OPTIMALITY AND UNIQUENESS OF THE D_4 ROOT SYSTEM

DAVID DE LAAT, NANDO LEIJENHORST, AND WILLEM DE MUINCK KEIZER

Abstract. We prove that the D_4 root system (the set of vertices of the regular 24-cell) is the unique optimal kissing configuration in $\mathbb{R}^4$, and is an optimal spherical code. For this, we use semidefinite programming to compute an exact optimal solution to the second level of the Lasserre hierarchy. We also improve the upper bound for the kissing number problem in $\mathbb{R}^6$ to 77.

Contents

1. Introduction

A kissing configuration in dimension n is a collection of nonoverlapping, equal-size spheres in $\mathbb{R}^n$ that touch (or "kiss") a central sphere of the same size. We will assume the spheres have unit radius, and we identify a kissing configuration with the set C of contact points with the central sphere. Such a set C is a spherical code with minimal angle at least $\pi/3$. The kissing number $k(n)$ in dimension n is the maximum size of such a set C.

In dimension four, a kissing configuration is given by the D_4 root system. This root system may be constructed as the set of all 24 vectors in $\mathbb{R}^4$ with integer coordinates and length $\sqrt{2}$, referred to as the roots. In this paper we normalize the roots to have unit length. Viewed geometrically, the roots form the vertices of the 24-cell, which is one of the six regular polytopes in dimension four. The possible inner products between distinct roots are 0, $\pm 1/2$, and -1. Hence D_4 is a kissing configuration in dimension 4 of size 24; that is, $k(4) \geq 24$. In 2008, Musin showed $k(4) = 24$; the D_4 root system is an optimal kissing configuration [38].

In this paper, we show it is unique. More precisely, we show that the D_4 root system is the only optimal kissing configuration in dimension four up to isometry. This implies it satisfies the stronger geometric condition of being an optimal spherical

Date: April 29, 2024.
2020 *Mathematics Subject Classification.* 90C22, 52C17.

code: it minimizes

$$t_{\max}(C) = \max_{\substack{x,y\in C\\ x\neq y}} \langle x, y\rangle$$

over all sets C consisting of 24 points in the unit sphere $S^3 = \{x \in \mathbb{R}^4 \mid \langle x, x\rangle = 1\}$. This is in contrast with the result by Cohn, Conway, Elkies and Kumar [7] that the D_4 root system is not universally optimal, meaning that there exists an absolutely monotonic function f (a smooth function with all derivatives nonnegative on $[-1, 1]$) for which D_4 does not minimize

$$\sum_{\substack{x,y\in C\\ x\neq y}} f(\langle x, y\rangle)$$

over all $C \subseteq S^3$ of size 24. In their paper, they conjecture that no universally optimal spherical code of 24 points exists in S^3. The combination of the D_4 root system being the unique optimal spherical code, but not a universally optimal spherical code, proves this conjecture.

The kissing number problem has a rich history, going back to a discussion between Newton and Gregory in 1694 on the correct value of $k(3)$, which was resolved in 1953 by Schütte and Van der Waerden [49]. Currently, the value of $k(n)$ is known for n = 1, 2, 3, 4, 8, and 24. For background on the kissing number problem, we refer to [45].

In 1973, Delsarte introduced the linear programming bound, which can be used to bound the sizes of codes over finite alphabets [19]. Delsarte, Goethals, and Seidel adapted this to the sphere, so that it can be used to compute upper bounds on $k(n)$ [20]. Remarkably, this bound is sharp in dimensions 8 and 24, where by a sharp bound we mean that the optimal objective value is exactly equal to the kissing number, without having to take the integer part. The optimal objective value is 240 in dimension 8 and 196560 in dimension 24, which coincides with the sizes of the kissing configurations obtained by taking the shortest nonzero vectors in the E_8 root lattice and the Leech lattice Λ_{24} [43, 32]. This proves optimality of those configurations, and since the bound is sharp, complementary slackness can be used to prove uniqueness [3].

In dimension four, the Delsarte bound was used to show $k(4) \leq 25$, which was the first improvement over Coxeter's upper bound of 26 from 1964 [42, 13]. In [1] it was shown that the Delsarte bound cannot be used to prove $k(4) = 24$, and Musin's optimality proof for the D_4 root system uses a strengthening of the Delsarte bound. However, this strengthening does not lead to a sharp bound.

The Delsarte bound is called a two-point bound since it considers constraints between pairs of points on the sphere. Bachoc and Vallentin developed the three-point semidefinite programming bound for spherical codes, adapted from Schrijver's three-point bound for binary codes [2, 48]. The three-point bound recovers the optimality results in dimensions 3 and 4 and improves the best-known upper bound for the kissing number problem in many other dimensions. To compute the three-point bound it is first reduced to a finite-dimensional problem by truncating an inverse Fourier transform, and since its introduction in 2008, all improvements to upper bounds on $k(n)$ have come from increasing this truncation degree [36, 34, 31].

The numerical data (see [31, Table 6.1] for the newest results), however, suggests that the three-point bound for the kissing number problem is not sharp in any dimension $3 \leq n \leq 24$ and any truncation degree, except for the cases $n = 8$ and

$n = 24$ where the Delsarte bound is already sharp. There has been considerable work on k-point bound generalizations of the three-point bound, but this has not yet resulted in sharp, or even improved, bounds for the kissing number problem (or spherical code problems in general) [39, 28, 40, 16, 37, 4].

Over the last decades, the moment/SOS approach by Lasserre and Parrilo (see [29, 30, 44]) has become an important tool in mathematical optimization and theoretical computer science. Applying the Lasserre hierarchy to the independent set problem in a finite graph gives a converging hierarchy of increasingly large semidefinite programs giving successively stronger upper bounds on the independence number. We can think of the kissing number problem as the independent set problem in the graph on the unit sphere S^{n-1}, where two distinct vertices $x, y \in S^{n-1}$ are adjacent if $\langle x, y\rangle > 1/2$. In [18], De Laat and Vallentin generalized this hierarchy to infinite graphs such as these, giving a hierarchy of $2t$-point bounds, where t is the level of the hierarchy. In principle, this solves the kissing number problem in any dimension, since this hierarchy converges in finitely many steps. In practice, computing the levels of this hierarchy beyond the first level (which reduces to the Delsarte linear programming bound) is challenging.

In this paper, we compute the second level of this hierarchy for spherical code problems. We show the second level of the hierarchy is sharp for the kissing number problem in dimension four (the upper bound is exactly 24) by computing an exact optimal solution. We then use complementary slackness to extract a uniqueness proof for the D_4 root system from the optimal solution.

This is the first time the second level of the Lasserre hierarchy has been computed for a spherical code problem and the first improvement over the three-point bounds for spherical codes. Previously, the second level of the Lasserre hierarchy has been computed for two problems on infinite graphs. Firstly, it has been computed for energy minimization on the two-dimensional sphere [15]. The techniques used there, however, become too expensive when going to higher dimensional spheres or further truncation degree of the inverse Fourier transform, and computing a sharp bound for the kissing number problem in dimension four would be prohibitively expensive with the techniques from that paper.

In [17], De Laat, Machado, and De Muinck Keizer compute the second and third levels of the hierarchy for the equiangular lines problem with a fixed angle θ. Here the corresponding graph on S^{n-1} has an edge between distinct points x and y if $\langle x, y\rangle \neq \pm\cos\theta$. Although this is an infinite graph, the quotient space $\mathcal{I}_t/\mathrm{O}(n)$, where $\mathcal{I}_t$ is the set of independent sets of size at most t and $\mathrm{O}(n)$ is the orthogonal group, is finite. Here $\mathrm{O}(n)$ acts on $\mathcal{I}_t$ by $g\{x_1, \ldots, x_k\} = \{gx_1, \ldots, gx_k\}$. For the kissing number problem, the quotient space $\mathcal{I}_t/\mathrm{O}(n)$ is infinite. Because of this, computing the hierarchy for the kissing number problem is more involved, and in this paper, we extend the techniques from [17] to do this.

The t-th level of the hierarchy is the optimization problem

$$\begin{aligned} &\text{minimize} && K(\emptyset, \emptyset)\\ &\text{subject to} && K \in \mathcal{C}(\mathcal{I}_t \times \mathcal{I}_t)_{\succeq 0},\\ & && A_t K(Q) \leq -1_{\mathcal{I}_{=1}}(Q),\ Q \in \mathcal{I}_{2t} \setminus \{\emptyset\}. \end{aligned} \tag{1.1}$$

Here $\mathcal{C}(\mathcal{I}_t \times \mathcal{I}_t)_{\succeq 0}$ is the cone of continuous, positive kernels on $\mathcal{I}_t$, where $\mathcal{I}_t$ inherits its topology from S^{n-1} (see [18]), $1_{\mathcal{I}_{=1}}$ is the indicator function of the set $\mathcal{I}_{=1}$ of one-element subsets of S^{n-1}, and

$$A_tK(Q) = \sum_{\substack{J_1,J_2\in\mathcal{I}_t\\ J_1\cup J_2=Q}} K(J_1,J_2).$$

Any feasible solution K provides an upper bound on $k(n)$; see the start of the proof of Lemma 5.1 for the argument. Moreover, if K is feasible, then the kernel

$$(J_1,J_2)\mapsto \int_{\mathrm{O}(n)} K(\gamma J_1,\gamma J_2)\,d\gamma$$

is also feasible and has the same objective value, from which it follows we may restrict to $\mathrm{O}(n)$-invariant kernels.

To reduce this to a finite-dimensional problem, we express such an $\mathrm{O}(n)$-invariant kernel K in terms of its inverse Fourier transform and truncate the series. For each $\lambda\in\mathbb{Z}^t$ with $\lambda_1\geq\ldots\geq\lambda_t\geq 0$, we define a unitary representation $\pi\colon \mathrm{O}(n)\to V$ and denote the space of continuous, $\mathrm{O}(n)$-equivariant maps from $\mathcal{I}_t$ to V by $\mathrm{Hom}_{\mathrm{O}(n)}(\mathcal{I}_t,V)$. We refer to $|\lambda|=\sum_i\lambda_i$ as the degree of π. We will construct a family $\{\psi_{\lambda,\ell}\}$ of elements in this space and define the matrix $Z_\lambda(J_1,J_2)$ by

$$Z_\lambda(J_1,J_2)_{\ell_1,\ell_2}=\langle\psi_{\lambda,\ell_1}(J_1),\psi_{\lambda,\ell_2}(J_2)\rangle,$$

where the inner product on V is used. For this, we use the construction by Gross and Kunze [25] of the spaces of invariants $V^{\mathrm{O}(n-t)}$ induced by the representations of $\mathrm{GL}(t)$.

For each λ, let $\widehat{K}_\lambda$ be a positive semidefinite matrix of the same size as Z_λ with only finitely many nonzero entries. Then the kernel $K\colon\mathcal{I}_t\times\mathcal{I}_t\to\mathbb{R}$ defined by

$$K(J_1,J_2)=\sum_{|\lambda|\leq d}\langle\widehat{K}_\lambda,Z_\lambda(J_1,J_2)\rangle,$$

is continuous, positive, and $\mathrm{O}(n)$-invariant. With the right choice of representations and families of equivariant functions, these approximate all continuous, positive, $\mathrm{O}(n)$-invariant kernels. This last statement will not be discussed in this paper, since it is not necessary for the main result.

We have two main technical contributions. In [17], the zonal matrices Z_λ are constructed for the case where $\mathcal{I}_t/\mathrm{O}(n)$ is finite and where there are only finitely many pointwise constraints. In Section 2, we give a construction for infinitely many orbits, and we give a rescaling so that the entries of $Z_\lambda(J_1,J_2)$ become polynomials in the inner products between the vectors in $J_1\cup J_2$. This allows us to reduce (1.1) to a finite-dimensional problem by truncating the inverse Fourier transform, and to write the constraints using sums-of-squares characterizations, which means we can use semidefinite programming to compute bounds.

Our second technical contribution concerns the computation of the zonal matrices for $t=2$. In our approach of generating the zonal matrices via representations of $\mathrm{O}(n)$ by induced by $\mathrm{GL}(t)$, we identify additional symmetries under certain actions of $\mathrm{O}(t)$ and $\mathrm{O}(n-t)$, and we use this to significantly reduce the amount of computations that need to be performed; see Section 3. To obtain a sharp bound for the kissing number problem in $\mathbb{R}^4$ we need the zonal matrices Z_λ with $|\lambda|\leq 14$, and for this these reductions are essential.

Cohn and Elkies [10] gave a noncompact adaptation of the Delsarte linear programming bound, and conjectured it gives the optimal sphere packing density in dimensions 8 and 24. Note that for noncompact problems, such as the sphere packing problem, one needs a sharp bound to prove optimality. In [51], Viazovska proved the groundbreaking result that the E_8 root lattice gives an optimal sphere packing in $\mathbb{R}^8$ by constructing an optimal solution to the Cohn-Elkies bound, after which optimality of the Leech lattice Λ_{24} was shown similarly in [11]. Currently, the sphere packing problem has been solved in dimensions 1, 2, 3, 8, and 24, where the proof for the three-dimensional case used a completely different approach [26].

It is conjectured that the D_4 lattice gives the optimal sphere packing in dimension four, where optimality among lattice packings has been known since 1873 [27]. A numerically sharp three-point bound for the lattice sphere packing problem in $\mathbb{R}^4$ has recently been computed in [9], but a (numerically) sharp bound for the general sphere packing problem is not known in dimension four. Since we show the second level of the Lasserre hierarchy is sharp for the kissing number problem in dimension four (just as the Delsarte bound is sharp in dimension 8 and 24), one might expect a noncompact adaptation (see also [12]) might be sharp for the sphere packing problem in dimension four (as is the Cohn-Elkies bound in dimensions 8 and 24). We therefore believe this might be a viable approach to solving the sphere packing problem in dimension four.

In this paper, we focus on the four-dimensional case. The reason for this is that computing the zonal matrices and solving the semidefinite programs is computationally expensive, and we have only performed computations for $|\lambda| \leq 16$. In the same way as for the three-point bounds, this is the truncation degree around which the bounds start to improve on the Delsarte bound. For the four-dimensional case of the kissing number problem, this results in a sharp bound, but it seems that in most other dimensions the degree is not yet high enough to get improved bounds. For the six-dimensional case, we report a small improvement in the upper bound from 78 to 77, which is the first improvement since the introduction of the three-point bound.

The paper is organized as follows. In Section 2, we construct a system of equivariant functions for which the corresponding zonal matrices consist of polynomials in the inner products. In Section 3, we show how these zonal matrices can be computed efficiently. In Section 4 we discuss the semidefinite programming formulation, and in Section 5 we discuss the applications.

2. Equivariant functions and zonal matrices

2.1. 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 for instance in [22]. 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 $\lambda_t \geq 0$.

Since we consider the second level of the Lasserre hierarchy, we will 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 standard representation with basis e_1, e_2, the signature $\lambda = (\lambda_1, \lambda_2)$ satisfies $\lambda_1 \geq \lambda_2 \geq 0$, and $m = \lambda_1 - \lambda_2$. 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

$$\begin{aligned}(2.1)\quad & \langle w_{k_1}, \rho(A) w_{k_2}\rangle \\ &= \det(A)^{\lambda_2} \sum_{l=0}^{m-k_1} \tbinom{m-k_2}{l}\tbinom{k_2}{m-k_1-l} A_{11}^{l} A_{21}^{m-k_2-l} A_{12}^{m-k_1-l} A_{22}^{k_2-(m-k_1-l)}.\end{aligned}$$

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

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

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

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

We have

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

and hence $d\rho(X)_{k_1 k_2} = -(m-k_2)\delta_{k_1, k_2+1} + k_2 \delta_{k_1, k_2-1}$. Such a formula may be obtained by considering a curve $c(t)$ with $c(0) = I$ and $c'(0) = X$ and considering the derivative of $\rho(c(t)) w_{k_2}$ evaluated at 0. For more background, we again refer to [22].

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

Following Gross and Kunze [25], we now define certain representations of $\mathrm{O}(n)$ 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

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

and the $n \times t$ matrix

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

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

$$f_w(\gamma) = \rho(\omega \gamma \epsilon) w. \tag{2.2}$$

Define the vector space of right translates of such functions by

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

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

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

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

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

On V, we define the inner product

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

By standard properties of the Haar measure, this makes V a unitary representation of $\mathrm{O}(n)$. It may be shown that with the inner product chosen in Section 2.1, the numbers

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

are real; see [17, Section 3].

For the main result of this paper, it is only required that V is a representation of the orthogonal group and $\Psi(W) \subseteq V^{\mathrm{O}(n-t)}$. However, it follows from the results in [25] that the above description is complete in the following sense. For $n > 2t$, the representations of $\mathrm{O}(n)$ defined above are irreducible and we have equality $\Psi(W) = V^{\mathrm{O}(n-t)}$. Moreover, all irreducible representations of $\mathrm{O}(n)$ with nontrivial invariants under $\mathrm{O}(n-t)$ are of this form for a unique λ. For $n = 2t$, a complete characterization of the irreducible representations and invariants is given in [25, Section 8], and using this it may be shown that the description of kernels in our approach is actually also complete in the case $n = 2t$. We defer the exact statement and verification to the upcoming PhD thesis of the third-named author.

2.3. **Equivariant functions.** In this section, we define a family of $\mathrm{O}(n)$-equivariant functions from $\mathcal{I}_t$ to the representation V as constructed in Section 2.2. The definition of these functions depends on the choice of representatives of the orbits of $\mathcal{I}_t$ under the action of $\mathrm{O}(n)$. Let

$$\begin{aligned} p_j(\{x,y\}) &= \langle x, y\rangle^j, \\ q_1(\{x,y\}) &= \sqrt{2(1+\langle x,y\rangle)}, \\ q_2(\{x,y\}) &= \sqrt{2(1-\langle x,y\rangle)}. \end{aligned}$$

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

$$\left\{\left(\frac{q_1(J)}{2}, \frac{q_2(J)}{2}, 0, \ldots, 0\right), \left(\frac{q_1(J)}{2}, -\frac{q_2(J)}{2}, 0, \ldots, 0\right)\right\}. \tag{2.3}$$

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 $s\colon \mathcal{I}_t \to \mathrm{O}(n)$ is a function such that $s(J)R = J$, where R is the orbit representative of the orbit $\mathrm{O}(n)J$. 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 $\mathcal{I}_{=i}/\mathrm{O}(n) \times \mathrm{O}(n)/\mathrm{O}(n-i)$. 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 $\mathcal{I}_{=2}/\mathrm{O}(n)$ and $\pi(s(J))\Psi(w_k)$ as a function on the factor $\mathrm{O}(n)/\mathrm{O}(n-2)$. 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 $\mathcal{I}_{=2}/\mathrm{O}(n)$, and that we obtain polynomial expressions.

Lemma 2.1. *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 $\mathrm{O}(n)$ 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 $\mathrm{O}(n)J$. For $g \in \mathrm{O}(n)$, we have $s(gJ)R = gJ$ and $gs(J)R = gJ$, so $s(gJ) = gs(J)h$ for some h in the stabilizer subgroup $\mathrm{Stab}_{\mathrm{O}(n)}(R)$. 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 $\mathrm{O}(n)$ with respect to R is $\mathrm{O}(n-1)$. By formula (2.1), 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 $\mathrm{O}(n)$ with respect to R is $S_2 \times \mathrm{O}(n-2)$, 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 2.2), we have $\pi(h)\Psi(w_k) = \Psi(w_k)$ for $h \in \mathrm{O}(n-2)$. 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 (2.1) 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 $\mathrm{O}(n-1)$ 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 (2.1), $\rho(\omega\gamma h\epsilon)w_k$ only depends on the second column of $\omega\gamma h\epsilon$, which is equal to the second column of $\omega\gamma\epsilon$, so

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

2.4. **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 $\mathrm{O}(n)$ 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.

Lemma 2.2. *Let (λ, i, j, k) be admissible. For fixed $w \in W$, the expression*

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

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

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 (2.1) 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

$$\psi_{\lambda,(i,j,k)}(J) = \langle x_1, x_2\rangle^j \rho(\omega\gamma\big[x_1 + x_2 \quad x_1 - x_2\big])w_k. \tag{2.4}$$

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

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

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

$$\rho\big(\omega\gamma\big[u \quad v\big]\big)w_k = \rho\big(\omega\gamma\big[u \quad -v\big]\big)w_k$$

for all orthonormal u and v. Since

$$\rho\Big(\omega\gamma\Big[\tfrac{x_1+x_2}{q_1(J)} \quad \tfrac{x_1-x_2}{q_2(J)}\Big]\Big) = \rho(\omega\gamma\big[x_1 + x_2 \quad x_1 - x_2\big])\rho\left(\begin{bmatrix}1/q_1(J) & 0\\ 0 & 1/q_2(J)\end{bmatrix}\right)$$

it follows that identity (2.4) holds whenever $\langle x_1, x_2\rangle \neq \pm 1$.

We will show (2.4) 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

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

A similar argument can be used to show (2.4) 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 2.3. *Fix i_1 and i_2 and let $J_1 = \{x_1, \ldots, x_{i_1}\}$ and $J_2 = \{y_1, \ldots, y_{i_2}\}$. 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 $x_1, \ldots, x_{i_1}, y_1, \ldots, y_{i_2}$.

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

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

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

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

By Lemma 2.2, this is a polynomial in the entries of the vectors $x_1, \ldots, x_{i_1}, y_1, \ldots, y_{i_2}$.

By Lemma 2.1, 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 $g \in \mathrm{O}(n)$. In other words, this is an $\mathrm{O}(n)$-invariant polynomial in the vectors $x_1, \ldots, x_{i_1}, y_1, \ldots, y_{i_2}$. By invariant theory (see, e.g., [22, §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. Efficient computation of the zonal matrices

In this section, we explain how we compute the zonal matrices from Section 2. 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 [17], 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

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

where we define

$$
P(S) = \int_{\mathrm{O}(n)} \langle \rho(\omega\gamma\epsilon)w_{k_1}, \rho(\omega\gamma S\epsilon)w_{k_2}\rangle \, d\gamma. \tag{3.1}
$$

Here $P(S)$ is a polynomial in the entries of the $n \times n$ matrix S, and the main focus of this section is the efficient computation of this polynomial.

3.1. **Additional symmetries.** To compute $P(S)$ using the matrix entries of ρ, one could directly use the expression

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

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

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

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

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:

$$
J_{l_1,l_2,[\alpha]} = \int_{\mathrm{O}(n)} \det(\bar{A})^{\lambda_2} \rho(A)^*_{k_1,l_1} \rho(B)_{l_2,k_2} \prod_{i=1}^{\lambda_2} B_{\alpha_{i1},1} B_{\alpha_{i2},2} \, d\gamma. \tag{3.2}
$$

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

$$
J_{l_1,l_1,[\sigma]} = c_{l_1,k_2,[\sigma]} J_{0,0,[e]} \tag{3.3}
$$

for all $0 \leq l_1 \leq m$ 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 $P(S)$ may then be computed as

$$P(S) = \sum_{l_1,\sigma} (-1)^{s(\sigma)} J_{l_1,l_1,[\sigma]}.$$

where the sum is over $0 \leq l_1 \leq m$ 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 (3.3) we can write this as

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

In summary, for fixed λ, k_1 and k_2, we need to compute only one integral of the form (3.2) using this approach.

We now show how to compute these coefficients. For $g \in \mathrm{O}(t)$, 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 $\mathrm{O}(n)$. 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

$$J_{l_1,l_2,[\alpha]} = \sum_{l_3,l_4,[\beta]} \det(\bar{g})^{\lambda_2} \rho(\bar{g})_{l_1,l_3} J_{l_3,l_4,[\beta]} \rho(g)_{l_2,l_4} \sum_{\zeta \in [\beta]} \prod_{i=1}^{\lambda_2} g_{\alpha_{i1},\zeta_{i1}} g_{\alpha_{i2},\zeta_{i2}}, \tag{3.4}$$

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 $\mathrm{O}(2)$ 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

$$\langle e_{l_1,l_2,[\alpha]}, \Phi(g) e_{l_3,l_4,[\beta]} \rangle = \det(g)^{\lambda_2} \rho(g)_{l_1,l_3} \rho(g)_{l_2,l_4} \sum_{\zeta\in[\beta]} \prod_{i=1}^{\lambda_2} g_{\alpha_{i1},\zeta_{i1}} g_{\alpha_{i2},\zeta_{i2}}. \tag{3.5}$$

Using (3.4) and (3.5) 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 $g \in \mathrm{O}(2)$. 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 (3.5) and take the derivative. Using the product rule, one obtains

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

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 2.1. 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 3.1. *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 3.2. *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 $J_{l_1,l_2,[\alpha]} = 0$. $\square$

We thus have the system $d\Phi(X)J = 0$ and certain components of J vanish due to Lemmas 3.1 and 3.2. 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

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

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

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

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 (3.3) holds.

3.2. **Real parts.** As shown in Section 3.1, to compute the zonal matrices we need to compute the quantity

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

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

$$(3.6)\quad 2\int_{\mathrm{O}(n)} \mathcal{R}\big(\det(\overline{\omega}\gamma\epsilon)^{\lambda_2} \rho(\omega\gamma\epsilon)^*_{k_1,0}\big) \mathcal{R}\big(\rho(\omega\gamma S\epsilon)_{0,k_2} ((\omega\gamma S\epsilon)_{1,1}(\omega\gamma S\epsilon)_{2,2})^{\lambda_2}\big)\, d\gamma,$$

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 adapt 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

$$(3.7)\qquad \int_{\mathrm{O}(n)} (\overline{\omega}\gamma\epsilon)^a (\omega\gamma S\epsilon)^b\, d\gamma = 2\int_{\mathrm{O}(n)} \mathcal{R}\left((\overline{\omega}\gamma\epsilon)^a\right) \mathcal{R}\left((\omega\gamma S\epsilon)^b\right)\, d\gamma$$

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

$$(3.8)\qquad \begin{aligned} \int_{\mathrm{O}(n)} (R_1^+\gamma\epsilon)^a (R_2^+\gamma S\epsilon)^b\, d\gamma &= \sum_{|u|=|v|=|\lambda|} I_{u,v} R_1^u R_2^v \\ &= \sum_{|s|=2|\lambda|} \sum_{\substack{u+v=s \\ |u|=|v|=|\lambda|}} I_{u,v} R_{11}^{u_1} R_{12}^{u_2} R_{21}^{v_1} R_{22}^{v_2}, \end{aligned}$$

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 (3.7), which is a real number by [17, 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 (3.7) is given by

$$(3.9)\qquad \sum_{|s|=|\lambda|} (-1)^{s_2} \sum_{\substack{u+v=2s \\ |u|=|v|=|\lambda|}} (-1)^{u_2} I_{u,v}.$$

A similar reasoning shows that the right-hand side of (3.7) 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 (3.8) to obtain the polynomial

$$\int_{\mathrm{O}(n)} (R_1^+ \gamma \epsilon)^a (R_1^+ \gamma S \epsilon)^b \, d\gamma = \sum_{|u|=|v|=|\lambda|} I_{u,v} R_1^u R_1^v = \sum_{|s|=2|\lambda|} \sum_{\substack{u+v=s \\ |u|=|v|=|\lambda|}} I_{u,v} R_1^s. \tag{3.10}$$

Similarly as before, we may substitute γ with $(g \oplus g \oplus I_{n-2t})\,\gamma$, and this leaves the polynomial (3.10) invariant by the invariance property of the Haar measure of $\mathrm{O}(n)$. We have $R_1^+ (g \oplus g \oplus I_{n-2t})\,\gamma = g R_1^+ \gamma$. Hence polynomial (3.10) 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

$$\left(R_{11}^2 + R_{12}^2\right)^{|\lambda|}. \tag{3.11}$$

Hence the s which occur in the sum in (3.10) must have even entries, and the polynomial (3.10) 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 (3.11), 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 (3.9), we now see that we may write the left-hand side of (3.7) 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 (3.7), which is what we wanted to show.

3.3. **Computational aspects.** In this section, we describe how to efficiently compute

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

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

Recall

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

Since (λ, i_1, j_1, k_1) is admissible, we have

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

for all $h \in \mathrm{O}(n - i_1)$, where as before we view h as a matrix in $\mathrm{O}(n)$ fixing the first i_1 coordinates. By the invariance property of the Haar measure, we have $P(hS) = P(S)$ for all $h \in \mathrm{O}(n - i_1)$. Write

$$S = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix},$$

where S_1 is of size $i_1 \times n$ and S_2 of size $(n - i_1) \times n$. Using invariant theory (see, e.g., [22, §F.1]) we see that P must be a polynomial in the entries of S_1 and the inner products between the columns of S_2. Furthermore, it follows from formula (2.1) for the matrix coefficients of ρ that only the entries from the first i_2 columns of S_1 are used, since the tuple (λ, i_2, j_2, k_2) is admissible.

Consider the ideal $\mathcal{J}$ generated by the entries of $S^{\mathsf{T}}S - I$ and the monomials $S_{i,j}$ for $i > i_1 + j$. We will describe a procedure to obtain from P a polynomial p in the entries of the top-left $i_1 \times i_2$ block of S such that the difference lies in $\mathcal{J}$. Since $S_{i,j} \in \mathcal{J}$ for $i > i_1 + j$, we may remove terms with such $S_{i,j}$ as the first step of the procedure. By the invariance property of P, we then obtain a polynomial of the form

$$\sum_{\alpha,a,b,c} C_{\alpha,a,b,c}(S_1)^\alpha (S_{3,1})^{2a}(S_{3,1}S_{3,2})^b(S_{3,2}^2 + S_{4,2}^2)^c \tag{3.12}$$

for some C, α, a, b, and c and the difference with P lies in $\mathcal{J}$. In (3.12) we replace every occurrence of $(S_{4,2})^2$ with $1-(S_{2,2})^2-(S_{1,2})^2-(S_{3,2})^2$, then every occurrence of $S_{3,1}S_{3,2}$ with $-S_{2,1}S_{2,2} - S_{1,1}S_{1,2}$, and finally every occurrence of $(S_{3,1})^2$ with $1 - (S_{2,1})^2 - (S_{1,1})^2$. Since all steps of the procedure may be performed by adding elements of $\mathcal{J}$, this procedure ends with a polynomial p in the top-left $i_1 \times i_2$ entries of S such that $P - p \in \mathcal{J}$.

Since $s(J_1)^{\mathsf{T}}s(J_2)$ is orthogonal, there exists an element $h \in \mathrm{O}(n - i_1)$ such that $(hS)_{i,j} = 0$ for $i > i_1 + j$. Then

$$P(s(J_1)^{\mathsf{T}}s(J_2)) = P(hs(J_1)^{\mathsf{T}}s(J_2)) = p(hs(J_1)^{\mathsf{T}}s(J_2)) = p(s(J_1)^{\mathsf{T}}s(J_2)),$$

where the second equality holds because $hs(J_1)^{\mathsf{T}}s(J_2)$ lies in the vanishing locus of $\mathcal{J}$. The expressions in the top-left $i_1 \times i_2$ block of $s(J_1)^{\mathsf{T}}s(J_2)$ involve fractions with denominators $q_1(J)$ and $q_2(J)$. We now describe how we avoid working with intermediate symbolic expressions.

From formula (2.1) for the matrix coefficients of the representation ρ of $\mathrm{GL}(2)$, it follows that every monomial in the expansion of (3.1) 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 from P, 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 $p(S)$, the number of variables from column l, where $l = 1, 2$, is at most $c_l(k_2)$, and differs from this by an even number.

Swapping k_1 and k_2, transposing S, and performing the same procedure as above gives a polynomial $\tilde{p}(S)$ in the top-left $i_1 \times i_2$ block of S such that P and $\tilde{p}$ are the same as functions on $\mathrm{O}(n)$. Now, for each monomial in $\tilde{p}(S)$ the number of variables from row l, with $l = 1, 2$, is at most $c_l(k_1)$, and differs from this by an even number. When we think of the polynomials $p(S)$ and $\tilde{p}(S)$ as functions in the top-left $i_1 \times i_2$ coordinates, they agree on an open set, and hence the polynomials are identical. This shows that in each monomial in $p(S)$ the number of variables from row l, with $l = 1, 2$, is also at most $c_l(k_1)$, and differs from this by an even number.

As discussed in the proof of Lemma 2.2, the (l_1, l_1) entry, with $1 \leq l_1, l_1 \leq 2$, of $s(J_1)^{\mathsf{T}} s(J_2)$ has denominator $q_{l_1}(J_{l_1}) q_{l_1}(J_{l_1})$. To obtain the zonal matrix entry, we may replace each monomial S^a in $p(S)$ 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 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 3.1 and 3.2, the integrand of $P(S)$ may be replaced by the product of

$$\mathcal{R}(\det(\overline{\omega}\gamma\epsilon)\rho(\overline{\omega}\gamma\epsilon)^*_{k_1,0}) \tag{3.13}$$

and

$$\mathcal{R}(\rho(\omega\gamma S\epsilon)_{0,k_2}((\omega\gamma S\epsilon)_{1,1}(\omega\gamma S\epsilon)_{2,2})^{\lambda_2}). \tag{3.14}$$

The integration over $\mathrm{O}(n)$ and the substitution procedure described above may be swapped. We first compute (3.14) explicitly as a polynomial in the variables $S_{i,j}$ with $i \leq i_1 + j$ and $j \leq i_2$ and the top-left $2t \times 2t$ block of γ. We then perform the above substitution procedure. Since by the above, we know that after integration over $\mathrm{O}(n)$ 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 $2t \times 2t$ block of γ.

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

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

This means that we do not have to work out the product of the whole polynomial (3.14) with (3.13). 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 [24]. This enables us to explicitly compute $p(S)$, from which we obtain the zonal matrix entry as explained above.

4. Semidefinite programming formulation

Let $d_1 \leq d_2 \leq \delta$ be positive integers with δ even. In our application to the D_4 root system, we use $d_1 = 14$ and $d_2 = \delta = 16$.

In the semidefinite program, we optimize over positive semidefinite matrices $\widehat{K}_\lambda$. Here the rows and columns are indexed by tuples (i, j, k) for which (λ, i, j, k) is admissible (see Section 2.3) and $|\lambda| + 2j \leq d_2$, and we similarly restrict the rows and columns of Z_λ. Let

$$K(J_1, J_2) = \sum_{|\lambda| \leq d_1} \langle \widehat{K}_\lambda, Z_\lambda(J_1, J_2) \rangle.$$

It follows from Proposition 2.3 that $A_2K(Q)$ is a polynomial in the inner products between the vectors in Q. Using Section 3, we can find polynomials $p_1, \dots, p_4$ in 0, 1, 3, and 6 variables, such that

$$\begin{aligned} p_1 &= A_2K(\{x_1\}), \\ p_2(\langle x_1, x_2 \rangle) &= A_2K(\{x_1, x_2\}), \\ p_3(\langle x_1, x_2 \rangle, \langle x_1, x_3 \rangle, \langle x_2, x_3 \rangle) &= A_2K(\{x_1, x_2, x_3\}), \\ p_4(\langle x_1, x_2 \rangle, \langle x_1, x_3 \rangle, \dots, \langle x_3, x_4 \rangle) &= A_2K(\{x_1, \dots, x_4\}). \end{aligned}$$

Here p_3 is S_3-invariant and p_4 is S_4-invariant, where S_3 acts by permuting variables and the action of S_4 is such that

$$\begin{aligned} &p_4(\langle x_{\sigma(1)}, x_{\sigma(2)} \rangle, \langle x_{\sigma(1)}, x_{\sigma(3)} \rangle, \dots, \langle x_{\sigma(3)}, x_{\sigma(4)} \rangle) \\ &\quad = p_4(\langle x_1, x_2 \rangle, \langle x_1, x_3 \rangle, \dots, \langle x_3, x_4 \rangle). \end{aligned}$$

for all $\sigma \in S_4$. Note that the polynomials $p_1, \dots, p_4$ are of degree at most d_2, and their coefficients depend linearly on the entries of the matrices $\widehat{K}_\lambda$.

The Fourier truncated version of (1.1) can be formulated as the following semidefinite program with polynomial inequality constraints:

$$\begin{array}{lll} \text{minimize} & (\widehat{K}_0)_{(0,0,0),(0,0,0)} & \\ \text{subject to} & \widehat{K}_\lambda \succeq 0, & |\lambda| \leq d_1, \\ & p_1 \leq -1, & \\ & p_2(u) \leq 0, & u \in [-1, \cos\theta], \\ & p_3(u_1, u_2, u_3) \leq 0, & (u_1, u_2, u_3) \in \Delta_3, \\ & p_4(u_1, \dots, u_6) \leq 0, & (u_1, \dots, u_6) \in \Delta_4. \end{array} \tag{4.1}$$

Let

$$G_3 = \begin{pmatrix} 1 & u_1 & u_2 \\ u_1 & 1 & u_3 \\ u_2 & u_3 & 1 \end{pmatrix} \quad \text{and} \quad G_4 = \begin{pmatrix} 1 & u_1 & u_2 & u_3 \\ u_1 & 1 & u_4 & u_5 \\ u_2 & u_4 & 1 & u_6 \\ u_3 & u_5 & u_6 & 1 \end{pmatrix}.$$

The semialgebraic set Δ_i consists of all $u \in \mathbb{R}^{\binom{i}{2}}$ with $(u_j + 1)(\cos\theta - u_j) \geq 0$ for all $1 \leq j \leq \binom{i}{2}$ and for which the determinants of all principal submatrices of size at least 3 of G_i are nonnegative.

We can now use sum-of-squares polynomials to relax this further to a semidefinite program. For the two-point constraint, for instance, we can use Lukács result (see, e.g., [46]) to replace the condition $p_2(u) \leq 0$ for $u \in [-1, \cos\theta]$ by

$$p_2(u) + s_0(u) + (u + 1)(\cos\theta - u)s_1(u) \equiv 0, \tag{4.2}$$

where s_0 and s_1 are sum-of-squares polynomials of degree δ and $\delta - 2$, respectively. Let $m_l(u)$ be a vector whose entries form a basis for the polynomials up to degree l. We can write

$$s_k(u) = \langle m_{\delta/2-k}(u) m_{\delta/2-k}(u)^{\mathsf{T}}, M_k \rangle,$$

where M_k is a positive semidefinite matrix. In this way, we can replace the two-point polynomial inequality constraint by two positive semidefinite matrices and several linear constraints that enforce the polynomial identity (4.2).

We can do something similar for the three-point and four-point constraints. Suppose $\Delta_i = \{u : g_k(u) \geq 0,\ k = 1, \ldots, l\}$, then we relax the polynomial inequality constraint to the identity

$$p_i(u) + \sum_{k=0}^{l} r_k(u) g_k(u) \equiv 0,$$

where we set $g_0(u) = 1$ and $r_k(u)$ is a sum-of-squares polynomial of degree at most $\delta - \deg(g_k)$. By Putinar's theorem [47] (see also [35, Chapter 13]), this relaxation converges to the original polynomial constraint when $\delta \to \infty$.

In the resulting semidefinite program, the positive semidefinite matrix variables for the four-point constraint will be far larger than any other matrix in the program. For this reason, exploiting the symmetries in the polynomials is essential.

If a semialgebraic set is invariant under the action of a group, then there exists a description in terms of invariant polynomials [5]. Let

$$\{q_1, \ldots, q_l\}$$

be an orbit of the polynomials describing Δ_i under the action of S_i. Then we can replace the polynomials in this orbit by the S_i-invariant polynomials

$$\sum_{\substack{B \subseteq \{1,\ldots,l\} \\ |B| = b}} \prod_{k \in B} q_k$$

for $b = 1, \ldots, l$; see, e.g., [34, 31] for the proof.

We may now assume the sum-of-squares polynomials for the three-point and four-point constraints are also invariant under the given action of the symmetric groups S_3 and S_4. This means that instead of using one large positive semidefinite matrix, we can use several smaller positive semidefinite matrices to model each sum-of-squares polynomial [23]. To do this explicitly we follow [31, Section 4].

This symmetry reduction involves the irreducible, unitary representations of S_3 and S_4. Although the irreducible, unitary representations we use involve irrational numbers, the irrationalities cancel in the final formulation, and the semidefinite program we obtain is rational whenever $\cos\theta$ is rational.

5. Applications to spherical codes

The results in this section depend on the proofs and verification code available at [14], including instructions on how to run the verification script. There we also make available the code we used for generating the proofs.

5.1. Optimality and uniqueness of 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 (4.1) with $n = 4$ and $\theta = \pi/3$. 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 [31] is essential. This solver supports arbitrary precision floating-point arithmetic and exploits the low-rank structure of the constraint matrices arising from enforcing the polynomial constraints through sampling at a unisolvent set [33]. 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. Since the dimension of the optimal face is lower than the dimension of the space given by the affine constraints, simply projecting the numerically optimal solution into the affine space does not work: the resulting matrix variables will generally not be positive semidefinite. Instead, we use the recently developed rounding heuristic from [8]. This method gives a major speedup over the rounding heuristic developed in [21], which is crucial for the size of semidefinite program we consider here.

Although a semidefinite program defined over the rationals does not necessarily admit a rational optimal solution (see, e.g., [41]), 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 $K(\emptyset, \emptyset) = 24$.

To verify that the exact solution is indeed feasible we check that the matrices are positive semidefinite by computing the Cholesky factorizations in rigorous ball arithmetic, and we check that the affine constraints hold in rational arithmetic. As part of the verification procedure, the zonal matrices Z_λ need to be constructed, which takes less than two days on a modern computer. The remainder of the verification procedure takes about an hour.

We obtain optimality as a spherical code as a consequence of the univariate polynomial p_2 corresponding to $A_2K|_{I_{=2}}$ having finitely many roots. A similar argument could have been used if a polynomial truncation of the Bachoc-Vallentin three-point bound had been sharp. In that case, however, it would have been immediate that the two-point constraint has finitely many roots (see [8, Section 3]), but for a truncation of the second level of the Lasserre hierarchy it is unclear whether p_2 can be identically zero when the bound is sharp.

Using Sturm sequences we can verify that p_2 has roots ± 1, $\pm 1/2$, and 0 in the interval $[-1, 1]$. That is, for distinct $x, y \in S^3$, $A_2K(\{x, y\}) = 0$ if and only if $\langle x, y\rangle \in \{-1, \pm 1/2, 0\}$. In the remainder of this section, we use this fact to show the D_4 root system is the unique optimal kissing configuration up to isometry and is an optimal spherical code.

Lemma 5.1. *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. Let K be the exact solution discussed above. By positivity of K and by the linear constraints

$$A_2K(Q) \leq -1_{\mathcal{I}_{=1}}(Q), \qquad Q \in \mathcal{I}_4 \setminus \{\emptyset\},$$

we have

$$0 \leq \sum_{\substack{J_1, J_2 \in \mathcal{I}_2 \\ J_1, J_2 \subseteq C}} K(J_1, J_2) = \sum_{\substack{Q \in \mathcal{I}_4 \\ Q \subseteq C}} A_2K(Q) \leq K(\emptyset, \emptyset) - |C|.$$

Since $|C| = 24$, equality holds throughout, so in particular, $A_2K(Q) = 0$ for all $Q \subseteq C$ with $|Q| \leq 4$. As mentioned above, for distinct $x, y \in S^3$, we have $A_2K(\{x, y\}) = 0$ if and only if $\langle x, y\rangle \in \{-1, \pm 1/2, 0\}$, which proves the lemma. □

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 5.2. *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 5.1. □

Theorem 5.3. *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 5.1, 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 5.1, 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, [50, 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 $\Phi_1, \dots, \Phi_k$, 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. □

We sketch an alternative proof of Theorem 5.3 which does not rely on the classification of irreducible root systems. Consider an optimal spherical code. As argued in the above proof, the code is antipodal. Now [6, Proposition 3.12] implies that the corresponding set of 12 lines is the union of 3 orthonormal bases, with lines in different bases not orthogonal. Choose one of these bases to define coordinates. The basis elements and their negatives give 8 points. By Lemma 5.1, the remaining 16 must have ±1/2 in every coordinate. The only possibility is to have every such point, and so the resulting configuration is unique.

5.2. **New bound in dimension six.** We also computed the second level of the hierarchy with the parameters $d_1 = 14$ and $d_2 = \delta = 16$ for the kissing number problem in dimensions 5, 6, 7, 10, 12, and 16. In dimension 6 this gives $k(6) \leq 77$, which improves on the previously best-known upper bound of $k(6) \leq 78$ obtained using the three-point bound [2]. As mentioned in the introduction, the degrees for which we perform computations are not yet high enough to get improvements in the other dimensions.

In dimension 6, this does not give a sharp bound, so that the optimal objective value and optimal solution potentially require high algebraic degree or bit size. This means the rounding procedure from [8] may not be able to find an exact feasible 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 $k(6) \leq 77$.

Acknowledgements

We are grateful to Henry Cohn and Fernando Oliveira for helpful discussions. We also thank Fernando Oliveira for letting us use his unpublished Julia code for invariant sum-of-squares polynomials, and we thank Henry Cohn for making us aware of the alternative proof below Theorem 5.3. We thank Joffrey Wallaart for ICT support for the computational part of this project.

References

[1] V. V. Arestov and A. G. Babenko, *On delsarte scheme of estimating the contact numbers*, Proc. of the Steklov Inst. of Math. (1997), no. 219, 36–65.

[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. MR2393433 doi:10.1090/S0894-0347-07-00589-9

[3] E. Bannai and N. J. A. Sloane, *Uniqueness of certain spherical codes*, Canadian Journal of Mathematics **33** (1981), no. 2, 437–449. doi:10.4153/CJM-1981-038-7

[4] B. Bekker and F. M. de Oliveira Filho, *On the convergence of the k-point bound for topological packing graphs*, preprint, 2023. arXiv:2306.02725

[5] L. Bröcker, *On symmetric semialgebraic sets and orbit spaces*, Banach Center Publ. **44** (1998), no. 1, 37–50.

[6] A. R. Calderbank, P. J. Cameron, W. M. Kantor, and J. J. Seidel, *$\mathbb{Z}_4$-kerdock codes, orthogonal spreads, and extremal euclidean line-sets*, Proc. Lond. Math. Soc. **75** (1997), no. 2, 436–480.

[7] H. Cohn, J. H. Conway, N. D. Elkies, and A. Kumar, *The D_4 root system is not universally optimal*, Experiment. Math. **16** (2007), no. 3, 313–320. MR2367321

[8] H. Cohn, D. de Laat, and N. Leijenhorst, *Optimality of spherical codes via exact semidefinite programming bounds*, preprint, 2024. arXiv:2403.16874

[9] H. Cohn, D. de Laat, and A. Salmon, *Three-point bounds for sphere packing*, preprint, 2022. arXiv:2206.15373

[10] H. Cohn and N. Elkies, *New upper bounds on sphere packings. I*, Ann. of Math. (2) **157** (2003), no. 2, 689–714. MR1973059 doi:10.4007/annals.2003.157.689

[11] 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

[12] H. Cohn and A. Salmon, *Sphere packing bounds via rescaling*, preprint, 2021. arXiv:2108.10936

[13] H. S. M. Coxeter, *An upper bound for the number of equal nonoverlapping spheres that can touch another of the same size*, Proc. Sympos. Pure Math., Vol. VII, Amer. Math. Soc., Providence, RI, 1963, pp. 53–71. MR164283

[14] D. de Laat, N. M. Leijenhorst, and W. H. H. de Muinck Keizer, *Data for "Optimality and uniqueness of the D_4 root system"*, data set, 4TU.ResearchData, 2024. doi:10.4121/74ce1c25-6fca-4680-8a36-e9c18e7e9594

[15] D. de Laat, *Moment methods in energy minimization: new bounds for Riesz minimal energy problems*, Trans. Amer. Math. Soc. **373** (2020), no. 2, 1407–1453. MR4068268 doi:10.1090/tran/7976

[16] D. de Laat, F. C. Machado, F. M. de Oliveira Filho, and F. Vallentin, *k-point semidefinite programming bounds for equiangular lines*, Math. Program. **194** (2022), no. 1-2, Ser. A, 533–567. MR4445463 doi:10.1007/s10107-021-01638-x

[17] D. de Laat, F. C. Machado, and W. de Muinck Keizer, *The lasserre hierarchy for equiangular lines with a fixed angle*, preprint, 2023. arXiv:2211.16471

[18] D. de Laat and F. Vallentin, *A semidefinite programming hierarchy for packing problems in discrete geometry*, Math. Program. **151** (2015), no. 2, Ser. B, 529–553. MR3348162 doi:10.1007/s10107-014-0843-4

[19] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. (1973), no. 10, vi+97. MR384310

[20] P. Delsarte, J.-M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geometriae Dedicata **6** (1977), no. 3, 363–388. MR0485471

[21] M. Dostert, D. de Laat, and P. Moustrou, *Exact semidefinite programming bounds for packing problems*, SIAM J. Optim. **31** (2021), no. 2, 1433–1458. MR4263438 doi:10.1137/20M1351692

[22] W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991. MR1153249 doi:10.1007/978-1-4612-0979-9

[23] K. Gatermann and P. A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Algebra **192** (2004), no. 1-3, 95–128. MR2067190 doi:10.1016/j.jpaa.2003.12.011

[24] T. Gorin and G. V. López, *Monomial integrals on the classical groups*, Journal of Mathematical Physics **49** (2008), no. 1, 013503. doi:10.1063/1.2830520

[25] K. I. Gross and R. A. Kunze, *Finite dimensional induction and new results on invariants for classical groups, I*, Am. J. Math **106** (1984), no. 4, 893–974.

[26] T. C. Hales, *A proof of the Kepler conjecture*, Ann. of Math. (2) **162** (2005), no. 3, 1065–1185. MR2179728 doi:10.4007/annals.2005.162.1065

[27] A. Korkine and G. Zolotareff, *Sur les formes quadratiques positives quaternaires*, Math. Ann. **5** (1872), no. 4, 581–583. MR1509795 doi:10.1007/BF01442912

[28] O. Kuryatnikova and J. Vera, *Approximating the cone of copositive kernels to estimate the stability number of infinite graphs*, Electron. Notes Discrete Math. **62** (2017), 303–308. doi:10.1016/j.endm.2017.10.052

[29] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. **11** (2000/01), no. 3, 796–817. MR1814045 doi:10.1137/S1052623400366802

[30] ———, *An explicit equivalent positive semidefinite program for nonlinear* 0-1 *programs*, SIAM J. Optim. **12** (2002), no. 3, 756–769. MR1884916 doi:10.1137/S1052623400380079

[31] N. Leijenhorst and D. de Laat, *Solving clustered low-rank semidefinite programs arising from polynomial optimization*, preprint, 2022. arXiv:arXiv:2202.12077

[32] V. I. Levenšteĭn, *On bounds for packings in n-dimensional euclidean space*, Dokl. Akad. Nauk SSSR **245** (1979), no. 6, 1299–1303. MR529659

[33] 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

[34] F. C. Machado and F. M. de Oliveira Filho, *Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry*, Experiment. Math. (2017), 362–369, arXiv:1609.05167.

[35] M. Marshall, *Positive polynomials and sums of squares*, Math. Surveys Monogr., no. 146, Amer. Math. Soc., 2008.

[36] H. D. Mittelmann and F. Vallentin, *High-accuracy semidefinite programming bounds for kissing numbers*, Experiment. Math. **19** (2010), no. 2, 175–179. MR2676746 doi:10.1080/10586458.2010.10129070

[37] O. R. Musin, *Semidefinite programming bounds for distance distribution of spherical codes*, preprint, 2023. arXiv:arXiv:2309.13854

[38] ———, *The kissing number in four dimensions*, Ann. of Math. (2) **168** (2008), no. 1, 1–32. MR2415397 doi:10.4007/annals.2008.168.1

[39] ———, *Multivariate positive definite functions on spheres*, Discrete geometry and algebraic combinatorics, Contemp. Math., vol. 625, Amer. Math. Soc., Providence, RI, 2014, pp. 177–190. MR3289412 doi:10.1090/conm/625/12498

[40] ———, *Towards a proof of the 24-cell conjecture*, Acta Math. Hungar. **155** (2018), no. 1, 184–199. doi:10.1007/s10474-018-0828-5

[41] J. Nie, K. Ranestad, and B. Sturmfels, *The algebraic degree of semidefinite programming*, Math. Program. **122** (2010), no. 2, 379–405. MR2546336 doi:10.1007/s10107-008-0253-6

[42] A. M. Odlyzko and N. J. A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Combin. Theory Ser. A **26** (1979), no. 2, 210–214. MR530296 doi:10.1016/0097-3165(79)90074-8

[43] A. M. Odlyzko and N. J. A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Combin. Theory Ser. A **26** (1979), no. 2, 210–214. MR530296 doi:10.1016/0097-3165(79)90074-8

[44] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, PhD thesis, Caltech, 2000.

[45] F. Pfender and G. M. Ziegler, *Kissing numbers, sphere packings, and some unexpected proofs*, Notices Amer. Math. Soc. **51** (2004), no. 8, 873–883. MR2145821

[46] V. Powers and B. Reznick, *Polynomials that are positive on an interval*, Trans. Amer. Math. Soc. **352** (2000), no. 10, 4677–4692.

[47] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J. **42** (1993), no. 3, 969–984. MR1254128 doi:10.1512/iumj.1993.42.42045

[48] A. Schrijver, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory **51** (2005), no. 8, 2859–2866. MR2236252 doi:10.1109/TIT.2005.851748

[49] K. Schütte and B. L. van der Waerden, *Das Problem der dreizehn Kugeln*, Math. Ann. **125** (1953), 325–334. MR53537 doi:10.1007/BF01343127

[50] N. J. A. Sloane, *Tables of sphere packings and spherical codes*, IEEE Trans. Inform. Theory **27** (1981), no. 3, 327–338. MR619118 doi:10.1109/TIT.1981.1056351

[51] M. S. Viazovska, *The sphere packing problem in dimension 8*, Ann. of Math. (2) **185** (2017), no. 3, 991–1015. doi:10.4007/annals.2017.185.3.7

D. de Laat, Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands
Email address: `d.delaat@tudelft.nl`

N. M. Leijenhorst, Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands
Email address: `n.m.leijenhorst@tudelft.nl`

W. H. H. de Muinck Keizer, Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands
Email address: `w.h.h.demuinckkeizer@tudelft.nl`