Discrete Comput Geom (2017) 58:449–481
DOI 10.1007/s00454-017-9882-y

New Upper Bounds for the Density of Translative Packings of Three-Dimensional Convex Bodies with Tetrahedral Symmetry

Maria Dostert[1] · **Cristóbal Guzmán**[2] · **Fernando Mário de Oliveira Filho**[3] · **Frank Vallentin**[1]

Received: 21 March 2016 / Revised: 10 January 2017 / Accepted: 19 February 2017 / Published online: 9 March 2017

Abstract In this paper we determine new upper bounds for the maximal density of translative packings of superballs in three dimensions (unit balls for the l_3^p-norm) and of Platonic and Archimedean solids having tetrahedral symmetry. Thereby, we improve Zong's recent upper bound for the maximal density of translative packings of regular tetrahedra from 0.3840 ... to 0.3745 ..., getting closer to the best known lower bound of 0.3673 ... We apply the linear programming bound of Cohn and Elkies which originally was designed for the classical problem of densest packings of round spheres. The proofs of our new upper bounds are computational and rigorous. Our main technical contribution is the use of invariant theory of pseudo-reflection groups in polynomial optimization.

Editor in Charge: Günter M. Ziegler

Maria Dostert
m.dostert@uni-koeln.de

Cristóbal Guzmán
crguzmanp@uc.cl

Fernando Mário de Oliveira Filho
fmario@gmail.com

Frank Vallentin
frank.vallentin@uni-koeln.de

1 Mathematisches Institut, Universität zu Köln, Weyertal 86–90, 50931 Cologne, Germany

2 Facultad de Matemáticas & Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile

3 Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão, 1010, São Paulo 05508-090, Brazil

Keywords Translative packings · Sums of Hermitian squares · Pseudo-reflections · Superballs · Platonic solids · Archimedean solids · Hilbert's 18th problem · Semidefinite programming · Interval arithmetic

Mathematics Subject Classification 52C17 · 90C22

1 Introduction

The most famous geometric packing problem is Kepler's conjecture from 1611: The density of any packing of equal-sized spheres into three-dimensional Euclidean space is never greater than $\pi/\sqrt{18} = 0.7404\ldots$ This density is achieved for example by the "cannonball" packing. In 1998 Hales and Ferguson solved Kepler's conjecture. Their proof is extremely complicated, involving more than 200 pages, intensive computer calculations, and checking of more than 5000 subproblems. They wrote a book [25] that contains the entire proof together with supporting material and commentary.

Very little is known if one goes beyond three-dimensional packings of spheres to packings of nonspherical objects. Considering nonspherical objects is interesting for many reasons. For example, using nonspherical objects one can model physical granular materials accurately. On the other hand, the mathematical difficulty increases substantially when one deals with nonspherical objects.

Jiao et al. [29] consider packings of three-dimensional superballs, which are unit balls of the l_3^p-norm, with $p \geq 1$:

$$B_3^p = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : |x_1|^p + |x_2|^p + |x_3|^p \leq 1\}.$$

Three-dimensional superballs can be synthesized experimentally as colloids, see Rossi et al. [41]. The name "superball" is attributed to the Danish inventor Piet Hein who used a "superellipse" with $p = 5/2$ in a design challenge of the redevelopment of the public square Sergels Torg in Stockholm (see also Rush and Sloane [42], and Gardner [18]). Life Magazine [26] quotes Piet Hein:

> Man is the animal that draws lines which he himself then stumbles over. In the whole pattern of civilization there have been two tendencies, one toward straight lines and rectangular patterns and one toward circular lines. There are reasons, mechanical and psychological, for both tendencies. Things made with straight lines fit well together and save space. And we can move easily—physically or mentally—around things made with round lines. But we are in a straitjacket, having to accept one or the other, when often some intermediate form would be better.

Back to the work of Jiao et al. [29]. They construct the densest known packings of B_3^p for many values of p. As motivation for their study Jiao, Stillinger, and Torquato write:

> Understanding the organizing principles that lead to the densest packings of nonspherical particles that do not tile space is of great practical and fundamental interest. Clearly, the effect of asphericity is an important feature to include on

> the way to characterizing more fully real dense granular media.
> [...]
> On the theoretical side, no results exist that rigorously prove the densest packings of other congruent non-space-tiling particles in three dimension.

Torquato and Jiao [49,50] extend the work on superballs to nonspherical non-differentiable shapes. They found dense packings of Platonic and of Archimedean solids.

> Very little is known about the densest packings of polyhedral particles that do not tile space, including the majority of the Platonic and Archimedean solids studied by the ancient Greeks. The difficulty in obtaining dense packings of polyhedra is related to their complex rotational degrees of freedom and to the non-smooth nature of their shapes.

The optimal, densest lattice packing of each Platonic or Archimedean solid is known. Minkowski [35] determines the densest lattice packing of regular octahedra. Hoylman [27] uses Minkowski's method to determine the densest lattice packing of regular tetrahedra. Betke and Henk [3] turn Minkowski's method into an implementable algorithm and find the densest lattice packings of all remaining Platonic and all Archimedean solids. Only two of the Platonic and Archimedean solids are not centrally symmetric, namely the tetrahedron and the truncated tetrahedron. These are also the only cases where Torquato and Jiao could use the extra freedom of rotating the solids to find new packings which are denser than the corresponding densest lattice packings. Also the dense superball packings of Jiao, Stillinger, and Torquato are lattice packings. Based on this evidence they formulate the following conjecture:

> The densest packings of the centrally symmetric Platonic and Archimedean solids are given by their corresponding optimal lattice packings. This is the analogue of Kepler's sphere conjecture for these solids.

For a convex body $\mathcal{K}$ in $\mathbb{R}^n$ it is natural to consider three kinds of increasingly restrictive packings: congruent packings, translative packings, and lattice packings. A packing of congruent copies of $\mathcal{K}$ has the form

$$\mathcal{P} = \bigcup_{i \in \mathbb{N}} (x_i + A_i\mathcal{K}), \quad \text{with } (x_i, A_i) \in \mathbb{R}^n \times \mathrm{SO}(n),\ i \in \mathbb{N},$$

where $x_i + A_i\mathcal{K}^\circ \cap x_j + A_j\mathcal{K}^\circ = \emptyset$ whenever the indices i and j are distinct. Here, $\mathcal{K}^\circ$ denotes the topological interior of the body $\mathcal{K}$ and

$$\mathrm{SO}(n) = \{A \in \mathbb{R}^{n\times n} : AA^{\mathsf{T}} = I_n,\ \det A = 1\}$$

denotes the special orthogonal group, an index-2 subgroup of the orthogonal group

$$\mathrm{O}(n) = \{A \in \mathbb{R}^{n\times n} : AA^{\mathsf{T}} = I_n\}.$$

The (upper) density of $\mathcal{P}$ is

$$\delta(\mathcal{P}) = \limsup_{r\to\infty} \sup_{c\in\mathbb{R}^n} \frac{\operatorname{vol}(B(c,r)\cap\mathcal{P})}{\operatorname{vol} B(c,r)},$$

where $B(c, r)$ is the Euclidean ball of radius r centered at c. A packing $\mathcal{P}$ is called a translative packing if each rotation A_i is identity. A translative packing is called a lattice packing if the set of x_i's forms a lattice.

Lattice packings are restrictive and many results about them are known. This is not the case for translative and congruent packings. While the conjecture of Torquato and Jiao ultimately aims at congruent packings, the objective of our paper is to develop tools coming from mathematical optimization which will be useful to make progress on the conjecture restricted to translative packings. In particular we prove new upper bounds for the density of densest translative packings of three-dimensional superballs and of Platonic and Archimedean solids with tetrahedral symmetry. We use the following theorem of Cohn and Elkies [6] for this.

Theorem 1.1 *Let $\mathcal{K}$ be a convex body in $\mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous L^1-function (a continuous function whose absolute value is Lebesgue integrable). Define the Fourier transform of f at u by*

$$\widehat{f}(u) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i u\cdot x}\, dx.$$

Suppose f satisfies the following conditions:

(i) $\widehat{f}(0) \geq 1$,
(ii) *f is of positive type, i.e. $\widehat{f}(u) \geq 0$ for every $u \in \mathbb{R}^n$,*
(iii) *$f(x) \leq 0$ whenever $\mathcal{K}^\circ \cap (x + \mathcal{K}^\circ) = \emptyset$.*

Then the density of any packing of translates of $\mathcal{K}$ in $\mathbb{R}^n$ is at most $f(0) \operatorname{vol} \mathcal{K}$.

One can find a proof of this theorem in Cohn and Kumar [7] or for the more general case of translative packings of multiple convex bodies $\mathcal{K}_1, \ldots, \mathcal{K}_N$ in de Laat et al. [33]. Originally, Cohn and Elkies [6] state the theorem only for admissible functions; these are functions for which the Poisson summation formula applies.

The Cohn–Elkies bound provides the basic framework for proving the best known upper bounds for the maximum density of sphere packings in dimensions 4, ..., 36. For some time it was conjectured to provide tight bounds in dimensions 8 and 24 and there was very strong numerical evidence to support this conjecture, see Cohn and Miller [9]. However, the only thing missing was a rigorous proof. Recently, in March 2016, such a proof was found by Viazovska [51] for dimension 8 and a few days later, building on Viazovska's breakthrough result, by Cohn et al. [8] for dimension 24. Here the explicit construction of optimal functions f uses the theory of quasimodular forms from analytic number theory.

De Laat et al. [33] have proposed a strengthening of the Cohn–Elkies bound and computed better upper bounds for the maximum density of sphere packings in dimensions 4, 5, 6, 7, and 9.

In all these calculations one can restrict the function f to be a radial function (a function whose values $f(x)$ only depend on the norm $\|x\|$ of the vector $x \in \mathbb{R}^n$) because of the rotational symmetry of the sphere. For the case of packings of nonspherical objects Cohn and Elkies [6] write:

> Unfortunately, when [the body we want to pack] is not a sphere, there does not seem to be a good analogue of the reduction to radial functions in Theorem [1.1]. That makes these cases somewhat less convenient to deal with.

Until now, the Cohn–Elkies bound has only been computed for packings of spheres. In this paper we show how to apply the Cohn–Elkies bound for nonspherical objects.

1.1 New Upper Bounds for Translative Packings

Before we describe our methods we report on the new upper bounds we obtained and compare them to the known lower and upper bounds. We give the new upper bounds for three-dimensional superball packings in Table 1, the new upper bounds for Platonic and Archimedean solid with tetrahedral symmetry are in Table 2.

1.1.1 Three-Dimensional Superballs

Jiao et al. [29,30] find dense packings of superballs B_3^p for all values of $p \geq 1$. Although they principally allow congruent packings in their computer simulations, the dense packings they find are all lattice packings. They subdivide the range $p \in [1, \infty)$ into four different regimes

$$p \in [1, 2\ln 3/\ln 4 = 1.5849\ldots] \cup [1.5849\ldots, 2] \cup [2, 2.3018\ldots] \cup [2.3018\ldots, \infty)$$

and give for each regime a family of lattices determining dense packings. When $p = 1$ then B_3^p is simply the regular octahedron, a Platonic solid. The optimal lattice packing of regular octahedra has been determined by Minkowski [35].

Recently, Ni et al. [36] claimed that for values of p lying in the interior of the first and respectively of the second regime, the packings of Jiao, Stillinger, and Torquato can be improved.

When $p = 2$, then B_3^p is the round unit ball of Euclidean space. The optimal lattice packing of B_3^2 has been determined by Gauss [20] using reduction theory of positive quadratic forms. Here, translative and congruent packings coincide because of the rotational symmetry of B_3^2. Hales [25] proved the optimality of the cannonball packing among all congruent packings. One should note that there is an uncountable family of non-lattice packings achieving the same density. The best upper bound obtainable from Theorem 1.1 is $0.7797\ldots$

For $p \geq 2.3018\ldots$ the densest known superball packings are given by the family of C_1-lattices which is defined by $\mathbb{Z}b_1 + \mathbb{Z}b_2 + \mathbb{Z}b_3$ with

$$b_1 = (2^{1-1/p}, 2^{1-1/p}, 0),\ b_2 = (2^{1-1/p}, 0, 2^{1-1/p}),\ b_3 = (2s + 2^{1-1/p}, -2s, -2s),$$

Table 1 Best known bounds for packings of three-dimensional superballs

Body	Lattice packing	
	Lower bound	Upper bound
B_3^1	$18/19 = 0.9473\ldots$ [35]	$18/19$ [35]
B_3^2	$\pi/\sqrt{18} = 0.7404\ldots$	$\pi/\sqrt{18}$ [20]
B_3^3	$0.8095\ldots$ [29]	*0.8236…*
B_3^4	$0.8698\ldots$ [29]	*0.8742…*
B_3^5	$0.9080\ldots$ [29]	*0.9224…*
B_3^6	$0.9318\ldots$ [29]	*0.9338…*
	Translative packing	
	Lower bound	Upper bound
B_3^1	$18/19 = 0.9473\ldots$ [35]	$0.9729\ldots$
B_2^2	$\pi/\sqrt{18} = 0.7404\ldots$	$\pi/\sqrt{18}$ [25]
B_3^3	$0.8095\ldots$ [29]	*0.8236…*
B_3^4	$0.8698\ldots$ [29]	*0.8742…*
B_3^5	$0.9080\ldots$ [29]	*0.9224…*
B_3^6	$0.9318\ldots$ [29]	*0.9338…*
	Congruent packing	
	Lower bound	Upper bound
B_3^1	$18/19 = 0.9473\ldots$ [35]	$1 - 1.4\ldots \cdot 10^{-12}$ [23]
B_2^2	$\pi/\sqrt{18} = 0.7404\ldots$	$\pi/\sqrt{18}$ [25]
B_3^3	$0.8095\ldots$ [29]	< 1
B_3^4	$0.8698\ldots$ [29]	< 1
B_3^5	$0.9080\ldots$ [29]	< 1
B_3^6	$0.9318\ldots$ [29]	< 1

New bounds obtained in this paper are written in italics

where s is the smallest positive root of the equation

$$(s + 2^{-1/p})^p + 2s^p - 1 = 0,$$

having density

$$\frac{\operatorname{vol} B_3^p}{2^{3-2/p}(3s + 2^{-1/p})}.$$

1.1.2 Platonic and Archimedean Solids with Tetrahedral Symmetry

We prove a new upper bound for the density of densest translative packings of regular tetrahedra and improve the upper bound of $0.3840\ldots$ recently obtained by Zong [53]

Table 2 Best known bounds for packings of three-dimensional Platonic and Archimedean solids with tetrahedral symmetry

Body	Lattice packing	
	Lower bound	Upper bound
Tetrahedron	$18/49 = 0.3673\ldots$ [24]	$18/49$ [27]
Truncated tetrahedron	$0.6809\ldots$ [3]	$0.6809\ldots$ [3]
Truncated cuboctahedron	$0.8493\ldots$ [3]	$0.8493\ldots$ [3]
Rhombicuboctahedron	$0.8758\ldots$ [3]	$0.8758\ldots$ [3]
Cuboctahedron	$0.9183\ldots$ [24]	$0.9183\ldots$ [27]
Truncated cube	$0.9737\ldots$ [3]	$0.9737\ldots$ [3]
	Translative packing	
	Lower bound	Upper bound
Tetrahedron	$18/49 = 0.3673\ldots$ [24]	*0.3745* ...
Truncated tetrahedron	$0.6809\ldots$ [3]	*0.7292* ...
Truncated cuboctahedron	$0.8493\ldots$ [3]	$0.8758\ldots$ [50]
Rhombicuboctahedron	$0.8758\ldots$ [3]	$0.8758\ldots$ [22]
Cuboctahedron	$0.9183\ldots$ [24]	*0.9364* ...
Truncated cube	$0.9737\ldots$ [3]	*0.9845* ...
	Congruent packing	
	Lower bound	Upper bound
Tetrahedron	$4000/4671 = 0.8563\ldots$ [5]	$1 - 2.6\ldots \cdot 10^{-25}$ [23]
Truncated tetrahedron	$207/208 = 0.9951\ldots$ [31], [11]	< 1
Truncated cuboctahedron	$0.8493\ldots$ [3]	$0.8758\ldots$ [50]
Rhombicuboctahedron	$0.8758\ldots$ [3]	$0.8758\ldots$ [22]
Cuboctahedron	$0.9183\ldots$ [24]	< 1
Truncated cube	$0.9737\ldots$ [3]	< 1

The octahedron, the cube and the truncated octahedron are omitted. New bounds obtained in this paper are written in italics

to $0.3745\ldots$ Groemer [24] shows that there is a lattice packing of regular tetrahedra which has density $0.3673\ldots$ and Hoylman [27] proves the optimality of Groemer's packing among lattice packings using Minkowski's method. Finding dense congruent packings of regular tetrahedra is fascinating. In fact, it is part of Hilbert's 18th problem. We refer to Lagarias and Zong [34] for the history of the tetrahedra packing problem and to Ziegler [52] for an overview on a race for the best construction.

As a corollary of our new bound for the tetrahedron, we improve Zong's bound for densest translative packings of the cuboctahedron from $0.9601\ldots$ to $0.9364\ldots$ This follows from Minkowski's observation that

$$\bigcup_{i \in \mathbb{N}} (x_i + \mathcal{K})$$

is a translative packing of $\mathcal{K}$ if and only if

$$\bigcup_{i\in\mathbb{N}}\Big(x_i + \frac{1}{2}(\mathcal{K}-\mathcal{K})\Big)$$

is a packing of $(\mathcal{K}-\mathcal{K})/2$, with

$$\mathcal{K}-\mathcal{K} = \{x-y : x, y \in \mathcal{K}\}$$

denoting the Minkowski difference of the body $\mathcal{K}$ with itself. The Minkowski difference of a regular tetrahedron with itself is the cuboctahedron whose volume is $2^3 \cdot 5/2$ times the volume of the regular tetrahedron.

We omitted the octahedron from Table 2 because it is B_3^1 in Table 1. We also omitted the cube and the truncated octahedron because with both solids one can tile three-dimensional space.

1.2 Computational Strategy

In this section we give a high level description of how we found suitable functions f for Theorem 1.1 for proving the new upper bounds.

The symmetry group of a convex body $\mathcal{K} \subseteq \mathbb{R}^n$ is

$$S(\mathcal{K}) = \{A \in \mathrm{O}(n) : A\mathcal{K} = \mathcal{K}\},$$

and when considering functions f for Theorem 1.1 the symmetry group of the Minkowski difference $\mathcal{K}-\mathcal{K}$ will be useful. For $A \in S(\mathcal{K}-\mathcal{K})$ we have

$$\begin{aligned} A^{-1}x + \mathcal{K}^\circ \cap \mathcal{K}^\circ = \emptyset &\iff x + A\mathcal{K}^\circ \cap A\mathcal{K}^\circ = \emptyset \iff x \notin A\mathcal{K}^\circ - A\mathcal{K}^\circ \\ &\iff x \notin A(\mathcal{K}^\circ - \mathcal{K}^\circ) \iff x \notin \mathcal{K}^\circ - \mathcal{K}^\circ \iff x + \mathcal{K}^\circ \cap \mathcal{K}^\circ = \emptyset. \end{aligned}$$

Hence we may assume without loss of generality that the function f we are seeking is invariant under the left action of $S(\mathcal{K}-\mathcal{K})$, i.e.

$$f(A^{-1}x) = f(x) \quad \text{for all } A \in S(\mathcal{K}-\mathcal{K}).$$

This assumption reduces the search space and also makes the third constraint $f(x) \leq 0$ whenever $x + \mathcal{K}^\circ \cap \mathcal{K}^\circ = \emptyset$ easier to model.

In the case of $\mathcal{K}$ being a superball B_3^p, with $p \geq 1$ and $p \neq 2$, the Minkowski difference $\mathcal{K}-\mathcal{K}$ is $2B_3^p$ and its symmetry group is a finite subgroup of the orthogonal group. It is the octahedral group (which is the same as the symmetry group of the regular cube $[-1,+1]^3$), which has 48 elements. The octahedral group is the reflection group B_3 which is generated by the three matrices

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{1}$$

where the first one is the reflection at the plane $x_1 = 0$, the second one is the reflection at the plane $x_2 + x_3 = 0$, and the last one is the reflection at the plane $x_1 - x_3 = 0$.

In the case of $\mathcal{K}$ being a Platonic or Archimedean solid with tetrahedral symmetry, the symmetry group of the Minkowski difference $\mathcal{K} - \mathcal{K}$ is the octahedral group, too.

We specify the function $f : \mathbb{R}^3 \to \mathbb{R}$ via its Fourier transform $\widehat{f}$. If f is invariant under the action of $S(\mathcal{K} - \mathcal{K})$ then the same is true for its Fourier transform $\widehat{f}$. Let g be a polynomial. We use the following template for the Fourier transform of f:

$$\widehat{f}(u) = g(u)e^{-\pi\|u\|^2}. \tag{2}$$

So $\widehat{f}$ is a Schwartz function (i.e. all derivates $D^\beta \widehat{f}(x)$ exist for all $x \in \mathbb{R}^n$ and all $\beta \in \mathbb{N}^n$ and $\sup\{|x^\alpha D^\beta \widehat{f}(x)| : x \in \mathbb{R}^n\} < \infty$ holds for all $\alpha, \beta \in \mathbb{N}^n$), implying that also f is a Schwartz function. In particular, f will be a continuous L^1-function.

If f is invariant under B_3 then so is the polynomial g which specifies $\widehat{f}$. This means that g lies in the ring of invariants of the group B_3 which is, by the theory of finite reflection groups, known to be freely generated by three basic invariants

$$\theta_1 = x_1^2 + x_2^2 + x_3^2, \quad \theta_2 = x_1^4 + x_2^4 + x_3^4, \quad \theta_3 = x_1^6 + x_2^6 + x_3^6. \tag{3}$$

Thus, we can assume that g lies in the polynomial ring $\mathbb{R}[\theta_1, \theta_2, \theta_3]$.

The first condition of Theorem 1.1 is a simple linear condition in the coefficients of the polynomial g, it says

$$g(0) \geq 1.$$

For the second condition we want the function f whose Fourier transform is given by (2) to be of positive type. This is true if and only if g is globally nonnegative. In general, checking that a polynomial is nonnegative everywhere is computationally difficult, it is an NP-hard problem. We use a standard relaxation of the global nonnegativity constraint by imposing a sufficient condition which is easier to check: We want that g can be written as a sum of squares, which we can formulate as a semidefinite condition. Furthermore, we can use the imposed B_3-invariance of g to simplify this semidefinite condition. In Sect. 2 we work out the theory of this simplification for the case of pseudo-reflection groups. In Sect. 3 we apply the theory to the finite reflection group B_3.

Although using this sum of squares relaxation works very well in practice, we are indeed restricting the search space of functions. Hilbert showed in 1888 that there are polynomials already in two variables which are globally nonnegative but which are not sum of squares. Hilbert's proof was nonconstructive and only in 1967 Motzkin published the first explicit example. Shortly afterwards Robinson showed that the B_3-invariant polynomial

$$x_1^6 + x_2^6 + x_3^6 - \left(x_1^4x_2^2 + x_1^2x_2^4 + x_1^4x_3^2 + x_1^2x_3^4 + x_2^4x_3^2 + x_2^2x_3^4\right) + 3x_1^2x_2^2x_3^2 \tag{4}$$

is nonnegative but not a sum of squares. We refer the interested reader to Reznick [40] for more on this.

For the third condition we first have to compute f from $\widehat{f}$. This is an easy linear algebra computation once we decompose g as a sum of products of radial polynomials times harmonic polynomials. We review this decomposition in Sect. 4.

Finally, we want that f be nonpositive outside of $\mathcal{K}^\circ - \mathcal{K}^\circ$. When $\mathcal{K} = B_3^p$ and when p is an even integer we can use another sufficient sum of squares condition:

$$f(x_1, x_2, x_3)e^{\pi\|x\|^2} + (x_1^p + x_2^p + x_3^p - 2)q_1(x_1, x_2, x_3) + q_2(x_1, x_2, x_3) = 0, \tag{5}$$

where q_1 and q_2 are B_3-invariant polynomials which can be written as sum of squares. This again can be expressed as a semidefinite condition.

So in the end we can find a good function f, minimizing $f(0)$, for Theorem 1.1 by solving a finite semidefinite programming problem, once we restrict the degrees of polynomials g, q_1, and q_2. We give an explicit finite-dimensional semidefinite programming formulation in Sect. 5.

A semidefinite programming problem—a rich generalization of linear programming—amounts to minimizing a linear function over an spectrahedron, the intersection of the cone of positive semidefinite matrices with an affine subspace. For solving semidefinite programming problems in practice one uses interior point methods. There are many very good software implementations of interior point methods available. For verifying that we proved a rigorous bound we only have to show that the solution the software gave to us is a function f which satisfies the conditions of Theorem 1.1. In Sect. 6 we explain this verification process in detail.

When p is not an even integer, the approach of using sum of squares in (5) breaks down. To get an upper bound we use a sum of squares condition for the next largest even integer p' and use a fine sample of points in the intersection of the set $B_3^p \setminus B_3^{p'}$ with the fundamental domain

$$0 \le x_1 \le x_2 \le x_3 \tag{6}$$

of the group B_3 to make sure that the function f is nonpositive there. With this we get a function f which almost satisfies the conditions of Theorem 1.1. It turns out, and we check this fact rigorously, that f satisfies the conditions for a slightly larger body $\alpha\mathcal{K}$ with α only slightly larger than one. Then we obtain the slightly weaker bound of $\alpha^3 f(0) \operatorname{vol} \mathcal{K}$.

When dealing with polytopes $\mathcal{K}$ we use a similar approach: We impose the sum of squares condition

$$f(x_1, x_2, x_3)e^{\pi\|x\|^2} + (x_1^2 + x_2^2 + x_3^2 - r)q_1(x_1, x_2, x_3) + q_2(x_1, x_2, x_3) = 0, \tag{7}$$

where r is the circumradius of the polytope $\mathcal{K}$. Again we use a fine sample of points in the intersection of the set $rB_3^2 \setminus \mathcal{K}$ with the fundamental domain (6) of the finite reflection group B_3 to make sure that the function f is nonpositive there.

1.3 Future Research

We end the introduction by showing directions and questions for possible future research.

Our bounds give hope that the Cohn–Elkies bound might be strong enough to prove optimality of the C_1-lattices for some values of p among all translative packings of superballs. Our computations were restricted, due to numerical difficulties, to polynomials g of rather small degrees and to sum of square certificates for nonnegativity. Does there exist a threshold $p' < \infty$ so that for all $p \geq p'$ the Cohn–Elkies bound is tight?

So the development of better computational techniques to compute the Cohn–Elkies bound would be very valuable. It also would be of interest to perform more computations. Pütz, in his master's thesis [38], computed bounds for Platonic and Archimedean solids with icosahedral symmetries. Bounds for superball or polytope packings in dimension 4 have not yet been computed.

When computing the bound for translative packings of B_3^p with odd p or for translative packings of polytopes we used sampling. This makes finding a rigorous proof more difficult. Is it possible to find a more convenient method to prove rigorous (and better) bounds?

Is it possible to apply Minkowski's method, or a variant of the algorithm of Betke and Henk to determine optimal lattice packings of three-dimensional superballs?

Cohn and Zhao [10] improve the asymptotic sphere packing bound by Kabatiansky and Levenshtein [32] slightly and show that the Cohn–Elkies bound is at least as strong as the Kabatiansky–Levenshtein bound. Elkies, Odlyzko, and Rush [15] improve the Minkowski–Hlawka lower bound for lattice packings of superballs. Fejes Tóth, Fodor, and Vígh [16] find upper bounds for congruent packings of n-dimensional regular cross polytopes when $n \geq 7$. How does the Cohn–Elkies bound behave asymptotically for translative superball packings?

With a generalization of the Cohn–Elkies bound one can also consider packings of congruent copies of a given body, but this is computationally even more challenging. This basic setup is explained in Oliveira and Vallentin [37] where they consider packings of congruent copies of regular pentagons in the Euclidean plane.

2 Sums of Hermitian Squares Invariant Under a Finite Group Generated by Pseudo-Reflections

Testing that a given real, multivariate polynomial is a sum of squares (SOS) is a fundamental computational problem in polynomial optimization and real algebraic geometry; see the recent book edited by Blekherman et al. [4].

Using the Gram matrix method this test can be reduced to the feasibility problem of semidefinite optimization: A real, multivariate polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ of degree $2d$ is an SOS if and only if there is a positive semidefinite matrix Q of size $\binom{n+d}{d} \times \binom{n+d}{d}$—a Gram matrix representation of p—so that

$$p(x_1, \ldots, x_n) = b(x_1, \ldots, x_n)^{\mathsf{T}} Q\, b(x_1, \ldots, x_n) \tag{8}$$

holds, where $b(x_1, \dots, x_n) \in \mathbb{R}[x_1, \dots, x_n]^{\binom{n+d}{d}}$ is a vector which contains a basis of the space of real polynomials up to degree d.

Gatermann and Parrilo [19] developed a general theory to simplify the matrices occurring in the Gram matrix method when the polynomial at hand is invariant under the action of a finite matrix group; see also Bachoc et al. [2].

In this section we work out this simplification for polynomials invariant under a finite group generated by pseudo-reflections. A pseudo-reflection is a linear transformation of $\mathbb{C}^n$ where precisely one eigenvalue is not equal to one. In particular, a reflection $x \mapsto x - \frac{2x \cdot v}{v \cdot v} v$ at a linear hyperplane orthogonal to a vector v is a pseudo-reflection.

In this case the computations required to apply the general theory of Gatermann and Parrilo can be done rather concretely on the basis of the theory developed by Shephard and Todd, Chevalley, and Serre (see for example the book by Humphreys [28], the survey by Stanley [44], or the book by Sturmfels [47]).

However, we deviate from the path set out by Gatermann and Parrilo in one important detail. Gatermann and Parrilo consider polynomials over the field of real numbers. When working with finite groups generated by pseudo-reflections it is more natural to work in the framework of Hermitian symmetric polynomials since we will use the Peter–Weyl theorem, the decomposition of the regular representation into irreducible unitary representations.

A polynomial $p \in \mathbb{C}[z_1, \dots, z_n, \overline{w}_1, \dots, \overline{w}_n] = \mathbb{C}[z, \overline{w}]$ is called a *Hermitian symmetric polynomial* if one of the following three equivalent conditions holds (see D'Angelo [13]):

- (i) The equality $p(z, \overline{w}) = \overline{p(w, \overline{z})}$ holds for all $z, w \in \mathbb{C}^n$.
- (ii) The function $z \mapsto p(z, \overline{z})$, with $z \in \mathbb{C}^n$, is real-valued.
- (iii) There is a Hermitian matrix $Q = (q_{\alpha\beta})$ so that one can represent p as $p(z, \overline{w}) = \sum_{\alpha,\beta} q_{\alpha\beta} z^\alpha \overline{w}^\beta$.

A Hermitian symmetric polynomial $p \in \mathbb{C}[z, \overline{w}]$ is a *sum of Hermitian squares* if there are polynomials $q_1, \dots, q_r \in \mathbb{C}[z]$ so that

$$p(z, \overline{w}) = \sum_{i=1}^{r} q_i(z) \overline{q_i(w)}$$

holds. In particular, a sum of Hermitian squares determines a real-valued nonnegative function by $z \mapsto p(z, \overline{z}) = \sum_{i=1}^{r} |q_i(z)|^2$. In fact, D'Angelo gave in [12, Defn. IV.5.1] eight positivity conditions for a Hermitian symmetric polynomial. He noted that being a sum of Hermitian squares is the strongest among them and that this condition is also easy to verify by the Gram matrix method after an obvious adaptation of (8):

$$p(z, \overline{w}) = b(z)^{\mathsf{T}} Q \overline{b(w)},$$

where Q is a Hermitian positive semidefinite matrix of size $\binom{n+d}{d} \times \binom{n+d}{d}$ and where $b(z) \in \mathbb{C}[z]^{\binom{n+d}{d}}$ is a vector which contains a basis of the space of complex polynomials up to degree d.

Now let us review the relevant theory of pseudo-reflection groups: Let $G \subseteq \mathrm{GL}_n(\mathbb{C})$ be a finite group generated by pseudo-reflections. It is acting on the polynomial ring $\mathbb{C}[x_1, \ldots, x_n] = \mathbb{C}[x]$ by

$$(gp)(x) = p(g^{-1}x) \quad \text{for } g \in G \text{ and } p \in \mathbb{C}[x].$$

The invariant ring is defined by

$$\mathbb{C}[x]^G = \{p \in \mathbb{C}[x] : gp = p \text{ for all } g \in G\}.$$

The invariant ring is generated by n homogeneous polynomials $\theta_1, \ldots, \theta_n$ which are algebraically independent. Thus,

$$\mathbb{C}[x]^G = \mathbb{C}[\theta_1, \ldots, \theta_n]$$

is a free algebra. Homogeneous, algebraically independent generators of the invariant ring are called basic invariants. They are not uniquely determined by the group, but their degrees $d_1, \ldots, d_n$ are.

The group action respects the grading of the polynomial ring. To determine the dimensions of the invariant subspaces of homogeneous polynomials

$$\mathrm{Hom}_k^G = \mathbb{C}[x]^G \cap \mathrm{Hom}_k$$

with

$$\mathrm{Hom}_k = \big\{p \in \mathbb{C}[x] : p(\alpha x) = \alpha^k p(x) \text{ for all } \alpha \in \mathbb{C}, \deg p = k\big\},$$

one can use Molien's series

$$\sum_{k=0}^{\infty} \dim \mathrm{Hom}_k^G \, t^k = \left(\prod_{i=1}^{n} (1 - t^{d_i})\right)^{-1}.$$

The coinvariant algebra is

$$\mathbb{C}[x]_G = \mathbb{C}[x]/I,$$

where $I = (\theta_1, \ldots, \theta_n)$ is the ideal generated by basic invariants. The coinvariant algebra is a graded algebra of finite dimension $|G|$. The dimensions of the homogeneous subspaces of $\mathbb{C}[x]_G$ are given by the Poincaré series

$$(1 - t)^{-n} \prod_{i=1}^{n} (1 - t^{d_i}).$$

In particular,

$$\mathbb{C}[x] = \mathbb{C}[x]^G \otimes \mathbb{C}[x]_G$$

holds. The action of G on the coinvariant algebra $\mathbb{C}[x]_G$ is equivalent to the regular representation of G. Let $\widehat{G}$ be the set of irreducible unitary representations of G up to equivalence. Then one can apply the Peter–Weyl theorem, see for example [48, Chap. 15]: There are homogeneous polynomials

$$\varphi_{ij}^{\pi}, \quad \text{with } \pi \in \widehat{G},\ 1 \leq i, j \leq d_\pi,$$

where d_π is the degree of π, which form a basis of the coinvariant algebra such that the transformation law

$$g\varphi_{ij}^{\pi} = (\pi(g)_j)^{\mathsf{T}} \begin{pmatrix} \varphi_{i1}^{\pi} \\ \vdots \\ \varphi_{id_\pi}^{\pi} \end{pmatrix}, \ i = 1, \ldots, d_\pi, \tag{9}$$

holds for all $g \in G$. Here, $\pi(g)_j$ denotes the j-th column of the unitary matrix $\pi(g) \in \mathrm{U}(d_\pi)$.

We extend the action of G from $\mathbb{C}[x]$ to $\mathbb{C}[z, \overline{w}]$ by

$$(gp)(z, \overline{w}) = p(g^{-1}z, \overline{g^{-1}w}) \quad \text{for } g \in G \text{ and } p \in \mathbb{C}[z, \overline{w}].$$

We define the ring of G-invariant Hermitian symmetric polynomials by

$$\mathbb{C}[z, \overline{w}]^G = \{p \in \mathbb{C}[z, \overline{w}] : gp = p \text{ for all } g \in G\}.$$

Now we set up all necessary notation for formulating the theorem which gives an explicit parametrization of the convex cone of G-invariant Hermitian symmetric polynomials which are Hermitian sum of squares. The following theorem can be derived from the real version of [19, Thm. 6.2]. So we omit the proof.

Theorem 2.1 *Let $G \subseteq \mathrm{GL}_n(\mathbb{C})$ be a finite group generated by pseudo-reflections. The convex cone of G-invariant Hermitian symmetric polynomials which can be written as sums of Hermitian squares equals*

$$\Big\{p \in \mathbb{C}[z, \overline{w}]^G : \ p(z, \overline{w}) = \sum_{\pi \in \widehat{G}} \langle P^{\pi}(z, \overline{w}), Q^{\pi}(z, \overline{w})\rangle, \\ P^{\pi}(z, \overline{w}) \text{ is a Hermitian SOS matrix polynomial in } \theta_i \Big\}.$$

Here $\langle A, B\rangle = \mathrm{Tr}(B^*A)$ denotes the trace inner product, the matrix $P^{\pi}(z, \overline{w})$ is a Hermitian SOS matrix polynomial in the variables $\theta_1, \ldots, \theta_n$, i.e. there is a matrix $L^{\pi}(z)$ with entries in $\mathbb{C}[z]^G = \mathbb{C}[\theta_1, \ldots, \theta_n]$ such that

$$P^{\pi}(z, \overline{w}) = L^{\pi}(z)\overline{L^{\pi}(w)}^{\mathsf{T}}$$

holds, and $Q^\pi(z,\overline{w}) \in (\mathbb{C}[z,\overline{w}]^G)^{d_\pi \times d_\pi}$ *is defined componentwise by*

$$[Q^\pi]_{kl}(z,\overline{w}) = \sum_{i=1}^{d_\pi} \varphi^\pi_{ki}(z)\overline{\varphi^\pi_{li}(w)}.$$

The computational value of this approach is that one only has to determine basic invariants $\theta_1,\dots,\theta_n$ and a suitable basis φ^π_{ij} of the coinvariant algebra which satisfies (9). These computations are *independent* of the degree of the polynomial p.

It turns out that for the octahedral group B_3 we consider for our application all irreducible unitary representations are orthogonal representations. In this case the previous theorem can be translated into the following version for the field of real numbers.

Theorem 2.2 *Let* $G \subseteq \mathrm{GL}_n(\mathbb{R})$ *be a finite group generated by pseudo-reflections so that all unitary irreducible representations* $\pi \in \widehat{G}$ *of* G *are orthogonal. The convex cone of* G*-invariant real polynomials which can be written as sums of squares equals*

$$\Big\{ p \in \mathbb{R}[x]^G : \ p(x) = \sum_{\pi \in \widehat{G}} \langle P^\pi(x), Q^\pi(x)\rangle, \quad P^\pi(x) \text{ is an SOS matrix polynomial in } \theta_i \Big\}.$$

Here the matrix $P^\pi(x)$ *is an SOS matrix polynomial in the variables* $\theta_1,\dots,\theta_n$*, i.e. there is a matrix* $L^\pi(x)$ *with entries in* $\mathbb{R}[x]^G = \mathbb{R}[\theta_1,\dots,\theta_n]$ *such that*

$$P^\pi(x) = L^\pi(x) L^\pi(x)^{\mathsf{T}}$$

holds, and $Q^\pi(x) \in (\mathbb{R}[x]^G)^{d_\pi \times d_\pi}$ *is defined componentwise by*

$$[Q^\pi]_{kl}(x) = \sum_{i=1}^{d_\pi} \varphi^\pi_{ki}(x)\varphi^\pi_{li}(x).$$

3 Real Sums of Squares Polynomials Invariant Under the Octahedral Group

In this section we specialize Theorem 2.2 to the symmetry group of the three-dimensional real octahedron, the octahedral group, which is the finite reflection group B_3 generated by the matrices (1). Since in the literature only very few cases of Theorems 2.1 or 2.2 are worked out explicitly, we give substantial amount of detail here.

We use the basic invariants $\theta_1, \theta_2, \theta_3$ as given in (3). Let χ_π be the character of the irreducible representations $\pi \in \widehat{\mathsf{B}_3}$. There are ten inequivalent irreducible unitary representations and the character table, which one computes with a computer algebra system or which one also can find in many text books on mathematical chemistry, is given in Table 3.

In the character table we use Mulliken symbols for concreteness. This scheme was suggested by Robert S. Mulliken, Nobel laureate in Chemistry in 1966. The symmetry

Table 3 Character table of B_3

	E	i	$3C_2$	$3\sigma_h$	$6C_2'$	$6\sigma_d$	$8C_3$	$6C_4$	$6S_4$	$8S_6$
A_{1g}	1	1	1	1	1	1	1	1	1	1
A_{1u}	1	−1	1	−1	1	−1	1	1	−1	−1
A_{2g}	1	1	1	1	−1	−1	1	−1	−1	1
A_{2u}	1	−1	1	−1	−1	1	1	−1	1	−1
E_g	2	2	2	2	0	0	−1	0	0	−1
E_u	2	−2	2	−2	0	0	−1	0	0	1
T_{1g}	3	3	−1	−1	−1	−1	0	1	1	0
T_{1u}	3	−3	−1	1	−1	1	0	1	−1	0
T_{2g}	3	3	−1	−1	1	1	0	−1	−1	0
T_{2u}	3	−3	−1	1	1	−1	0	−1	1	0

Rows are indexed by characters, and columns are indexed by conjugacy classes

group of the regular three-dimensional octahedron coincides with the one of the regular cube. In the following we describe the conjugacy classes of B_3 geometrically by looking at the symmetries of the cube: E is the identity of the group (E from German *Einheit*), i is the inverse operation $i(x) = -x$, $3C_2$ are the three clockwise rotations by $180°$ through the axis of the facet centers, $3\sigma_h$ are the three reflections through planes which are parallel to pairs of facets (σ from *Spiegelung*), $6C_2'$ are the six clockwise rotations by $180°$ through the axis of the edge centers, $6\sigma_d$ are the six reflections through the planes given by the diagonals of the facets, $8C_3$ are the eight clockwise rotations by $120°$ through the diagonals of the cube, $6C_4$ are the six clockwise rotations by $90°$ through the axis of the facet centers, $6S_4$ are the six rotation-reflections by $90°$ through the axis of the facet centers, and $8S_6$ are the eight rotation-reflections by $60°$ through the diagonals of the cube. One-dimensional characters are given by the letter A, two-dimensional characters are specified by the letter E, and the three-dimensional ones by T. The subscript g (*gerade*) or u (*ungerade*) is used to distinguish between $\chi(i) = 1$ and $\chi(i) = -1$.

The Molien series of $\mathbb{C}[x_1, x_2, x_3]^{\mathsf{B}_3}$ is

$$\frac{1}{(1-t^2)(1-t^4)(1-t^6)} = 1 + t^2 + 2t^4 + 3t^6 + 4t^8 + 5t^{10} + 7t^{12} + 8t^{14} + 10t^{16} + 12t^{18} + \cdots$$

The coinvariant algebra $\mathbb{C}[x]_G = \mathbb{C}[x]/I$ with $I = (\theta_1, \theta_2, \theta_3)$ decomposes into

$$V = V_0 \oplus \cdots \oplus V_9$$

according to the grading by degree, where the dimensions of the spaces V_k, with $k = 0, \ldots, 9$ can be read off by the Poincaré series

$$1 + 3t + 5t^2 + 7t^3 + 8t^4 + 8t^5 + 7t^6 + 5t^7 + 3t^8 + t^9.$$

The group action respects the grading. It turns out that all irreducible unitary representations occur multiplicity-free in the V_k's and that all of them are orthogonal representations. Serre [43, Chap. 2.6, Thm. 8] gives a formula which can be used to decompose a finite-dimensional representation into its isotypic components. Consider the representation

$$\rho_k \colon \mathsf{B}_3 \to \mathrm{GL}(\mathrm{Hom}_k), \quad \rho_k(g)(p) \mapsto gp,$$

and consider a unitary irreducible representation $\pi \in \widehat{\mathsf{B}}_3$. Then the image of the linear map

$$p_k^\pi \colon \mathrm{Hom}_k \to \mathrm{Hom}_k, \quad p_k^\pi = \frac{d_\pi}{|\mathsf{B}_3|} \sum_{g \in \mathsf{B}_3} \chi_\pi(g^{-1}) \rho_k(g)$$

gives the subspace V_k^π of Hom_k which is the isotypic component of Hom_k having type π.

We choose the smallest degree k_π so that there is a nontrivial isotypic component of Hom_{k_π} having type π, and this choice of k_π implies that this isotypic component $V_{k_\pi}^\pi$ is actually an irreducible subspace. Then we equip this irreducible subspace with a B_3-invariant inner product and compute an orthonormal basis by Gram–Schmidt orthonormalization. This orthonormal basis gives polynomials φ_{1j}^π, with $j = 1, \dots, d_\pi$, which we need for applying Theorem 2.2. The results are displayed in Table 4.

The next task is to find the other polynomials φ_{ij}^π, with $i = 2, \dots, d_\pi$, which transform according to (9). We use the algorithm of Serre [43, Chap. 2.7, Prop. 8] for this.

Define

$$p_{k,ij}^\pi \colon \mathrm{Hom}_k \to \mathrm{Hom}_k \quad \text{by} \quad p_{k,ij}^\pi = \frac{d_\pi}{|\mathsf{B}_3|} \sum_{g \in \mathsf{B}_3} \pi_{ji}(g^{-1}) \rho_k(g),$$

where $\pi(g) \in \mathrm{U}(d_\pi)$ is the unitary matrix which we get by considering the matrix representation $\rho_{k_\pi}(g)$ restricted to the irreducible subspace of $V_{k_\pi}^\pi$ of Hom_{k_π} having type π and expressed in terms of the orthonormal basis φ_{1j}^π, with $j = 1, \dots, d_\pi$, we just computed. Denote the image $p_{k,ii}^\pi(V_k^\pi)$ by $V_{k,i}^\pi$, where $i = 1, \dots, d_\pi$. Then we have the decomposition

$$V_k^\pi = V_{k,1}^\pi \oplus \cdots \oplus V_{k,d_\pi}^\pi.$$

Consider a nonzero vector $\varphi_{k1}^\pi \in V_{k,1}^\pi$. Define $\varphi_{ki}^\pi = p_{k,i1}^\pi(\varphi_{k1}^\pi)$. Then,

$$\rho_k(g)(\varphi_{ki}^\pi) = \sum_{j=1}^{d_\pi} \pi_{ji}(g) \varphi_{kj}^\pi$$

Table 4 Orthonormal basis φ^{π}_{1j}, with $j = 1, \ldots, d_\pi$, of subspaces $V^{\pi}_{k_\pi}$

A_{1g}	1
A_{1u}	$x_1x_2x_3$
A_{2g}	$x_1^4x_2^2 - x_1^4x_3^2 - x_1^2x_2^4 + x_1^2x_3^4 + x_2^4x_3^2 - x_2^2x_3^4$
A_{2u}	$x_1^5x_2^3x_3 - x_1^5x_2x_3^3 - x_1^3x_2^5x_3 + x_1^3x_2x_3^5 + x_1x_2^5x_3^3 - x_1x_2^3x_3^5$
E_g	$x_1^3x_2x_3 - x_1x_2x_3^3$
	$\frac{\sqrt{3}}{3}x_1^3x_2x_3 - \frac{2\sqrt{3}}{3}x_1x_2^3x_3 + \frac{\sqrt{3}}{3}x_1x_2x_3^3$
E_u	$x_1^2 - x_3^2$
	$\frac{\sqrt{3}}{3}x_1^2 - \frac{2\sqrt{3}}{3}x_2^2 + \frac{\sqrt{3}}{3}x_3^2$
T_{1g}	$x_1^3x_2 - x_1x_2^3$
	$x_1^3x_3 - x_1x_3^3$
	$x_2^3x_3 - x_2x_3^3$
T_{1u}	$x_1^2x_2 - x_2x_3^2$
	$x_1^2x_3 - x_2^2x_3$
	$x_1x_2^2 - x_1x_3^2$
T_{2g}	x_1x_2
	x_1x_3
	x_2x_3
T_{2u}	x_1
	x_2
	x_3

holds for all $g \in G$, as we wanted. With this information we can construct the matrices Q^π. We give them in Table 5.

4 Computing the Fourier Transform

As explained in the introduction we define the function f which we want to use in Theorem 1.1 through its Fourier transform $\widehat{f}(u) = g(u)e^{-\pi\|u\|^2}$ where g is a polynomial. In order to verify the third condition of the theorem, we have to compute f from $\widehat{f}$. In other words, we have to compute the Fourier transform of $u \mapsto \widehat{f}(-u)$. In this section we explain how to do this. We first consider the general case, when g is an arbitrary complex polynomial in n variables. Then we show how some of the computations can be simplified when we assume that g is B_3-invariant. In the end, since the Fourier transform is linear, we have to solve a certain system of linear equations. A similar calculation was done by Dunkl [14]. He even gives explicit algebraic solutions.

Consider the following decomposition of complex polynomials in n variables of degree at most d:

Table 5 Matrices Q^{π} for the group B_3 given in upper triangular row-major order (in the consecutive order of row entries of upper triangular matrices)

A_{1g}	1
A_{1u}	$\theta_1^3 - 3\theta_1\theta_2 + 2\theta_3$
A_{2g}	$-\theta_1^6 + 9\theta_1^4\theta_2 - 8\theta_1^3\theta_3 - 21\theta_1^2\theta_2^2 + 36\theta_1\theta_2\theta_3 + 3\theta_2^3 - 18\theta_3^2$
A_{2u}	$-\theta_1^9 + 12\theta_1^7\theta_2 - 10\theta_1^6\theta_3 - 48\theta_1^5\theta_2^2 + 78\theta_1^4\theta_2\theta_3 + 66\theta_1^3\theta_2^3 - 34\theta_1^3\theta_3^2 - 150\theta_1^2\theta_2^2\theta_3$ $-9\theta_1\theta_2^4 + 126\theta_1\theta_2\theta_3^2 + 6\theta_2^3\theta_3 - 36\theta_3^3$
E_g	$-2\theta_1^5 + 12\theta_1^3\theta_2 - 4\theta_1^2\theta_3 - 18\theta_1\theta_2^2 + 12\theta_2\theta_3$
	$-2\theta_1^4\theta_2 + 6\theta_1^3\theta_3 + 6\theta_1^2\theta_2^2 - 22\theta_1\theta_2\theta_3 + 12\theta_3^2$
	$\theta_1^7 - 9\theta_1^5\theta_2 + 10\theta_1^4\theta_3 + 19\theta_1^3\theta_2^2 - 36\theta_1^2\theta_2\theta_3 - 3\theta_1\theta_2^3 + 16\theta_1\theta_3^2 + 2\theta_2^2\theta_3$
E_u	$-2\theta_1^2 + 6\theta_2$
	$-2\theta_1\theta_2 + 6\theta_3$
	$\theta_1^4 - 6\theta_1^2\theta_2 + 8\theta_1\theta_3 + \theta_2^2$
T_{1g}	$12\theta_1\theta_3 - 12\theta_2^2$
	$2\theta_1^5 - 12\theta_1^3\theta_2 + 16\theta_1^2\theta_3 + 6\theta_1\theta_2^2 - 12\theta_2\theta_3$
	$2\theta_1^6 - 12\theta_1^4\theta_2 + 10\theta_1^3\theta_3 + 12\theta_1^2\theta_2^2 - 6\theta_1\theta_2\theta_3 - 6\theta_2^3$
	$2\theta_1^6 - 10\theta_1^4\theta_2 + 10\theta_1^3\theta_3 + 10\theta_1\theta_2\theta_3 - 12\theta_3^2$
	$\theta_1^7 - 3\theta_1^5\theta_2 + 2\theta_1^4\theta_3 - 9\theta_1^3\theta_2^2 + 24\theta_1^2\theta_2\theta_3 + 3\theta_1\theta_2^3 - 12\theta_1\theta_3^2 - 6\theta_2^2\theta_3$
	$4\theta_1^6\theta_2 - 3\theta_1^5\theta_3 - 21\theta_1^4\theta_2^2 + 32\theta_1^3\theta_2\theta_3 + 12\theta_1^2\theta_2^3 - 12\theta_1^2\theta_3^2 - 9\theta_1\theta_2^2\theta_3 - 3\theta_2^4$
T_{1u}	$-12\theta_1^3 + 48\theta_1\theta_2 - 36\theta_3$
	$-6\theta_1^4 + 24\theta_1^2\theta_2 - 12\theta_1\theta_3 - 6\theta_2^2$
	$-6\theta_1^3\theta_2 + 6\theta_1^2\theta_3 + 18\theta_1\theta_2^2 - 18\theta_2\theta_3$
	$-2\theta_1^5 + 6\theta_1^3\theta_2 + 2\theta_1^2\theta_3 - 6\theta_2\theta_3$
	$\theta_1^6 - 9\theta_1^4\theta_2 + 8\theta_1^3\theta_3 + 15\theta_1^2\theta_2^2 - 12\theta_1\theta_2\theta_3 - 3\theta_2^3$
	$\theta_1^7 - 6\theta_1^5\theta_2 + 5\theta_1^4\theta_3 + 3\theta_1^3\theta_2^2 + 6\theta_1\theta_2^3 - 9\theta_2^2\theta_3$
T_{2g}	$3\theta_1^2 - 3\theta_2$
	$6\theta_1\theta_2 - 6\theta_3$
	$-\theta_1^4 + 6\theta_1^2\theta_2 - 2\theta_1\theta_3 - 3\theta_2^2$
	$-2\theta_1^4 + 12\theta_1^2\theta_2 - 10\theta_1\theta_3$
	$-\theta_1^5 + 4\theta_1^3\theta_2 - 2\theta_1^2\theta_3 + 3\theta_1\theta_2^2 - 4\theta_2\theta_3$
	$-2\theta_1^4\theta_2 + \theta_1^3\theta_3 + 9\theta_1^2\theta_2^2 - 7\theta_1\theta_2\theta_3 - 3\theta_2^3 + 2\theta_3^2$
T_{2u}	$6\theta_1$
	$6\theta_2$
	$6\theta_3$
	$6\theta_3$
	$\theta_1^4 - 6\theta_1^2\theta_2 + 8\theta_1\theta_3 + 3\theta_2^2$
	$\theta_1^5 - 5\theta_1^3\theta_2 + 5\theta_1^2\theta_3 + 5\theta_2\theta_3$

$$\mathbb{C}[x]_{\leq d} = \bigoplus_{j=0}^{d} \mathrm{Hom}_j = \bigoplus_{j=0}^{d} \bigoplus_{\substack{r,k \\ 2r+k=j}} \|x\|^{2r}\, \mathrm{Harm}_k, \tag{10}$$

where

$$\mathrm{Harm}_k = \Big\{h \in \mathrm{Hom}_k : \Delta h = \Big(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\Big)h = 0\Big\}$$

is the space of (homogeneous) harmonic polynomials of degree k. In other words, harmonic polynomials of degree k are the kernel of the Laplace operator

$$\mathrm{Harm}_k = \ker \Delta, \quad \Delta\colon\ \mathrm{Hom}_k \to \mathrm{Hom}_{k-2},$$

where

$$\dim \mathrm{Harm}_k = \dim \mathrm{Hom}_k - \dim \mathrm{Hom}_{k-2} = \binom{n+k-1}{k} - \binom{n+k-3}{k-2}.$$

The existence of decomposition (10) is classical; one can find a proof, for example, in the book by Stein and Weiss [45, Thm. IV.2.10]. Decomposition (10) together with the following proposition shows how to compute f from g by solving a system of linear equations. The proposition in particular shows that the function $x \mapsto h_k(x)e^{-\pi\|x\|^2}$ with $h_k \in \mathrm{Harm}_k$ is an eigenfunction of the Fourier transform with eigenvalue i^{-k}.

Proposition 4.1 *Let*

$$f(x) = h_k(x)\|x\|^{2r}e^{-\pi\|x\|^2}$$

be a Schwartz function with $h_k \in \mathrm{Harm}_k$*. The Fourier transform of* f *is*

$$\widehat{f}(u) = (i^{-k}h_k(u)) \cdot \pi^{-r} r! L_r^{n/2+k-1}(\pi\|u\|^2)e^{-\pi\|u\|^2},$$

where $L_r^{n/2+k-1}$ *is the Laguerre polynomial of degree* r *with parameter* $n/2+k-1$.

In general, Laguerre polynomials L_r^{α} with parameter α are orthogonal polynomials for the inner product $\int_0^{\infty} f(x)g(x)x^{\alpha}e^{-x}\,dx$, see the book by Andrews et al. [1] for more details.

Proof Using Stein and Weiss [45, Thm. IV.3.10] one sees that the Fourier transform of f is

$$\widehat{f}(u) = (i^{-k}h_k(u)) \cdot 2\pi\|u\|^{-(n/2-1+k)} \cdot \int_0^{\infty} s^{n/2+2r+k} J_{n/2-1+k}(2\pi s\|u\|)e^{-\pi s^2}\,ds,$$

where J_α is the Bessel function of the first kind of order α. By Andrews, Askey, and Roy [1, Coro. 4.11.8] the integral above equals

$$\frac{\Gamma(n/2+r+k)(\sqrt{\pi}\|u\|)^{n/2-1+k}e^{-\pi\|u\|^2}}{2(\sqrt{\pi})^{n/2+2r+k+1}\Gamma(n/2+k)}\ {}_1F_1\left(\begin{matrix}-r\\ n/2+k\end{matrix};\pi\|u\|^2\right),$$

where ${}_1F_1$ denotes the hypergeometric series. Hence,

$$\widehat{f}(u)=(i^{-k}h_k(u))\cdot\pi^{-r}\frac{\Gamma(n/2+r+k)}{\Gamma(n/2+k)}\ {}_1F_1\left(\begin{matrix}-r\\ n/2+k\end{matrix};\pi\|u\|^2\right)e^{-\pi\|u\|^2}.$$

The hypergeometric series becomes a Laguerre polynomial ([1, (6.2.2)])

$${}_1F_1\left(\begin{matrix}-r\\ n/2+k\end{matrix};\pi\|u\|^2\right)=\frac{r!}{(n/2+k)_r}\,L_r^{n/2+k-1}(\pi\|u\|^2).$$

Combining the last two equations gives the desired result. □

If one assumes that the polynomial g is B_3-invariant one can save quite some computations. Instead of working with decomposition (10) we can work with a B_3-invariant decomposition because the Laplacian Δ commutes with the action of the orthogonal group:

$$\mathbb{C}[x]_{\le d}^{\mathsf{B}_3}=\bigoplus_{j=0}^{d}\mathrm{Hom}_j^{\mathsf{B}_3}=\bigoplus_{j=0}^{d}\ \bigoplus_{\substack{r,k\\2r+k=j}}\theta_1^r\,\mathrm{Harm}_k^{\mathsf{B}_3}.$$

To see the computational advantage, let us compare the dimensions of the harmonic subspaces. We generally have $\dim\mathrm{Harm}_k=2k+1$ when $n=3$, but the Molien series counting the dimensions of the invariant harmonic subspaces (see Goethals and Seidel [21]) is

$$\begin{aligned}\sum_{t=0}^{\infty}\dim\mathrm{Harm}_k^{\mathsf{B}_3}t^k&=\frac{1}{(1-t^4)(1-t^6)}\\&=1+t^4+t^6+t^8+t^{10}+2t^{12}+t^{14}+2t^{16}+2t^{18}+\cdots\end{aligned}$$

5 Semidefinite Programming Formulation

We now present in detail the semidefinite program we use to find good functions f satisfying the conditions of Theorem 1.1 when $\mathcal{K}\subseteq\mathbb{R}^3$ is such that the Minkowski difference $\mathcal{K}-\mathcal{K}$ is invariant under the action of B_3. This is the case, e.g., when $\mathcal{K}$ is a three-dimensional superball or a Platonic or an Archimedean solid with tetrahedral symmetry.

5.1 Representation of the Function f via Its Fourier Transform

Recall that we specify the function $f\colon \mathbb{R}^3 \to \mathbb{R}$ via its Fourier transform. Given a real polynomial $g \in \mathbb{R}[x] = \mathbb{R}[x_1, x_2, x_3]$, we define

$$\widehat{f}(u) = g(u)e^{-\pi \|u\|^2}. \tag{11}$$

We deal exclusively with B_3-invariant functions, so we take the polynomial g above B_3-invariant. Functions invariant under B_3 are even, and so are their Fourier transforms. From this it follows that there is no loss of generality in considering real-valued Fourier transforms and so there is also no loss of generality in requiring that g be a real polynomial. This simplifies the use of semidefinite programming considerably since we only have to optimize over the cone of real positive semidefinite matrices and not over the larger cone of Hermitian positive semidefinite matrices.

The function f is of positive type if and only if g is a nonnegative polynomial. Since it is computationally difficult to work with nonnegative polynomials, we require instead that g be a sum of squares, thus restricting [see for example the Robinson polynomial (4)] the set of functions f that we work with.

Theorem 2.2 provides a parametrization of the cone of SOS polynomials invariant under B_3, like g. In the theorem, each matrix P^π is an SOS matrix polynomial, that is, there is a matrix L^π whose entries are invariant polynomials such that $P^\pi = L^\pi(L^\pi)^{\mathsf{T}}$. To find an SOS polynomial g, we may then fix the maximum degree a polynomial in P^π can have, and use the fact (cf. Gatermann and Parrilo [19, Defn. 2.2]) that $S \in \mathbb{R}[x]^{n\times n}$ is an SOS matrix if and only if the polynomial $y^{\mathsf{T}} S y \in \mathbb{R}[x, y]$ is a sum of squares, where $y = (y_1, \ldots, y_n)$ are new variables. This, together with (8), suggests a way to represent g with one positive semidefinite matrix for each of the irreducible representations of B_3.

In our formulation we use a derived parametrization that produces numerically stabler problems providing bounds that can be rigorously shown to be correct. Our approach is as follows. For each irreducible unitary representation $\pi \in \widehat{\mathsf{B}}_3$, let $\Phi^\pi = \big(\varphi^\pi_{ij}\big)_{i,j=1}^{d_\pi}$ where the φ^π_{ij}'s were defined in Sects. 2 and 3. Then $Q^\pi = \Phi^\pi(\Phi^\pi)^{\mathsf{T}}$. Each row of Φ^π contains homogeneous invariant polynomials all of the same degree; we say the *degree* of a row is the degree of the polynomials in it.

Let $\mathcal{A}$ be some basis of $\mathbb{R}[x]^{\mathsf{B}_3}$ consisting of homogeneous polynomials. For an integer $t \geq 0$ and each $\pi \in \widehat{\mathsf{B}}_3$, let $\mathcal{I}^t_\pi$ be the set of pairs (a, r), where $a \in \mathcal{A}$ and $1 \leq r \leq d_\pi$ indexes a row of Φ^π such that the degree of a plus the degree of the row r is at most t. For each $\pi \in \widehat{\mathsf{B}}_3$ we may then consider the matrix $V^{\pi,t}$ with rows and columns indexed by $\mathcal{I}^t_\pi$ such that

$$V^{\pi,t}_{(a,r),(b,s)} = abQ^\pi_{rs}.$$

Notice that the entry $((a, r), (b, s))$ of $V^{\pi,t}$ has degree equal to $\deg a + \deg b + \deg Q^\pi_{rs} \leq 2t$.

In our formulation we will fix an odd[1] positive integer d and let

$$g(x) = \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle V^{\pi,d}(x), R^{\pi} \rangle \tag{12}$$

be the polynomial that defines $\widehat{f}$, where R^{π} are positive semidefinite matrices of the appropriate sizes. Notice that this, together with the construction of the $V^{\pi,d}$ matrices, implies that g is a sum of squares polynomial of degree at most $2d$ invariant under B_3, and that vice versa all sum of squares B_3-invariant polynomials of degree at most $2d$ are of this form.

The function f is the Fourier inverse of $\widehat{f}$. In Sect. 4, we have seen how the inverse can be computed when $\widehat{f}$ is given by an invariant polynomial as in (11). In fact, there is a linear transformation $\mathcal{F}\colon \mathbb{R}[x]^{\mathsf{B}_3} \to \mathbb{R}[x]^{\mathsf{B}_3}$ such that

$$f(x) = \mathcal{F}[g](x)e^{-\pi\|x\|^2}.$$

In particular, if g is given as in (12), then

$$f(x) = e^{-\pi\|x\|^2} \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](x), R^{\pi} \rangle,$$

where by applying $\mathcal{F}$ to a matrix we apply it to each entry and get a matrix as a result.

With this, we can easily see how to express condition (i) of Theorem 1.1. It becomes

$$\sum_{\pi \in \widehat{\mathsf{B}}_3} \langle V^{\pi,d}(0), R^{\pi} \rangle \geq 1.$$

The bound provided by the theorem is then vol $\mathcal{K}$ times

$$f(0) = \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](0), R^{\pi} \rangle;$$

this will be the objective function of our semidefinite program.

5.2 Nonpositivity Constraint

We impose the condition $f(x) \leq 0$ when $x \notin \mathcal{K}^{\circ} - \mathcal{K}^{\circ}$ in two steps by breaking the domain in which the function has to be nonpositive in two parts, an unbounded and a bounded one. We then deal with the unbounded part with an SOS constraint and with the bounded part via sampling.

[1] The reason why we pick odd d is so that the resulting problem admits a strictly feasible solution. This will be better explained in Sect. 6.1.

Let us consider first the unbounded part of the domain. Let s be a B_3-invariant polynomial such that

$$\mathcal{K}^\circ - \mathcal{K}^\circ \subseteq \{\, x \in \mathbb{R}^3 : s(x) < 0 \,\},$$

where the set on the right-hand side is bounded. For instance, if δ is the maximum norm of a vector in $\mathcal{K} - \mathcal{K}$, then we may take $s(x) = \|x\|^2 - \delta^2$.

If there are SOS polynomials q_1 and q_2 such that

$$\mathcal{F}[g](x) = -s(x)q_1(x) - q_2(x), \tag{13}$$

then f will be nonpositive in $\{\, x \in \mathbb{R}^3 : s(x) \geq 0 \,\}$. Now, g is invariant and hence $\mathcal{F}[g]$ is invariant. Since s is also invariant, we may take both q_1 and q_2 invariant without loss of generality. So we may use for q_1 and q_2 a parametrization similar to the one we used for g, but here it is important to be careful with the choice of degrees of q_1 and q_2.

In principle, the degrees of q_1 and q_2 can be anything as long as they are high enough so that the identity above may hold. In practice, it is a good idea to limit the degrees of q_1 and q_2 as much as possible. For instance, if q_2 is allowed to have a larger degree than $\mathcal{F}[g]$, then it is certainly not possible to represent it in our parametrization with positive *definite* matrices, and this will make it very difficult to rigorously prove that the numbers we obtain are indeed bounds.

We fixed the degree of g to be at most $2d$ for some odd d. Then $\mathcal{F}[g]$ also has degree at most $2d$. Since s is invariant, it has an even degree, say $2d_s$. We will impose $\deg q_1 \leq 2d - 2d_s$ and $\deg q_2 \leq 2d$. Now we may parametrize q_1 using positive semidefinite matrices S_1^π for $\pi \in \widehat{\mathsf{B}}_3$ and q_2 using positive semidefinite matrices S_2^π for $\pi \in \widehat{\mathsf{B}}_3$, rewriting (13) as

$$\sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](x), R^\pi \rangle + \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle s(x)V^{\pi,d-d_s}(x), S_1^\pi \rangle + \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle V^{\pi,d}(x), S_2^\pi \rangle = 0.$$

Notice this is a polynomial identity, which should be translated into linear constraints in our semidefinite program. To do so we need to express all polynomials in a common basis. A natural choice here is a basis of the invariant ring $\mathbb{R}[x]^{\mathsf{B}_3}$, since we work exclusively with invariant polynomials.

Now we still need to ensure that f is nonpositive in the bounded set

$$\mathcal{D} = \{\, x \in \mathbb{R}^3 : s(x) < 0 \,\} \setminus (\mathcal{K}^\circ - \mathcal{K}^\circ).$$

We do so by using a finite sample of points in $\mathcal{D}$ and adding linear constraints requiring that $\mathcal{F}[g]$ be nonpositive for each point in the sample. The idea is that, if we select enough points, then these constraints should ensure that f is nonpositive everywhere in $\mathcal{D}$.

So we choose a finite set $\mathcal{S} \subseteq \mathcal{D}$. Because of the invariance of g, and hence of $\mathcal{F}[g]$, we may restrict ourselves to points in the fundamental domain of B_3 or, in other words, we may restrict ourselves to points (x_1, x_2, x_3) with $0 \le x_1 \le x_2 \le x_3$. Then we add the constraints

$$\sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](x), R^\pi \rangle \le 0 \qquad \text{for all } x \in \mathcal{S}$$

to our problem.

5.3 Full Formulation

Here is the semidefinite programming problem we solve. Recall that d is an odd positive integer.

$$\begin{aligned} \min \sum_{\pi \in \widehat{\mathsf{B}}_3} & \langle \mathcal{F}[V^{\pi,d}](0), R^\pi \rangle \\ \text{(a)} \quad & \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle V^{\pi,d}(0), R^\pi \rangle \ge 1, \\ \text{(b)} \quad & \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](x), R^\pi \rangle + \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle s(x) V^{\pi,d-d_s}(x), S_1^\pi \rangle \\ & \qquad + \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle V^{\pi,d}(x), S_2^\pi \rangle = 0, \\ \text{(c)} \quad & \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](x), R^\pi \rangle \le 0 \qquad \text{for all } x \in \mathcal{S}, \\ & R^\pi, S_1^\pi, \text{ and } S_2^\pi \text{ are positive semidefinite.} \end{aligned} \tag{14}$$

As mentioned before, to express the SOS constraint above, which is in fact a polynomial identity, we need to express all polynomials involved in a given common basis. We use for this a basis of $\mathbb{R}[x]^{\mathsf{B}_3}$.

6 Rigorous Verification

In this section we discuss how the numerical results obtained can be turned into rigorous bounds. For the remainder of this section, $\mathcal{K} \subseteq \mathbb{R}^3$ will be a convex body such that $\mathcal{K} - \mathcal{K}$ is B_3-invariant.

6.1 Solving the Problem and Checking the SOS Constraint

We input problem (14) to a semidefinite programming solver. In doing so, we are using floating-point numbers to represent the input data. Solvers also use floating-point numbers in their numerical calculations, so the solution obtained at the end is likely not feasible. *But*, if it is close enough to being feasible, then it can be turned into a feasible solution. To this end it is important to find a solution in which the minimum eigenvalue of any matrix R^π, S_1^π and S_2^π is much larger than the maximum violation of any constraint.

Here is where it becomes important to make the formulation as tight as possible, for instance by picking the degrees of polynomials g, q_1, and q_2 in (13) correctly, so that (14) admits a *strictly feasible* solution, that is, a solution in which every matrix is positive definite. It is also for this reason that we have chosen d odd, since for even d the resulting problem is not strictly feasible.

To obtain such a solution we use a two-step approach. First we solve our problem to get an estimate z^* on the optimal value. Many interior point solvers work exclusively with positive definite solutions, but at the end round the solution to a face of the positive semidefinite cone. So the resulting solution matrices might have zero eigenvalues. To overcome this problem, we then pick some small error η (we usually pick something like $\eta = 10^{-5}$) and remove the objective function of the problem, adding it as a constraint like

$$\sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](0), R^\pi \rangle \le z^* + \eta.$$

So we sacrifice a bit of the optimal value. Most solvers, however, when dealing with feasibility problems, i.e., problems without an objective function, return a solution in the *analytic center* if a solution is found, and that solution will have positive definite matrices with large minimum eigenvalues. Of course, how large the minimum eigenvalues will be depends on the choice of η.

It is also important to use a solver able to work with high-precision floating-point numbers. Solvers working with double-precision floating-point arithmetic have failed to find feasible solutions of our problem because of numerical stability issues. Moreover, by using high-precision arithmetic we will get in the end a solution that is only slightly violated, which is our goal. To solve our problems, we used the SDPA-GMP solver [17].

Say then we have a solution $(R^\pi, S_1^\pi, S_2^\pi)$ with the desired property, that is, a solution in which the minimum eigenvalues are much larger than the maximum constraint violation. Our next step is to get a bound on the minimum eigenvalue of each matrix involved. We do it as follows. For each matrix A in the solution, we use binary search to find $\lambda_A > 0$ close to the minimum eigenvalue of A so that $A - \lambda_A I$ has a Cholesky decomposition LL^T. This we do with high-precision floating-point arithmetic. Then we use instead of A the matrix $\tilde{A} = LL^\mathsf{T} + \lambda_A I$. We have then a positive definite matrix and a bound on its minimum eigenvalue. We use interval arithmetic with high-precision floating-point arithmetic [39] to represent the new solution $(\tilde{R}^\pi, \tilde{S}_1^\pi, \tilde{S}_2^\pi)$ obtained in this way.

Now we can easily compute how violated the normalization constraint (a) of (14) is using interval arithmetic; if it is violated then we can multiply the solution by a positive number so as to have it satisfied. It is also easy to compute the objective value of the solution. We can also use interval arithmetic to compute an upper bound on the maximum violation of the SOS constraint (b) in (14). To do so, we compute the absolute value of the coefficient of

$$r(x) = \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle \mathcal{F}[V^{\pi,d}](x), \tilde{R}^\pi \rangle + \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle s(x) V^{\pi,d-d_s}(x), \tilde{S}_1^\pi \rangle + \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle V^{\pi,d}(x), \tilde{S}_2^\pi \rangle$$

with largest absolute value. Here we should note that matrices $V^{\pi,m}$ can be expressed using only rationals (if we work with a basis of $\mathbb{R}[x]^{\mathsf{B}_3}$ whose elements have only rational coefficients, as we actually do), and these rationals can be approximated with interval arithmetic. For the matrices $\mathcal{F}[V^{\pi,d}]$ of Fourier inverses we also need irrationals (namely, powers of π), but these can be approximated with interval arithmetic.

We want to change $(\tilde{R}^\pi, \tilde{S}_1^\pi, \tilde{S}_2^\pi)$ in order to make r identically zero. Notice r is invariant and has degree up to $2d$. By construction of the $V^{\pi,d}$ matrices, r can be expressed as a linear combination of their entries. In other words, there are matrices T^π such that

$$r(x) = \sum_{\pi \in \widehat{\mathsf{B}}_3} \langle V^{\pi,d}(x), T^\pi \rangle.$$

Then $(\tilde{R}^\pi, \tilde{S}_1^\pi, \tilde{S}_2^\pi - T^\pi)$ satisfies the SOS constraint (b). If the numbers in T^π are small enough compared to the minimum eigenvalue of $\tilde{S}_2^\pi$, then $\tilde{S}_2^\pi - T^\pi$ will be positive semidefinite, and we will have obtained a solution satisfying the SOS constraint in (14). Namely, it suffices to require

$$\|T^\pi\| \le \lambda_{\tilde{S}_2^\pi}$$

for all $\pi \in \widehat{\mathsf{B}}_3$, where $\lambda_{\tilde{S}_2^\pi}$ is any lower bound on the minimum eigenvalue of $\tilde{S}_2^\pi$, which may be obtained as explained above. Here, $\|A\| = \langle A, A\rangle^{1/2}$ is the Frobenius norm of matrix A.

To estimate $\|T^\pi\|$ we use the following approach in which we do not explicitly determine T^π. We find a maximal linearly independent subset $\mathcal{B}$ of polynomials inside the set of all entries of the $V^{\pi,d}$ matrices for $\pi \in \widehat{\mathsf{B}}_3$. Now we create the matrix A with rows indexed by all monomials occurring in a polynomial in $\mathcal{B}$ and columns indexed by $\mathcal{B}$. An entry (m, a) of A, where m is a monomial and $a \in \mathcal{B}$, contains the coefficient of monomial m in polynomial a. Then we find a submatrix $\hat{A}$ of A consisting of $|\mathcal{B}|$ linearly-independent rows of A and we compute $\hat{A}^{-1}$ using rational arithmetic. We may compute

$$\|\hat{A}^{-1}\|_\infty = \max_{i=1,\ldots,|\mathcal{B}|} \sum_{j=1}^{|\mathcal{B}|} |\hat{A}^{-1}_{ij}|$$

and observe that the maximum absolute value of any coefficient of the expansion of r in basis $\mathcal{B}$ is at most $\|\hat{A}^{-1}\|_\infty \|r\|_\infty$, where $\|r\|_\infty$ is the maximum absolute value of any coefficient of r. In this way we may get an estimate on $\|T^\pi\|$.

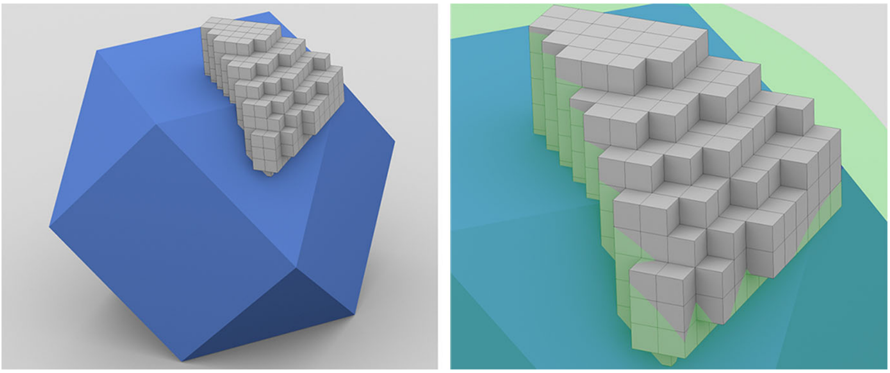

Fig. 1 Here we have an initial cube partition for $\mathcal{D}'$ where $\mathcal{K}$ is a regular tetrahedron so that $\mathcal{K} - \mathcal{K}$ is the cuboctahedron with circumradius 1. We take $\alpha = 1.02$ and $s(x) = \|x\|^2 - 1$. On the *left* we show the whole cuboctahedron and the partition. On the *right* we show the partition in detail; the unit sphere is also shown in *green*

6.2 Checking the Sample Constraints

Even if one uses a great number of sample constraints in condition (c) it is unlikely that the resulting function will be nonpositive in $\mathcal{D}$ as required. The sample constraints cannot accurately detect the boundary of $\mathcal{D}$. However, for some small factor $\alpha > 1$, which we hope will be small if the sample was fine enough, the function will be nonpositive in the domain

$$\mathcal{D}' = \{\, x \in \mathbb{R}^3 : s(x) < 0 \,\} \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ).$$

One may quickly estimate a good value for α by testing the function on a fine grid of points. Then all that is left to do is check that the function is really nonpositive in $\mathcal{D}'$. Of course, the larger the α, the worse the bound will be because it needs to be multiplied by α^3.

In our approach we use interval arithmetic to evaluate the polynomial $\mathcal{F}[g]$, so as to obtain rigorous results. We consider a partition of $\mathbb{R}^3$ into cubes of side-length δ for some small δ and we let $\mathcal{C}$ be the set of all partition cubes that contain at least one point $(x_1, x_2, x_3) \in \mathcal{D}'$ with $0 \le x_1 \le x_2 \le x_3$. Note that $\mathcal{C}$ is finite and covers $\mathcal{D}'$; Fig. 1 shows an example initial partition when $\mathcal{K}$ is the regular tetrahedron.

We then check that for every point in $\bigcup_{C \in \mathcal{C}} C$ the polynomial $\mathcal{F}[g]$ is nonpositive. We do that as follows.

First, for every cube $C \in \mathcal{C}$ we compute an upper bound of the norm of the gradient of $\mathcal{F}[g]$, a number ν_C such that

$$\|\nabla \mathcal{F}[g](x)\| \le \nu_C \tag{15}$$

for all $x \in C$. This is easy to do with interval arithmetic. We have the coefficients of $\mathcal{F}[g]$ represented by intervals. A cube $C = [x_1, y_1] \times [x_2, y_2] \times [x_3, y_3]$ is the product of three intervals. We then only have to compute $\nabla \mathcal{F}[g]([x_1, y_1], [x_2, y_2], [x_3, y_3])$

using interval arithmetic. This will give us a vector of intervals $([l_1, u_1], [l_2, u_2], [l_3, u_3])$ such that for all $(x_1, x_2, x_3) \in C$ we have

$$(l_1, l_2, l_3) \le \nabla \mathcal{F}[g](x_1, x_2, x_3) \le (u_1, u_2, u_3).$$

From this it is easy to compute a number ν_C satisfying (15).

Next, for a fixed integer $N \ge 1$, say, we uniformly divide each side of the cube C into N intervals, obtaining a grid of points inside of C. In other words, if x_C is the lower-left corner of C, we consider the set of points

$$C_N = \{ x_C + (a, b, c)\delta/N : 0 \le a, b, c \le N,\ a, b, c \in \mathbb{N} \}.$$

At least one point of C belongs to $\mathcal{D}'$, and hence at least one point of C_N is not in $\alpha(\mathcal{K}^\circ - \mathcal{K}^\circ)$. Let then $d(C, N)$ be the maximum minimum distance from any point of $C \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ)$ to a point of $C_N \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ)$ and let

$$\mu(C, N) = \max \{ \mathcal{F}[g](x) : x \in C_N \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ) \}.$$

If $\mu(C, N) > 0$, then the function is not nonpositive in the required domain. We hope however that, for our choice of $\alpha > 1$, we will have $\mu(C, N) < 0$. Suppose that this is the case. Given $x \in C \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ)$, let x' be the point in $C_N \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ)$ closest to x. By the mean-value theorem we have that

$$|\mathcal{F}[g](x) - \mathcal{F}[g](x')| \le \nu_C \|x - x'\| \le \nu_C d(C, N).$$

So, if

$$\nu_C d(C, N) \le |\mu(C, N)| \le |\mathcal{F}[g](x')|, \tag{16}$$

then $\mathcal{F}[g](x) \le 0$. Checking condition (16) for all $x' \in C_N \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ)$ gives us a sufficient condition that allows us to conclude that $\mathcal{F}[g](x) \le 0$ for all $x \in C \backslash \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ)$.

We still have to estimate $d(C, N)$, but that is a simple matter. There are two cases. If $C \cap \alpha(\mathcal{K}^\circ - \mathcal{K}^\circ) \ne \emptyset$, then $d(C, N) \le (\delta/N)\sqrt{3}$; if not, then $d(C, N) \le (\delta/(2N))\sqrt{3}$.

So our strategy is to process each cube in $\mathcal{C}$. For each cube, we start with $N = 2$ and check if that is enough to conclude that the function is nonpositive in the cube. If not, then we increase N. Once all cubes have been processed, we know that the function is nonpositive everywhere in the domain. Finally, notice that we always use interval arithmetic to perform all computations, thus obtaining rigorous results at the end, once the procedure terminates.

There is only one extra issue that is conceptually simple but that makes things technically harder. Computing with interval arithmetic is very slow, and hence if too many cubes would require dense grids (say, with hundreds of points per side), the computation would take several months. The size of the grid required by a cube is however directly proportional to the upper bound on the norm of the gradient, which

Table 6 List of rigorous bounds together with the factor α we needed in the verification

Body	Upper bound	Factor α
Regular octahedron (B_3^1)	0.972912750	1.001
B_3^3	0.823611150	1.002
B_3^4	0.874257405	1
B_3^5	0.922441815	1.005
B_3^6	0.933843309	1
Regular tetrahedron	0.374568355	1.02
Truncated cube	0.984519783	1.003
Truncated tetrahedron	0.729209804	1.023

is better the smaller the cube is. So, by taking smaller δ, we can improve on the grid sizes. But by changing δ globally, we increase the total number of cubes, possibly slowing down the total computation time.

A better strategy is as follows: if the grid size required by a cube is greater than a certain threshold (we use 30), then we split the cube at its center creating eight new cubes and keeping only those that intersect the domain. Then we process the resulting cubes instead, which are smaller and therefore lead to better grid sizes. This splitting process is carried out recursively, up to a certain maximum depth.

Finally, when one estimates the required grid size it may happen that, from one iteration to another, the grid size N is increased only slightly. This should be avoided, since computing the function is quite expensive. So our approach is as follows: first, we carry out the whole verification procedure using double-precision floating-point arithmetic for function evaluation, but not for the other computations. This is quite fast, finishing in a few hours. Then we use the estimated grid size for each cube in a checking routine that remakes all calculations using interval arithmetic.

6.3 Further Implementation Details

Table 6 contains the list of bounds we computed and rigorously verified using the approach described in this section.

The solutions, as well as the verification scripts and programs, can be found as ancillary files from the `arXiv.org` e-print archive.

The procedure to verify the SOS constraints was implemented as a Sage [46] script `verify.sage` and runs in Sage 6.2; see the documentation file `README_SOSChecking`. The approach to test that the function is nonpositive in the domain was implemented as a C++11 program called `checker` using the MPFI library [39] for interval arithmetic. Verification time was in all but one case under 2 days; in the case of the regular octahedron it took several weeks. More documentation of the C++11 program can be found in `README_SampleChecking` and a description of the classes in `docu`.

It is interesting to observe that the polynomials $\mathcal{F}[g]$ obtained from the solutions to the semidefinite programs we consider provide interesting low degree polyno-

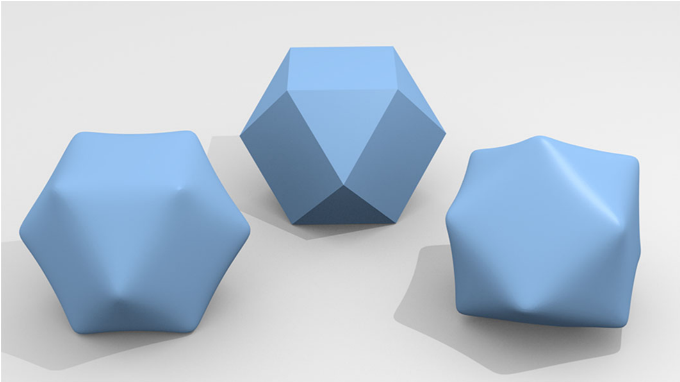

Fig. 2 The cuboctahedron (*center*) with two rotations of the set $\{ x \in \mathbb{R}^3 : \mathcal{F}[g](x) \geq 0 \}$, where g is the polynomial given by the solution to our problem for the regular tetrahedron

mial approximations of the Minkowski difference $\mathcal{K} - \mathcal{K}$; we used $d = 13$ in our computations, so that g and $\mathcal{F}[g]$ have degree 26. Figure 2 shows the cuboctahedron, which is the Minkowski difference of two regular tetrahedra, and the region $\{ x \in \mathbb{R}^3 : \mathcal{F}[g](x) \geq 0 \}$, where g is given by the solution to our problem for the tetrahedron. Notice how the region approximates the cuboctahedron. In fact, the upper bounds we computed are also bounds for translative packings of the nonconvex bodies determined by these polynomial approximations.

Acknowledgements Frank Vallentin thanks Peter Littelmann for a helpful discussion. We also thank the referees for their thorough comments which helped to improve the paper. Frank Vallentin was partially supported by VIDI Grant 639.032.917 from the Netherlands Organization for Scientific Research (NWO).

References

1. Andrews, G.E., Askey, R., Roy, R.: Special Functions. Encyclopedia of Mathematics and Its Applications, vol. 71. Cambridge University Press, Cambridge (1999)
2. Bachoc, C., Gijswijt, D.C., Schrijver, A., Vallentin, F.: Invariant semidefinite programs. In: Anjos, M.F., Lasserre, J.B. (eds.) Handbook on Semidefinite, Conic and Polynomial Optimization. International Series in Operations Research and Management Science, vol. 166, pp. 219–269. Springer, New York (2012). arXiv:1007.2905
3. Betke, U., Henk, M.: Densest lattice packings of 3-polytopes. Comput. Geom. **16**(3), 157–186 (2000). arXiv:math/9909172v1
4. Blekherman, G., Parrilo, P.A., Thomas, R.R. (eds.): Semidefinite Optimization and Convex Algebraic Geometry. MOS-SIAM Series on Optimization, vol. 13. SIAM, Philadelphia (2013)
5. Chen, E.R., Engel, M., Glotzer, S.C.: Dense crystalline dimer packings of regular tetrahedra. Discrete Comput. Geom. **44**(2), 253–280 (2010). arXiv:1001.0586
6. Cohn, H., Elkies, N.: New upper bounds on sphere packings I. Ann. Math. **157**(2), 689–714 (2003). arXiv:math/0110009
7. Cohn, H., Kumar, A.: Universally optimal distribution of points on spheres. J. Am. Math. Soc. **20**(1), 99–148 (2007). arXiv:math/0607446
8. Cohn, H., Kumar, A., Miller, S.D., Radchenko, D., Viazovska, M.: The sphere packing problem in dimension 24 (2016). arXiv:1603.06518
9. Cohn, H., Miller, S.D.: Some properties of optimal functions for sphere packing in dimensions 8 and 24 (2016). arXiv:1603.04759

10. Cohn, H., Zhao, Y.: Sphere packing bounds via spherical codes. Duke Math. J. **163**(10), 1965–2002 (2014). arXiv:1212.5966
11. Damasceno, P.F., Engel, M., Glotzer, S.C.: Crystalline assemblies and densest packings of a family of truncated tetrahedra and the role of directional entropic forces. ACS Nano **6**(1), 609–614 (2012). arXiv:1109.1323
12. D'Angelo, J.P.: Inequalities from Complex Analysis. Carus Mathematical Monographs, vol. 28. Mathematical Association of America, Washington (2002)
13. D'Angelo, J.P.: Hermitian analogues of Hilbert's 17th problem. Adv. Math. **226**(5), 4607–4637 (2011). arXiv:1012.2479
14. Dunkl, C.F.: Cube group invariant spherical harmonics and Krawtchouk polynomials. Math. Z. **177**(4), 561–577 (1981)
15. Elkies, N.D., Odlyzko, A.M., Rush, J.A.: On the packing densities of superballs and other bodies. Invent. Math. **105**(3), 613–639 (1991)
16. Fejes Tóth, G., Fodor, F., Vígh, V.: The packing density of the n-dimensional cross-polytope. Discrete Comput. Geom. **54**(1), 182–194 (2015). arXiv:1503.04571v1
17. Fujisawa, K., Fukuda, M., Kobayashi, K., Kojima, M., Nakata, K., Nakata, M., Yamashita, M.: SDPA (SemiDefinite Programming Algorithm) User's Manual—Version 7.0.5. Research Report B-448. Tokyo Institute of Technology, Tokyo (2008) http://www.is.titech.ac.jp/research/research-report/B/B-448, http://sdpa.sourceforge.net
18. Gardner, M.: Mathematical Carnival, 2nd edn. Mathematical Association of America, Washington, DC (1989)
19. Gatermann, K., Parrilo, P.A.: Symmetry groups, semidefinite programs, and sums of squares. J. Pure Appl. Algebra **192**(1–3), 95–128 (2004). arXiv:math/0211450
20. Gauß, C.F.: Recension der "Untersuchung über die Eigenschaften der positiven ternären quadratischen Formen von Ludwig August Seeber". J. Reine Angew. Math. **20**, 312–320 (1840)
21. Goethals, J.M., Seidel, J.J.: Cubature formulae, polytopes, and spherical designs. In: Davis, C., Grünbaum, B., Sherk, F.A. (eds.) The Geometric Vein, pp. 203–218. Springer, New York (1981)
22. de Graaf, J., van Roij, R., Dijkstra, M.: Dense regular packings of irregular nonconvex particles. Phys. Rev. Lett. **107**(15), # 155501 (2011)
23. Gravel, S., Elser, V., Kallus, Y.: Upper bound on the packing density of regular tetrahedra and octahedra. Discrete Comput. Geom. **46**, 799–818 (2011). arXiv:1008.2830
24. Groemer, H.: Über die dichteste gitterförmige Lagerung kongruenter Tetraeder. Monatsh. Math. **66**(1), 12–15 (1962)
25. Hales, T.C., Ferguson, S.P.: In: Lagarias, J.C. (ed.) The Kepler Conjecture: The Hales–Ferguson Proof. Springer, New York (2011)
26. Hicks, J.: A Poet with a Slide Rule: Piet Hein Bestrides Art and Science. Life Magazine, October 14, 55–66 (1966)
27. Hoylman, F.J.: The densest lattice packing of tetrahedra. Bull. Am. Math. Soc. **76**(1), 135–137 (1970)
28. Humphreys, J.E.: Reflection Groups and Coxeter Groups. Cambridge Studies in Advanced Mathematics, vol. 29. Cambridge University Press, Cambridge (1990)
29. Jiao, Y., Stillinger, F.H., Torquato, S.: Optimal packings of superballs. Phys. Rev. E **79**(4), 041309 (2009). arXiv:0902.1504
30. Jiao, Y., Stillinger, F.H., Torquato, S.: Erratum: optimal packings of superballs [Phys. Rev. E 79, 041309 (2009)]. Phys. Rev. E **84**(6), 069902 (2011)
31. Jiao, Y., Torquato, S.: A packing of truncated tetrahedra that nearly fills all of space and its melting properties. J. Chem. Phys. **135**(15), 151101 (2011). arXiv:1107.2300
32. Kabatiansky, G.A., Levenshtein, V.I.: Bounds for packings on a sphere and in space. Probl. Inf. Transm. **14**(1), 1–17 (1978)
33. de Laat, D., de Oliveira Filho, F.M., Vallentin, F.: Upper bounds for packings of spheres of several radii. Forum Math. Sigma **2**, e23 (2014). arXiv:1206.2608
34. Lagarias, J.C., Zong, C.: Mysteries in packing regular tetrahedra. Notices Am. Math. Soc. **59**(11), 1540–1549 (2012)
35. Minkowski, H.: Dichteste gitterförmige Lagerung kongruenter Körper. Nachr. Ges. Wiss. Göttingen, 311–355 (1904)
36. Ni, R., Gantapara, A.P., de Graaf, J., van Roij, R., Dijkstra, M.: Phase diagram of colloidal hard superballs: from cubes via spheres to octahedra. Soft Matter **8**, 8826–8834 (2012). arXiv:1111.4357

37. de Oliveira Filho, F.M., Vallentin, F.: Computing upper bounds for packing densities of congruent copies of a convex body I (2013). arXiv:1308.4893
38. Pütz, A.: Translative Packings of Convex Bodies and Finite Reflection Groups. Master's Thesis, University of Cologne (2016)
39. Revol, N., Rouillier, F.: Motivations for an arbitrary precision interval arithmetic and the MPFI library. Reliab. Comput. **11**(4), 275–290 (2005)
40. Reznick, B.: Some concrete aspects of Hilbert's 17th problem. In: Delzell, C.N., Madden, J.J. (eds.) Real Algebraic Geometry and Ordered Structures. Contemporary Mathematics, vol. 253, pp. 251–272. American Mathematical Society, Providence, RI (2000)
41. Rossi, L., Sacanna, S., Irvine, W.T.M., Chaikin, P.M., Pine, D.J., Philipse, A.P.: Cubic crystals from cubic colloids. Soft Matter **7**, 4139–4142 (2011)
42. Rush, J.A., Sloane, N.J.A.: An improvement to the Minkowski–Hlawka bound for packing superballs. Mathematika **34**(1), 8–18 (1987)
43. Serre, J.-P.: Linear Representations of Finite Groups. Graduate Texts in Mathematics, vol. 42. Springer, New York (1977)
44. Stanley, R.P.: Invariants of finite groups and their application to combinatorics. Bull. Am. Math. Soc. (N.S.) **1**(3), 475–511 (1979)
45. Stein, E.M., Weiss, G.L.: Introduction to Fourier Analysis on Euclidean Spaces. Princeton Mathematical Series, vol. 3. Princeton University Press, Princeton (1971)
46. Stein, W.A., *et al.*: Sage Mathematics Software (Version 6.2). The Sage Development Team (http://www.sagemath.org) (2012)
47. Sturmfels, B.: Algorithms in Invariant Theory. Texts and Monographs in Symbolic Computation. Springer, Vienna (1993)
48. Terras, A.: Fourier Analysis on Finite Groups and Applications. London Mathematical Society Student Texts, vol. 43. Cambridge University Press, Cambridge (1999)
49. Torquato, S., Jiao, Y.: Dense packings of the Platonic and Archimedean solids. Nature **460**, 876–879 (2009). arXiv:0908.4107
50. Torquato, S., Jiao, Y.: Dense packings of polyhedra: Platonic and Archimedean solids. Phys. Rev. E **80**(4), 041104 (2009)
51. Viazovska, M.S.: The sphere packing problem in dimension 8 (2016). arXiv:1603.04246
52. Ziegler, G.M.: Three mathematics competitions. In: Schleicher, D., Lackmann, M. (eds.) An Invitation to Mathematics: From Competitions to Research, pp. 195–205. Springer, Heidelberg (2011)
53. Zong, C.: On the translative packing densities of tetrahedra and cuboctahedra. Adv. Math. **260**, 130–190 (2014). arXiv:1208.0420