m+1),(m+1)]^{n-1}$. By Poisson summation, we then have

$$\int_{\mathbb{R}^{n-1}} f(x)\,dx = \widehat{f}(0) = \sum_{k\in(m+1)\mathbb{Z}^{n-1}} \widehat{f}(k) = \frac{1}{(m+1)^{n-1}} \sum_{k\in\frac{1}{m+1}\mathbb{Z}^{n-1}} f(k).$$

By again using Poisson summation, we have

$$\begin{aligned}\sum_{k\in\frac{1}{m+1}\mathbb{Z}^{n-1}} f(k) &= (m+1)^{n-1}\widehat{f}(0)\\ &= (m+1)^{n-1} \sum_{k\in(m+1)\mathbb{Z}^{n-1}} \widehat{f}(k) e^{2\pi i\langle k,s\rangle}\\ &= \sum_{k\in\frac{1}{(m+1)}\mathbb{Z}^{n-1}} f(k+s)\end{aligned}$$

for any $s \in \mathbb{R}^{n-1}$. To avoid evaluation of f on a removable singularity, we set

$$s = \left(\frac{1}{m(m+1)n}, \frac{2}{m(m+1)n}, \ldots, \frac{n-1}{m(m+1)n}\right),$$

and compute the integral as

$$\int_{\mathbb{R}^{n-1}} f(x)\,dx = \frac{1}{(m+1)^{n-1}} \sum_{k\in\frac{1}{(m+1)}\mathbb{Z}^{n-1}} f(k+s).$$

Because of the sines and cosines in the formula of $\widehat{\chi}_{H_{n-1}}$ (see Appendix 10.5), each series $\sum_k f(k+s)$ can be split into $(2(m+1))^{n-1}$ series of rational functions over the lattice $\mathbb{Z}^{n-1}$, one for each value appearing for the sines and cosines.

Computing these series exactly is the computational bottleneck in our approach. Maple is the only system among the computer algebra systems we tried in which we could successfully compute the sums of all the series, and for this we needed several ad hoc tricks. For instance, for series over $\mathbb{Z}^3$, we first compute the outer series over $\mathbb{Z}$, and then use partial fraction decomposition to simplify the terms before summing over $\mathbb{Z}^2$. Here a complicating factor was that after partial fraction decomposition the series of some terms no longer converge, which meant we needed to recombine certain terms in ad hoc ways before summing over $\mathbb{Z}^2$.

Unfortunately, Maple 2022.2 contains a bug where it acts nondeterministically and sometimes gives a completely wrong answer. We, therefore, check that the sum computed with Maple deviates at most 0.1% from the sum we obtain by truncating the series to the box $[-C/(m+1), C/(m+1)]^{n-1}$, where for $n = 3$ we use $C = 400$ (or $C = 1300$ for the cases where this is not sufficient) and for $n = 4$ we use $C = 25$. We give text files containing the series in Maple syntax and for each series an interval containing the correct sum given by decimal expressions of a midpoint and a radius. We also give a Julia file with the code to obtain the rigorous bounds using the data in these text files.

For $(n, m) = (3, 1)$ and $(n, m) = (3, 2)$, Table 10.6.1 gives the results for the subspaces $F = \operatorname{span}\{g_\lambda : \lambda \in S_d\}$, where

$$S_d = \Gamma_n\{m\lambda \in \mathbb{Z}^{n-1} : \|\lambda\|_1 \leq d\}.$$

This proves the case $(3, 2)$ of Theorem 10.1 (the case $(3, 1)$ was already done in Section 10.4). For $(n, m) = (4, 1)$, we only compute the series exactly for $\lambda = 0$ (partly because of the aforementioned Maple bug), which gives $c_{4,1} \leq 447/3500 \approx 0.1277$ as used in Theorem 10.1. Numerically we estimate that by using more values of λ this can be improved to $c_{4,1} \leq 0.026$. This would improve the lower bound in Theorem 10.1 from 0.9787 to 0.9956.

TABLE 10.6.1. Upper bounds on the best possible value for $c_{n,m}$ obtained through a parametrization using shifts S_d^2.

d	$c_{3,1}$	$c_{3,2}$
0	0.144444445	1.488888889
1	0.077710979	1.401604735
2	0.077580416	1.401343568
3	0.077261926	1.400616000
4	0.077247720	1.400581457
5	0.077213324	1.400506625

CHAPTER 11

Discussion and future work

This chapter consists of a number of remarks, questions and possible directions for future research.

Lasserre hierarchies

In Chapter 2, we gave a framework that encompasses all known Lasserre hierarchies for problems in discrete geometry on compact spaces. It would be interesting to consider similar hierarchies for problems that do not immediately fit in this framework. One example is the extension to non-compact spaces, which is considered by Cohn and Salmon for the sphere packing problem [**41**]. In dimension 4, the optimal sphere packing is conjectured to be the D_4 root lattice. Our result that the second level of the Lasserre hierarchy is sharp for the kissing number in dimension 4, where we prove that the unique optimal solution is the D_4 root system, indicates that the second level of the Lasserre hierarchy for the sphere packing problem may also be sharp.

Currently, the only equality constraints that are used in the Lasserre hierarchies in discrete geometry are the cardinality constraints of Example 2.6. It would be interesting to consider a problem that requires different or additional equality constraints. Does this influence the finite convergence of the hierarchy? In the convergence proof, we obtain through Lemma 2.9 a measure of the form

$$\lambda = \int_R \chi_R d\sigma(R),$$

where $\chi_R = \sum_{Q \subseteq R} \delta_Q$ and σ is a probability measure. If such a measure satisfies $\lambda(g) = b$ for some continuous function g and $b \in \mathbb{R}$, that does not imply that

$$\chi_R(g) = b$$

almost everywhere on the support of σ, which would be required for the convergence proof. Therefore, the convergence of the hierarchy will depend on the exact equality constraints that are required.

Solving semidefinite programs

In Chapter 5, we introduced a high-precision semidefinite programming solver, which uses low-rank structures present in the problem. Further development of the solver would be useful. In particular, pre- and post-processing steps would simplify the use of the solver, as would a better integration with the Julia optimization ecosystem. Here, one could think of removing linearly dependent free variables and linearly dependent constraints.

Rounding

As with many heuristics, the rounding procedure of Chapter 6 can be further improved. In the first step, we use the uniqueness of the row-reduced echelon form to find the exact kernel vectors that describe the optimal face. However, even if there is a basis of the kernel of a rational matrix of relatively small bitsize, the row-reduced echelon form can have large bitsize, thereby requiring a high-precision solution for a good enough numerical approximation. It would be interesting to find a different method (with similar computational complexity) that would reduce the precision needed to find the correct kernel vectors.

The size of the final solution seems to depend mainly on the number of variables taken into account during the last step of the rounding procedure. Taking more variables into account gives a closer approximation to the numerical solution at the cost of a larger bitsize. If the small non-zero eigenvalues of the numerical solution matrix decrease through the transformation of Section 6.3, a better approximation is warranted, leading to a solution of larger bitsize. Since only the span of the first k basis vectors corresponding to the transformation is fixed, it might be possible to avoid such a decrease in the small eigenvalues by constructing a different transformation matrix.

Spherical Codes

There are many semidefinite programming bounds available for the spherical cap packing problem: the k-th level of the Lasserre hierarchy which gives $2k$-point bounds, the k-point bounds of [**89**] consisting of relaxations of the k-th level of the Lasserre hierarchy, the bounds described by Musin in [**115**], and the bounds which use the approximation of the copositive cone by Kuryatnikova and Vera [**83**]. Some of these bounds are known to be related, and such a relation has been used by Bekker and Oliveira in [**6**] to prove the convergence of the k-point bounds in [**89**]. However, typically the relation stated is between different levels of the hierarchy. It would be interesting to numerically compare the strength of the different hierarchies. With the solver of Chapter 5, it is now possible to compare the three and four-point bounds of the different hierarchies. Such a numerical study might give an indication which path of research would be most promising for other problems in discrete geometry.

Energy minimization

In Chapter 8, we conjecture that the second level of the Lasserre hierarchy is sharp for some families of configurations. In these cases, the potential is given by the harmonic energy potential, and the polynomial degree needed increases with the dimension. Because of this, the extrapolation techniques as used in [**90**] to prove asymptotic bounds cannot be used. However, maybe it is possible to avoid such an increase in the number of representations needed by using non-polynomial kernels, or to avoid an increase in the polynomial degree by using Hermite interpolation of fixed degree.

Analytic number theory

In Chapter 10, we use 3 and 4-point correlations to bound the fraction of non-trivial zeros of certain Dirichlet L-functions with multiplicity at most 2 and 3.

This is in contrast to the problems in discrete geometry, where k-point bounds give bounds on the same quantity as the 2-point bound. Is it possible to use k-point bounds with $k > 2$ to bound the fraction of simple zeros?

Bibliography

1. G. E. Andrews, R. Askey, and R. Roy, *Special Functions*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, Cambridge, 1999. doi:10.1017/CBO9781107325937
2. C. Bachoc and F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. **21** (2008), no. 3, 909–924. doi:10.1090/S0894-0347-07-00589-9
3. C. Bachoc and F. Vallentin, *Optimality and uniqueness of the* $(4, 10, 1/6)$ *spherical code*, J. Combin. Theory Ser. A **116** (2009), no. 1, 195–204. doi:10.1016/j.jcta.2008.05.001
4. B. Ballinger, G. Blekherman, H. Cohn, N. Giansiracusa, E. Kelly, and A. Schürmann, *Experimental study of energy-minimizing point configurations on spheres*, Experiment. Math. **18** (2009), no. 3, 257–283.
5. A. Barvinok, *A Course in Convexity*, Graduate Studies in Mathematics, vol. 54, American Mathematical Society, 2002.
6. B. Bekker and F. M. de Oliveira Filho, *On the convergence of the k-point bound for topological packing graphs*, arXiv:2306.02725, June 2023. doi:10.48550/arXiv.2306.02725
7. S. J. Benson and Y. Ye, *DSDP5 - software for semidefinite programming*, ACM Trans. Math. Software **34** (2008), no. 3, Art. 16, 20. doi:10.1145/1356052.1356057
8. J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A Fresh Approach to Numerical Computing*, SIAM Rev. **59** (2017), no. 1, 65–98. doi:10.1137/141000671
9. A. R. Booker, *Simple zeros of degree 2 L-functions*, J. Eur. Math. Soc. (JEMS) **18** (2016), no. 4, 813–823. doi:10.4171/JEMS/603
10. A. R. Booker, P. J. Cho, and M. Kim, *Simple zeros of automorphic L-functions*, Compos. Math. **155** (2019), no. 6, 1224–1243. doi:10.1112/s0010437x19007279
11. A. R. Booker, M. B. Milinovich, and N. Ng, *Quantitative estimates for simple zeros of L-functions*, Mathematika **65** (2019), no. 2, 375–399. doi:10.1112/s0025579318000530
12. B. Borchers, *CSDP, A C library for semidefinite programming*, Optimization Methods and Software **11** (1999), no. 1-4, 613–623. doi:10.1080/10556789908805765
13. B. Borda, *Lattice points in algebraic cross-polytopes and simplices*, Discrete Comput. Geom. **60** (2018), no. 1, 145–169. doi:10.1007/s00454-017-9946-z
14. K. Böröczky, Jr., *Finite packing and covering*, Cambridge Tracts in Mathematics, vol. 154, Cambridge University Press, Cambridge, 2004. doi:10.1017/CBO9780511546587
15. S. Borodachov, *Optimal Antipodal Configuration of* $2d$ *Points on a Sphere in* $\mathbb{R}^d$ *for Covering*, arXiv:2210.12472, October 2022. doi:10.48550/arXiv.2210.12472
16. S. Borodachov, *Polarization Problem on a Higher-Dimensional Sphere for a Simplex*, Discrete Comput Geom **67** (2022), no. 2, 525–542. doi:10.1007/s00454-021-00308-1
17. S. Borodachov, *Absolute minima of potentials of a certain class of spherical designs*, Applied and Numerical Harmonic Analysis, Volume dedicated to Edward Saff's 80-th birthday (A. Martinez-Finkelshtein, A. Stokols, D. Bilyk, and E. Jacob, eds.), Springer, 2025, pp. 101–129.
18. S. V. Borodachov, D. P. Hardin, and E. B. Saff, *Discrete Energy on Rectifiable Sets*, Springer Monographs in Mathematics, Springer, New York, NY, 2019. doi:10.1007/978-0-387-84808-2
19. S. P. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, UK ; New York, March 2004.
20. P. G. Boyvalenkov, P. D. Dragnev, D. P. Hardin, E. B. Saff, and M. M. Stoyanova, *On polarization of spherical codes and designs*, Journal of Mathematical Analysis and Applications **524** (2023), no. 1, 127065. doi:10.1016/j.jmaa.2023.127065
21. A. E. Brouwer, *The uniqueness of the strongly regular graph on 77 points*, J. Graph Theory **7** (1983), no. 4, 455–461. doi:10.1002/jgt.3190070411

22. A. E. Brouwer and W. H. Haemers, *The Gewirtz Graph: An Exercise in the Theory of Graph Spectra*, European Journal of Combinatorics **14** (1993), no. 5, 397–407. doi:10.1006/eujc.1993.1044
23. A. E. Brouwer and H. Van Maldeghem, *Strongly regular graphs*, Encyclopedia of Mathematics and Its Applications, vol. 182, Cambridge University Press, Cambridge, 2022. doi:10.1017/9781009057226
24. H. M. Bui and D. R. Heath-Brown, *On simple zeros of the Riemann zeta-function*, Bull. Lond. Math. Soc. **45** (2013), no. 5, 953–961. doi:10.1112/blms/bdt026
25. A. Calderbank, P. Cameron, W. Kantor, and J. Seidel, *$\mathbb{Z}_4$-kerdock codes, orthogonal spreads, and extremal euclidean line-sets*, Proceedings of the London Mathematical Society. Third series **75** (1997), no. 2, 436–480. doi:10.1112/S0024611597000403
26. P. J. Cameron, J. M. Goethals, and J. J. Seidel, *Strongly regular graphs having strongly regular subconstituents*, Journal of Algebra **55** (1978), no. 2, 257–280. doi:10.1016/0021-8693(78)90220-X
27. P. J. Cameron and J. J. Seidel, *Quadratic forms over* GF(2), Indagationes Mathematicae (Proceedings) **76** (1973), no. 1, 1–8. doi:10.1016/1385-7258(73)90014-0
28. E. Carneiro, V. Chandee, F. Littmann, and M. B. Milinovich, *Hilbert spaces and the pair correlation of zeros of the Riemann zeta-function*, J. Reine Angew. Math. **725** (2017), 143–182. doi:10.1515/crelle-2014-0078
29. E. Carneiro, A. Chirre, and M. B. Milinovich, *Hilbert spaces and low-lying zeros of L-functions*, Adv. Math. **410** (2022), Paper No. 108748, 48. doi:10.1016/j.aim.2022.108748
30. A. Y. Cheer and D. A. Goldston, *Simple zeros of the Riemann zeta-function*, Proc. Amer. Math. Soc. **118** (1993), no. 2, 365–372. doi:10.2307/2160310
31. A. Chirre, F. Gonçalves, and D. de Laat, *Pair correlation estimates for the zeros of the zeta function via semidefinite programming*, Advances in Mathematics **361** (2020), 106926. doi:10.1016/j.aim.2019.106926
32. H. Cohn, *Table of spherical codes*, `https://dspace.mit.edu/handle/1721.1/153543`, February 2024.
33. H. Cohn and N. Elkies, *New upper bounds on sphere packings. I*, Ann. of Math. (2) **157** (2003), no. 2, 689–714. doi:10.4007/annals.2003.157.689
34. H. Cohn and A. Kumar, *Universally optimal distribution of points on spheres*, J. Amer. Math. Soc. **20** (2007), no. 1, 99–148. doi:10.1090/S0894-0347-06-00546-7
35. 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. doi:10.4007/annals.2017.185.3.8
36. H. Cohn, A. Kumar, and G. Minton, *Optimal simplices and codes in projective spaces*, Geometry & Topology **20** (2016), no. 3, 1289–1357. doi:10.2140/gt.2016.20.1289
37. H. Cohn, D. de Laat, and N. Leijenhorst, *Data for "Optimality of spherical codes via exact semidefinite programming bounds"*, `https://hdl.handle.net/1721.1/153928`, 2024.
38. H. Cohn, D. de Laat, and N. Leijenhorst, *Optimality of spherical codes via exact semidefinite programming bounds*, arXiv:2403.16874, March 2024.
39. H. Cohn, D. de Laat, and A. Salmon, *Three-point bounds for sphere packing*, arXiv:2206.15373, June 2022.
40. H. Cohn and A. Li, *Improved kissing numbers in seventeen through twenty-one dimensions*, arXiv:2411.04916, November 2024. doi:10.48550/arXiv.2411.04916
41. H. Cohn and A. Salmon, *Sphere packing bounds via rescaling*, arXiv:2108.10936, August 2021. doi:10.48550/arXiv.2108.10936
42. H. Cohn and J. Woo, *Three-point bounds for energy minimization*, J. Amer. Math. Soc. **25** (2012), no. 4, 929–958. doi:10.1090/S0894-0347-2012-00737-1
43. J. B. Conrey and A. Ghosh, *Simple zeros of the Ramanujan τ-Dirichlet series*, Invent. Math. **94** (1988), no. 2, 403–419. doi:10.1007/BF01394330
44. J. B. Conrey, A. Ghosh, and S. M. Gonek, *Mean values of the Riemann zeta-function with application to the distribution of zeros*, Number Theory, Trace Formulas and Discrete Groups (Oslo, 1987), Academic Press, Boston, MA, 1989, pp. 185–199. doi:10.1016/B978-0-12-067570-8.50017-2
45. J. B. Conrey, A. Ghosh, and S. M. Gonek, *Simple zeros of the Riemann zeta-function*, Proc. London Math. Soc. (3) **76** (1998), no. 3, 497–522. doi:10.1112/S0024611598000306

46. J. H. Conway and N. J. A. Sloane, *The cell structures of certain lattices*, Miscellanea Mathematica, Springer, Berlin, 1991, pp. 71–107.
47. P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geom Dedicata **6** (1977), no. 3, 363–388. doi:10.1007/BF03187604
48. M. M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, Algorithms and Combinatorics, vol. 15, Springer Berlin Heidelberg, Berlin, Heidelberg, 1997. doi:10.1007/978-3-642-04295-9
49. J. Dieudonné, *Sur la Séparation des Ensembles Convexes.*, Mathematische Annalen **163** (1966), 1–3.
50. J. D. Dixon, *Exact solution of linear equations using P-adic expansions*, Numer. Math. **40** (1982), no. 1, 137–141. doi:10.1007/BF01459082
51. M. Dostert, D. de Laat, and P. Moustrou, *Exact Semidefinite Programming Bounds for Packing Problems*, SIAM J. Optim. **31** (2021), no. 2, 1433–1458. doi:10.1137/20M1351692
52. A. V. Efimov, *An analog of Rudin's theorem for continuous radial positive definite functions of several variables*, Proc. Steklov Inst. Math. **284** (2014), S79–S86. doi:10.1134/S0081543814020072
53. T. Ericson and V. Zinoviev, *Codes on Euclidean spheres*, North-Holland Mathematical Library, vol. 63, North-Holland Publishing Co., Amsterdam, 2001.
54. A. de Faveri, *Simple zeros of* GL(2) *L-functions*, Journal of the European Mathematical Society (2024). doi:10.4171/jems/1559
55. L. Fejes Tóth, G. Fejes Tóth, and W. Kuperberg, *Lagerungen: Arrangements in the Plane, on the Sphere, and in Space*, Grundlehren Der Mathematischen Wissenschaften, vol. 360, Springer International Publishing, Cham, 1953. doi:10.1007/978-3-031-21800-2
56. C. Fieker, W. Hart, T. Hofmann, and F. Johansson, *Nemo/Hecke: Computer algebra and number theory packages for the Julia programming language*, ISSAC'17—Proceedings of the 2017 ACM International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2017, arXiv:1705.06134, pp. 157–164. doi:10.1145/3087604.3087611
57. W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991. doi:10.1007/978-1-4612-0979-9
58. K. Gatermann and P. A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, Journal of Pure and Applied Algebra **192** (2004), no. 1, 95–128. doi:10.1016/j.jpaa.2003.12.011
59. A. Gewirtz, *The uniqueness of g(2,2,10,56)*, Trans. New York Acad. Sci. **31** (1969), no. 6 Series II, 656–675. doi:10.1111/j.2164-0947.1969.tb01990.x
60. D. A. Goldston, S. M. Gonek, A. E. Özlük, and C. Snyder, *On the pair correlation of zeros of the Riemann zeta-function*, Proc. London Math. Soc. (3) **80** (2000), no. 1, 31–49. doi:10.1112/S0024611500012211
61. T. Gorin and G. V. López, *Monomial integrals on the classical groups*, J. Math. Phys. **49** (2008), no. 1, 013503, 20. doi:10.1063/1.2830520
62. R. Greenfeld and N. Lev, *Fuglede's spectral set conjecture for convex polytopes*, Anal. PDE **10** (2017), no. 6, 1497–1538. doi:10.2140/apde.2017.10.1497
63. K. I. Gross and R. A. Kunze, *Finite-dimensional induction and new results on invariants for classical groups. I*, Amer. J. Math. **106** (1984), no. 4, 893–974. doi:10.2307/2374328
64. M. Halická, E. de Klerk, and C. Roos, *On the convergence of the central path in semidefinite optimization*, SIAM J. Optim. **12** (2002), no. 4, 1090–1099. doi:10.1137/S1052623401390793
65. A. R. Hammons, Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, *The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes*, IEEE Trans. Inform. Theory **40** (1994), no. 2, 301–319. doi:10.1109/18.312154
66. R. H. Hardin, N. J. A. Sloane, and W. D. Smith, *Spherical Codes*, 1994, `https://cohn.mit.edu/sloane/`.
67. T. Hartman, D. Mazáč, and L. Rastelli, *Sphere packing and quantum gravity*, J. High Energ. Phys. **2019** (2019), no. 12, 48. doi:10.1007/JHEP12(2019)048
68. V. Havin and B. Jöricke, *The uncertainty principle in harmonic analysis*, Ergebnisse Der Mathematik Und Ihrer Grenzgebiete (3), vol. 28, Springer-Verlag, Berlin, 1994. doi:10.1007/978-3-642-78377-7
69. D. A. Hejhal, *On the triple correlation of zeros of the zeta function*, International Mathematics Research Notices **1994** (1994), no. 7, 293–302. doi:10.1155/S1073792894000334
70. C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM J. Optim. **6** (1996), no. 2, 342–361. doi:10.1137/0806020

71. D. G. Higman and C. C. Sims, *A simple group of order* $44,352,000$, Math. Z. **105** (1968), 110–113. doi:10.1007/BF01110435
72. R. D. Hill and S. R. Waters, *On the cone of positive semidefinite matrices*, Linear Algebra Appl. **90** (1987), 81–88. doi:10.1016/0024-3795(87)90307-7
73. A. J. Hoffman and R. R. Singleton, *On Moore graphs with diameters 2 and 3*, IBM J. Res. Develop. **4** (1960), 497–504. doi:10.1147/rd.45.0497
74. E. Hubert and G. Labahn, *Rational invariants of scalings from Hermite normal forms*, ISSAC 2012—Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2012, pp. 219–226. doi:10.1145/2442829.2442862
75. H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. doi:10.1090/coll/053
76. F. Johansson, *Arb: Efficient Arbitrary-Precision Midpoint-Radius Interval Arithmetic*, IEEE Transactions on Computers **66** (2017), no. 8, 1281–1292. doi:10.1109/TC.2017.2690633
77. A. M. Kerdock, *A class of low-rate nonlinear binary codes*, Information and Control **20** (1972), no. 2, 182–187. doi:10.1016/S0019-9958(72)90376-2
78. V. L. Klee, *Separation Properties of Convex Cones*, Proceedings of the American Mathematical Society **6** (1955), no. 2, 313–318. doi:10.2307/2032366
79. M. H. Klin and A. J. Woldar, *The strongly regular graph with parameters (100,22,0,6): Hidden history and beyond*, Acta Univ. M. Belii Ser. Math. **25** (2017), 5–62.
80. M. Kojima, S. Shindoh, and S. Hara, *Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices*, SIAM J. Optim. **7** (1997), no. 1, 86–125. doi:10.1137/S1052623494269035
81. H. König, *Isometric imbeddings of Euclidean spaces into finite dimensional l_p-spaces*, Banach Center Publications **34** (1995), 79–87. doi:10.4064/-34-1-79-87
82. J. Kupka and K. Prikry, *The measurability of uncountable unions*, Amer. Math. Monthly **91** (1984), no. 2, 85–97. doi:10.2307/2322101
83. 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. doi:10.1016/j.endm.2017.10.052
84. D. de Laat, *Moment methods in extremal geometry*, Ph.D. thesis, Delft University of Technology, 2016, `https://repository.tudelft.nl/record/uuid:fce81f72-8261-484d-b9f4-d3d0c26f0473`.
85. D. de Laat, *Moment methods in energy minimization: New bounds for Riesz minimal energy problems*, Trans. Amer. Math. Soc. **373** (2019), no. 2, 1407–1453. doi:10.1090/tran/7976
86. D. de Laat, N. Leijenhorst, and W. H. H. de Muinck Keizer, *Data for "Optimality and uniqueness of the D_4 root system"*, doi:10.4121/74ce1c25-6fca-4680-8a36-e9c18e7e9594, 2024.
87. D. de Laat, N. Leijenhorst, and W. H. H. de Muinck Keizer, *Data for "Traceless projections of tensors with applications to hierarchies in discrete geometry"*, doi:10.4121/72a584f9-0146-4def-a203-fbdca05b641c, 2025.
88. D. de Laat, N. M. Leijenhorst, and W. H. H. de Muinck Keizer, *Traceless projection of tensors with applications to hierarchies in discrete geometry*, In preparation.
89. 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, 533–567. doi:10.1007/s10107-021-01638-x
90. D. de Laat, W. H. H. de Muinck Keizer, and F. C. Machado, *The Lasserre hierarchy for equiangular lines with a fixed angle*, arXiv:2211.16471, November 2022.
91. 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), e23. doi:10.1017/fms.2014.24
92. D. de Laat and F. Vallentin, *A semidefinite programming hierarchy for packing problems in discrete geometry*, Math. Program. **151** (2015), no. 2, 529–553. doi:10.1007/s10107-014-0843-4
93. W. Landry and D. Simmons-Duffin, *Scaling the semidefinite program solver SDPB*, arXiv:1909.09745, September 2019.
94. J. B. Lasserre, *Global Optimization with Polynomials and the Problem of Moments*, SIAM J. Optim. **11** (2001), no. 3, 796–817. doi:10.1137/S1052623400366802
95. J. B. Lasserre, *An explicit equivalent positive semidefinite program for nonlinear 0-1 programs*, SIAM J. Optim. **12** (2002), no. 3, 756–769. doi:10.1137/S1052623400380079

96. M. Laurent, *A Comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre Relaxations for 0-1 Programming*, Mathematics of Operations Research **28** (2003), no. 3, 470–496.
97. N. Leijenhorst, *Code accompanying the thesis "On the computation of three and four-point bounds in discrete geometry and analytic number theory"*, available at 4TU.ResearchData, 2025.
98. N. Leijenhorst and D. de Laat, *Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput. **16** (2024), no. 3, 503–534. doi:10.1007/s12532-024-00264-w
99. A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, *Factoring polynomials with rational coefficients*, Math. Ann. **261** (1982), no. 4, 515–534. doi:10.1007/BF01457454
100. N. Lev and M. Matolcsi, *The Fuglede conjecture for convex domains is true in all dimensions*, Acta Math. **228** (2022), no. 2, 385–420. doi:10.4310/acta.2022.v228.n2.a3
101. V. I. Levenshtein, *Bounds for the maximal cardinality of a code with bounded modules of the inner product*, Soviet Math. Dokl. **25** (1982), no. 2, 526–531.
102. J. Lofberg and P. Parrilo, *From coefficients to samples: A new approach to SOS optimization*, 2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601) (Nassau), IEEE, 2004, pp. 3154–3159. doi:10.1109/CDC.2004.1428957
103. 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. doi:10.1080/10586458.2017.1286273
104. A. L. Mackay, *The packing of three-dimensional spheres on the surface of a four-dimensional hypersphere*, J. Phys. A: Math. Gen. **13** (1980), no. 11, 3373. doi:10.1088/0305-4470/13/11/011
105. S. Mehrotra, *On the Implementation of a Primal-Dual Interior Point Method*, SIAM J. Optim. **2** (1992), no. 4, 575–601. doi:10.1137/0802028
106. D. M. Mesner, *An investigation of certain combinatorial properties of partially balanced incomplete block experimental designs and association schemes, with a detailed study of designs of Latin square and related types*, Ph.D. thesis, Michigan State University, 1923. doi:10.25335/3Z1W-1B88
107. M. B. Milinovich and N. Ng, *Simple zeros of modular L-functions*, Proc. Lond. Math. Soc. (3) **109** (2014), no. 6, 1465–1506. doi:10.1112/plms/pdu041
108. H. D. Mittelmann and F. Vallentin, *High accuracy semidefinite programming bounds for kissing numbers*, Expt. Math **19** (2010), no. 2, 175–179. doi:10.1080/10586458.2010.10129070
109. D. Monniaux and P. Corbineau, *On the generation of Positivstellensatz witnesses in degenerate cases*, Interactive Theorem Proving, Lecture Notes in Comput. Sci., vol. 6898, Springer, Heidelberg, 2011, `http://arxiv.org/abs/1105.4421`, pp. 249–264.
110. R. D. C. Monteiro, *Primal-dual path-following algorithms for semidefinite programming*, SIAM J. Optim. **7** (1997), no. 3, 663–678. doi:10.1137/S1052623495293056
111. H. L. Montgomery, *The pair correlation of zeros of the zeta function*, Proceedings of Symposia in Pure Mathematics, vol. 24, American Mathematical Society, Providence, Rhode Island, 1973, pp. 181–193. doi:10.1090/pspum/024/9944
112. H. L. Montgomery, *Distribution of the zeros of the Riemann zeta function*, Proceedings of the International Congress of Mathematicians (Vancouver, B.C., 1974), Vol. 1, Canad. Math. Congr., Montreal, QC, 1975, pp. 379–381.
113. W. H. H. de Muinck Keizer, *On zonal Stiefel harmonics with applications in discrete geometry*, Ph.D. thesis, Delft University of Technology, 2025.
114. O. R. Musin, *The kissing number in four dimensions*, Ann. of Math. (2) **168** (2008), no. 1, 1–32. doi:10.4007/annals.2008.168.1
115. O. R. Musin, *Multivariate positive definite functions on spheres*, arXiv:math/0701083 (2008).
116. J. Nie, K. Ranestad, and B. Sturmfels, *The algebraic degree of semidefinite programming*, Math. Program. **122** (2010), no. 2, 379–405. doi:10.1007/s10107-008-0253-6
117. A. W. Nordstrom and J. P. Robinson, *An optimum nonlinear code*, Information and Control **11** (1967), no. 5, 613–616. doi:10.1016/S0019-9958(67)90835-2
118. D. Papp and S. Yıldız, *Sum-of-squares optimization without semidefinite programming*, SIAM J. Optim. **29** (2019), no. 1, 822–851. doi:10.1137/17M1160124
119. P. A. Parrilo and B. Sturmfels, *Minimizing polynomial functions*, Algorithmic and Quantitative Real Algebraic Geometry (Piscataway, NJ, 2001), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 60, Amer. Math. Soc., Providence, RI, 2003, `https://arxiv.org/abs/math/0103170`, pp. 83–99.

120. G. Pataki, *On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues*, Math. Oper. Res. **23** (1998), no. 2, 339–358. doi:10.1287/moor.23.2.339
121. F. N. Permenter, *Reduction methods in semidefinite and conic optimization*, Phd thesis, Massachusetts Institute of Technology, 2017.
122. K. Phelps, *Enumeration of Kerdock codes of length 64*, Des. Codes Cryptogr. **77** (2015), no. 2-3, 357–363. doi:10.1007/s10623-015-0053-y
123. M. Putinar, *Positive Polynomials on Compact Semi-algebraic Sets*, Indiana University Mathematics Journal **42** (1993), no. 3, 969–984.
124. C. Riener, J. Rolfes, and F. Vallentin, *A semidefinite programming hierarchy for covering problems in discrete geometry*, arXiv:2312.11267, December 2023.
125. J. H. Rolfes, *Convex Optimization Techniques for Geometric Covering Problems*, Ph.D. thesis, University of Cologne, 2019, p. 115.
126. Z. Rudnick and P. Sarnak, *The n-level correlations of zeros of the zeta function*, C. R. Acad. Sci. Paris Sér. I Math. **319** (1994), no. 10, 1027–1032.
127. Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81** (1996), no. 2, 269–322. doi:10.1215/S0012-7094-96-08115-6
128. I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108. doi:10.1215/S0012-7094-42-00908-6
129. K. Schütte and B. L. van der Waerden, *Das Problem der dreizehn Kugeln*, Math. Ann. **125** (1953), 325–334. doi:10.1007/BF01343127
130. J.-P. Serre, *Compact groups*, Linear Representations of Finite Groups (J.-P. Serre, ed.), Graduate Texts in Mathematics, Springer, New York, NY, 1977, pp. 32–34. doi:10.1007/978-1-4684-9458-7_4
131. J.-P. Serre, *Linear representations of finite groups*, corr. 5th print ed., Graduate Texts in Mathematics, no. 42, Springer-Verlag, New York, 1996.
132. D. Simmons-Duffin, *A Semidefinite Program Solver for the Conformal Bootstrap*, arXiv:1502.02033, February 2015.
133. N. J. A. Sloane, *Tables of sphere packings and spherical codes*, IEEE Trans. Inform. Theory **27** (1981), no. 3, 327–338. doi:10.1109/TIT.1981.1056351
134. S. L. Snover, *The uniqueness of the Nordstrom-Robinson and the Golay binary codes*, ProQuest LLC, Ann Arbor, MI, 1973. doi:10.25335/M56D5PM3R
135. K. Sono, *A note on simple zeros of primitive Dirichlet L-functions*, Bull. Aust. Math. Soc. **93** (2016), no. 1, 19–30. doi:10.1017/S0004972715000623
136. E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, vol. No. 30, Princeton University Press, Princeton, NJ, 1970.
137. K.-C. Toh, *A Note on the Calculation of Step-Lengths in Interior-Point Methods for Semidefinite Programming*, Comput. Optim. Appl. **21** (2002), no. 3, 301–310. doi:10.1023/A:1013777203597
138. K.-C. Toh, M. J. Todd, and R. H. Tütüncü, *On the Implementation and Usage of SDPT3 – A Matlab Software Package for Semidefinite-Quadratic-Linear Programming, Version 4.0*, Handbook on Semidefinite, Conic and Polynomial Optimization (M. F. Anjos and J. B. Lasserre, eds.), vol. 166, Springer US, Boston, MA, 2012, pp. 715–754. doi:10.1007/978-1-4614-0769-0_25
139. B. A. Venkov, *On a class of Euclidean polyhedra*, Vestnik Leningrad. Univ. Ser. Mat. Fiz. Him. **9** (1954), no. 2, 11–31.
140. M. Viazovska, *The sphere packing problem in dimension 8*, Ann. Math. **185** (2017), no. 3, 991–1015. doi:10.4007/annals.2017.185.3.7
141. D. V. Widder, *The Laplace Transform*, Princeton Mathematical Series, vol. vol. 6, Princeton University Press, Princeton, NJ, 1941.
142. S. Willard, *General topology*, Dover Publications, Inc., Mineola, NY, 1970.
143. M. Yamashita, K. Fujisawa, K. Nakata, M. Nakata, M. Fukuda, K. Kobayashi, and K. Goto, *A high-performance software package for semidefinite programs: SDPA 7*, Research Report B-460, Dept. of Mathematical and Computing Science, Tokyo Institute of Technology, Tokyo, Japan, January 2010.
144. V. A. Yudin, *The minimum of potential energy of a System of point charges*, Discrete Math. Appl. **3** (1993), no. 1, 75–82. doi:10.1515/dma.1993.3.1.75

Acknowledgements

As with nearly all things, finishing a PhD is not something you can do alone. First and foremost, I would like to thank my advisor David de Laat for introducing me to this wonderful research topic, and for his guidance, encouragement, and feedback.

Thanks to my collaborators Henry Cohn, Felipe Gonçalves, Willem de Muinck Keizer, David, Fernando de Oliveira, Andreas Spomer, Frank Vallentin, Marc Christian Zimmerman for the discussions and the work done together.

Thanks to everyone on the fourth floor for all the seminars, group outings, and interesting lunches. A special thanks to Bram Bekker, Fernando, and Willem, for the discussions related to discrete geometry.

I would also like to extend my thanks to the people in Cologne, for their hospitality during my visits there, and to all the organizers of the conferences and seminars I have been to.

Thanks to the committee members, Dion Gijswijt, Mark Veraar, Monique Laurent, Etienne de Klerk, and Philippe Moustrou, for taking the time to read this thesis.

And finally, thanks to my family for their support and love.

Nando Leijenhorst
Delft, April 2025

Curriculum Vitæ

Nando Matthias Leijenhorst

21-05-1998 Born in Delft, the Netherlands.

Education

2010–2016 VWO
Christelijk Lyceum Delft

2016–2019 Double BSc. Applied Mathematics & Applied Physics
Delft University of Technology
Thesis: Quantum error correction: Decoders for the toric code
Supervisors: Dr. M. Caspers
Dr. D. Elkouss

2019–2021 MSc. Applied Mathematics
Delft University of Technology
Thesis: A solver for clustered low-rank SDPs arising from multivariate polynomial (matrix) programs
Supervisor: Dr. D. de Laat

2021–2025 Phd. in Mathematics
Delft University of Technology
Thesis: On the computation of three and four-point bounds in discrete geometry and analytic number theory
Promotor: Prof. Dr. D. Gijswijt
Supervisor: Dr. D. de Laat

List of Publications

(1) F. Gonçalves, D. de Laat and N. Leijenhorst, *Multiplicity of nontrivial zeros of primitive L-functions via higher-level correlations*, Math. Comp. (2024), arXiv:2303.01095, doi:10.1090/mcom/4005

(2) N. Leijenhorst and D. de Laat, *Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput. **16** (2024), no. 3, 503-534, doi:10.1007/s12532-024-00264-w

(3) H. Cohn, D. de Laat and N. Leijenhorst, *Optimality of spherical codes via exact semidefinite programming bounds*, 2024, arXiv:2403.16874

(4) D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer, *Optimality and uniqueness of the D_4 root system*, 2024, arXiv:2404.18794

(5) D. de Laat, N. M. Leijenhorst and W. H. H. de Muinck Keizer, *Traceless projection of tensors with applications to hierarchies in discrete geometry*, In preparation

(6) N. Leijenhorst, A. Spomer, F. Vallentin and M. C. Zimmerman, *Semidefinite approaches to covering and polarization on the sphere*, In preparation