# UPPER BOUNDS FOR PACKINGS OF SPHERES OF SEVERAL RADII

DAVID DE LAAT[1], FERNANDO MÁRIO DE OLIVEIRA FILHO[2] and
FRANK VALLENTIN[3]

[1] Delft Institute of Applied Mathematics, Delft University of Technology, P.O. Box 5031,
2600 GA Delft, The Netherlands;
email: mail@daviddelaat.nl

[2] Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão, 1010,
05508-090 São Paulo, Brazil;
email: fmario@gmail.com

[3] Mathematisches Institut, Universität zu Köln, Weyertal 86-90, 50931 Köln, Germany;
email: frank.vallentin@uni-koeln.de

## Abstract

We give theorems that can be used to upper bound the densities of packings of different spherical caps in the unit sphere and of translates of different convex bodies in Euclidean space. These theorems extend the linear programming bounds for packings of spherical caps and of convex bodies through the use of semidefinite programming. We perform explicit computations, obtaining new bounds for packings of spherical caps of two different sizes and for binary sphere packings. We also slightly improve the bounds for the classical problem of packing identical spheres.

2010 Mathematics Subject Classification: 52C17, 90C22 (primary)

## 1. Introduction

How densely can one pack given objects into a given container? Problems of this sort, generally called *packing problems*, are fundamental problems in geometric optimization.

An important example having a rich history is the sphere packing problem. Here, one tries to place equal-sized spheres with pairwise-disjoint interiors into $n$-dimensional Euclidean space while maximizing the fraction of covered space.

In two dimensions the best packing is given by placing open disks centered at the points of the hexagonal lattice. In three dimensions, the statement that the best sphere packing has density $\pi/\sqrt{18} = 0.7404\ldots$ was known as Kepler's conjecture; it was proved by Hales [21] in 1998 by means of a computer-assisted proof.

Currently, one of the best methods for obtaining upper bounds for the density of sphere packings is due to Cohn and Elkies [8]. In 2003, they used linear programming to obtain the best-known upper bounds for the densities of sphere packings in dimensions 4, ..., 36. They almost closed the gap between lower and upper bounds in dimensions 8 and 24. Their method is the noncompact version of the linear programming method of Delsarte *et al.* [12] for upper bounding the densities of packings of spherical caps on the unit sphere.

From a physical point of view, packings of spheres of different sizes are relevant, as they can be used to model chemical mixtures which consist of multiple atoms or, more generally, to model the structure of composite material. For more about technological applications of these kinds of systems of polydisperse, totally impenetrable spheres, we refer to Torquato [40, Ch. 6]. In recent work, Hopkins *et al.* [26, 27] presented lower bounds for the densities of packings of spheres of two different sizes, also called *binary sphere packings*.

In coding theory, packings of spheres of different sizes are important in the design of error-correcting codes which can be used for unequal error protection. Masnick and Wolf [32] were the first who considered codes with this property.

In this paper, we extend the linear programming method of Cohn and Elkies to obtain new upper bounds for the densities of multiple-size sphere packings. We also extend the linear programming method of Delsarte, Goethals, and Seidel to obtain new upper bounds for the densities of multiple-size spherical cap packings.

We perform explicit calculations for binary packings in both cases using semidefinite, instead of linear, programming. In particular, we complement the constructive lower bounds of Hopkins, Jiao, Stillinger, and Torquato by nonconstructive upper bounds. Insights gained from our computational approach are then used to improve known upper bounds for the densities of monodisperse sphere packings in dimensions 4, ... 9, except 8. The bounds we present improve on the best-known bounds due to Cohn and Elkies [8].

## 1.1. Methods and theorems.
We model the packing problems using tools from combinatorial optimization. All possible positions of the objects which we can use for the packing are vertices of a graph, and we draw edges between two vertices whenever the two corresponding objects cannot be simultaneously present in the packing because they overlap in their interiors. Now every independent set in this conflict graph gives a valid packing, and *vice versa*.

To determine the density of the packing, we use vertex weights, since we want to distinguish between 'small' and 'big' objects. For finite graphs, it is known that the weighted independence number can be upper bounded by the weighted theta number. Our theorems for packings of spherical caps and spheres are infinite-dimensional analogs of this result.

Let $G = (V, E)$ be a finite graph. A set $I \subseteq V$ is *independent* if no two vertices in $I$ are adjacent. Given a weight function $w \colon V \to \mathbb{R}_{\geqslant 0}$, the *weighted independence number* of $G$ is the maximum weight of an independent set; that is,

$$\alpha_w(G) = \max\left\{ \sum_{x \in I} w(x) : I \subseteq V \text{ is independent} \right\}.$$

Finding $\alpha_w(G)$ is an NP-hard problem.

Grötschel *et al.* [20] defined a graph parameter that gives an upper bound for $\alpha_w$ and which can be computed efficiently by semidefinite optimization. It can be presented in many different yet equivalent ways, but the way that is convenient for us is

$$
\begin{aligned}
\vartheta'_w(G) = \min \quad & M \\
& K - (w^{1/2})(w^{1/2})^{\mathsf{T}} \quad \text{is positive semidefinite,} \\
& K(x, x) \leqslant M \quad\quad\quad\quad\ \text{for all } x \in V, \\
& K(x, y) \leqslant 0 \quad\quad\quad\quad\ \ \text{for all } \{x, y\} \notin E \text{ where } x \neq y, \\
& M \in \mathbb{R}, K \in \mathbb{R}^{V \times V} \quad \text{is symmetric.}
\end{aligned}
$$

Here, we give a proof of the fact that $\vartheta'_w(G)$ upper bounds $\alpha_w(G)$. In a sense, after discarding the analytical arguments in the proofs of Theorems 1.2 and 1.3, we are left with this simple proof.

THEOREM 1.1. *For any finite graph $G = (V, E)$ with weight function $w \colon V \to \mathbb{R}_{\geqslant 0}$, we have $\alpha_w(G) \leqslant \vartheta'_w(G)$.*

*Proof.* Let $I \subseteq V$ be an independent set of nonzero weight, and let $K \in \mathbb{R}^{V \times V}$, $M \in \mathbb{R}$ be a feasible solution of $\vartheta'_w(G)$. Consider the sum

$$\sum_{x, y \in I} w(x)^{1/2} w(y)^{1/2} K(x, y).$$

This sum is at least

$$\sum_{x, y \in I} w(x)^{1/2} w(y)^{1/2} w(x)^{1/2} w(y)^{1/2} = \left( \sum_{x \in I} w(x) \right)^2$$

because $K - (w^{1/2})(w^{1/2})^{\mathsf{T}}$ is positive semidefinite.

The sum is also at most

$$\sum_{x \in I} w(x) K(x, x) \leqslant M \sum_{x \in I} w(x)$$

because $K(x, x) \leqslant M$ and because $K(x, y) \leqslant 0$ whenever $x \neq y$, as $I$ forms an independent set. Now, combining both inequalities proves the theorem. □

**Multiple-size spherical cap packings.** We first consider packings of spherical caps of several radii on the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : x \cdot x = 1\}$. The spherical cap with angle $\alpha \in [0, \pi]$ and center $x \in S^{n-1}$ is given by

$$C(x, \alpha) = \{y \in S^{n-1} : x \cdot y \geqslant \cos \alpha\}.$$

Its normalized volume equals

$$w(\alpha) = \frac{\omega_{n-1}(S^{n-2})}{\omega_n(S^{n-1})} \int_{\cos \alpha}^1 (1 - u^2)^{(n-3)/2} \, du,$$

where $\omega_n(S^{n-1}) = (2\pi^{n/2})/\Gamma(n/2)$ is the surface area of the unit sphere. Two spherical caps $C(x_1, \alpha_1)$ and $C(x_2, \alpha_2)$ intersect in their topological interiors if and only if the inner product of $x_1$ and $x_2$ lies in the interval $(\cos(\alpha_1 + \alpha_2), 1]$. Conversely, we have

$$C(x_1, \alpha_1)^\circ \cap C(x_2, \alpha_2)^\circ = \emptyset \iff x_1 \cdot x_2 \leqslant \cos(\alpha_1 + \alpha_2).$$

A *packing* of spherical caps with angles $\alpha_1, \ldots, \alpha_N$ is a union of any number of spherical caps with these angles and pairwise-disjoint interiors. The density of the packing is the sum of the normalized volumes of the constituting spherical caps.

The optimal packing density is given by the weighted independence number of the *spherical cap packing graph*. This is the graph with vertex set $S^{n-1} \times \{1, \ldots, N\}$, where a vertex $(x, i)$ has weight $w(\alpha_i)$, and where two distinct vertices $(x, i)$ and $(y, j)$ are adjacent if $\cos(\alpha_i + \alpha_j) < x \cdot y$.

In Section 2, we will extend the weighted theta prime number to the spherical cap packing graph. There we will also derive Theorem 1.2 below, which gives upper bounds for the densities of packings of spherical caps. We will show that the sharpest bound given by this theorem is in fact equal to the theta prime number.

In what follows, we denote by $P_k^n$ the Jacobi polynomial $P_k^{((n-3)/2, (n-3)/2)}$ of degree $k$, normalized so that $P_k^n(1) = 1$. Jacobi polynomials are orthogonal polynomials defined on the interval $[-1, 1]$ with respect to the measure $(1 - u^2)^{(n-3)/2} du$. See, for instance, Andrews *et al.* [1] for more information.

THEOREM 1.2. *Let $\alpha_1, \ldots, \alpha_N \in (0, \pi]$ be angles, and, for $i, j = 1, \ldots, N$ and $k \geqslant 0$, let $f_{ij,k}$ be real numbers such that $f_{ij,k} = f_{ji,k}$ and $\sum_{k=0}^{\infty} |f_{ij,k}| < \infty$ for all $i, j$. Write*

$$f_{ij}(u) = \sum_{k=0}^{\infty} f_{ij,k} P_k^n(u). \tag{1.1}$$

*Suppose that the functions $f_{ij}$ satisfy the following conditions.*

(i) $(f_{ij,0} - w(\alpha_i)^{1/2} w(\alpha_j)^{1/2})_{i,j=1}^{N}$ *is positive semidefinite.*

(ii) $(f_{ij,k})_{i,j=1}^{N}$ *is positive semidefinite for $k \geqslant 1$.*

(iii) $f_{ij}(u) \leqslant 0$ *whenever $-1 \leqslant u \leqslant \cos(\alpha_i + \alpha_j)$.*

*Then the density of every packing of spherical caps with angles $\alpha_1, \ldots, \alpha_N$ on the unit sphere $S^{n-1}$ is at most $\max\{f_{ii}(1) : i = 1, \ldots, N\}$.*

When $N = 1$, Theorem 1.2 reduces to the linear programming bound for spherical cap packings of Delsarte *et al.* [**12**]. In Section 4, we use semidefinite programming instead of linear programming to perform explicit computations for $N = 2$.

**Translational packings of bodies and multiple-size sphere packings.** We now deal with packings of spheres with several radii in $\mathbb{R}^n$. Theorem 1.3 presented below can be used to find upper bounds for the densities of such packings. In fact, it is more general, and it can be applied to packings of translates of different convex bodies.

Let $\mathcal{K}_1, \ldots, \mathcal{K}_N$ be convex bodies in $\mathbb{R}^n$. A translational *packing* of $\mathcal{K}_1, \ldots, \mathcal{K}_N$ is a union of translations of these bodies in which any two copies have disjoint interiors. The *density* of a packing is the fraction of space covered by it. There are different ways to formalize this definition, and questions appear as to whether every packing has a density and so on. We postpone further discussion on this matter until Section 3, where we give a proof of Theorem 1.3.

Our theorem can be seen as an analog of the weighted theta prime number $\vartheta_w'$ for the infinite graph $G$ whose vertex set is $\mathbb{R}^n \times \{1, \ldots, N\}$ and in which vertices $(x, i)$ and $(y, j)$ are adjacent if $x + \mathcal{K}_i$ and $y + \mathcal{K}_j$ have disjoint interiors. The weight function we consider assigns weight $\mathrm{vol}\,\mathcal{K}_i$ to vertex $(x, i) \in \mathbb{R}^n \times \{1, \ldots, N\}$. We will say more about this interpretation in Section 3.

For the statement of the theorem, we need some basic facts from harmonic analysis. Let $f : \mathbb{R}^n \to \mathbb{C}$ be an $L^1$ function. For $u \in \mathbb{R}^n$, the Fourier transform

of $f$ at $u$ is

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

We say that function $f$ is a *Schwartz function* (also called a *rapidly decreasing function*) if it is infinitely differentiable, and if any derivative of $f$, multiplied by any power of the variables $x_1, \ldots, x_n$, is a bounded function. The Fourier transform of a Schwartz function is a Schwartz function, too. A Schwartz function can be recovered from its Fourier transform by means of the *inversion formula*:

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(u) e^{2\pi i u \cdot x} \, du$$

for all $x \in \mathbb{R}^n$.

THEOREM 1.3. *Let $\mathcal{K}_1, \ldots, \mathcal{K}_N$ be convex bodies in $\mathbb{R}^n$, and let $f \colon \mathbb{R}^n \to \mathbb{R}^{N \times N}$ be a matrix-valued function whose every component $f_{ij}$ is a Schwartz function. Suppose that $f$ satisfies the following conditions.*

(i) *The matrix $\left( \hat{f}_{ij}(0) - (\text{vol } \mathcal{K}_i)^{1/2} (\text{vol } \mathcal{K}_j)^{1/2} \right)_{i,j=1}^{N}$ is positive semidefinite.*

(ii) *The matrix of Fourier transforms $\left( \hat{f}_{ij}(u) \right)_{i,j=1}^{N}$ is positive semidefinite for every $u \in \mathbb{R}^n \setminus \{0\}$.*

(iii) *$f_{ij}(x) \leqslant 0$ whenever $\mathcal{K}_i^\circ \cap (x + \mathcal{K}_j^\circ) = \emptyset$.*

*Then the density of any packing of translates of $\mathcal{K}_1, \ldots, \mathcal{K}_N$ in the Euclidean space $\mathbb{R}^n$ is at most $\max\{f_{ii}(0) : i = 1, \ldots, N\}$.*

We give a proof of this theorem in Section 3. When $N = 1$, and when the convex body $\mathcal{K}_1$ is centrally symmetric (an assumption which is in fact not needed), then this theorem reduces to the linear programming method of Cohn and Elkies [8].

We apply this theorem to obtain upper bounds for the densities of binary sphere packings, as we discuss in Section 1.3.

## 1.2. Computational results for binary spherical cap packings.

We applied Theorem 1.2 to compute upper bounds for the densities of binary spherical cap packings. The results we obtained are summarized in the plots of Figure 1.

For $n = 3$, Florian [14, 15] provides a geometric upper bound for the density of a spherical cap packing. He shows that the density of a packing on $S^2$ of spherical caps with angles $\alpha_1, \ldots, \alpha_N \in (0, \pi/3)$ is at most

$$\max_{1 \leqslant i \leqslant j \leqslant k \leqslant N} D(\alpha_i, \alpha_j, \alpha_k),$$

(a) $n = 3$

(b) SDP bound/geometric bound for $n = 3$

(c) $n = 4$

(d) $n = 5$
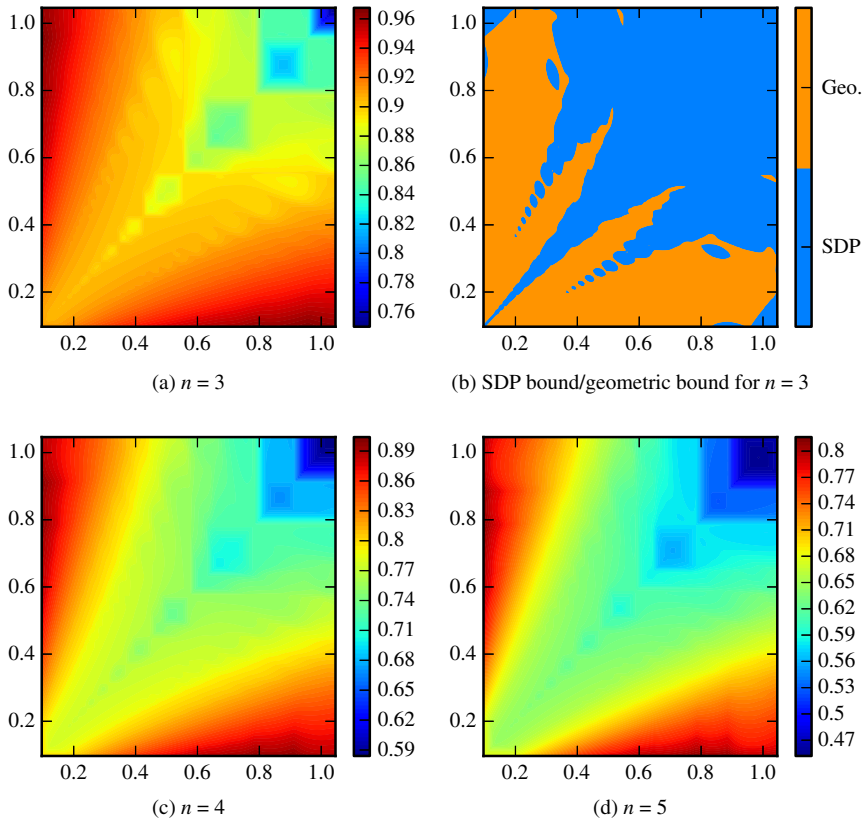
Figure 1. Upper bounds on the packing density for $N = 2$. The horizontal and vertical axes represent the spherical cap angle; the colors indicate the density, or in the case of plot (b) whether the SDP bound or the geometric bound is sharper.

where $D(\alpha_i, \alpha_j, \alpha_k)$ is defined as follows. Let $\mathcal{T}$ be a spherical triangle in $S^2$ such that, if we center the spherical caps with angles $\alpha_i$, $\alpha_j$, and $\alpha_k$ at the vertices of $\mathcal{T}$, then the caps intersect pairwise at their boundaries. The number $D(\alpha_i, \alpha_j, \alpha_k)$ is then defined as the fraction of the area of $\mathcal{T}$ covered by the caps.

In Figure 1(b), we see that for $N = 2$ it depends on the angles whether the geometric or the semidefinite programming (SDP) bound is sharper. In particular, we see that near the diagonal the semidefinite programming bound is at least as good as the geometric bound; see also Figure 1(a).

We can construct natural multiple-size spherical cap packings by taking the incircles of the faces of spherical Archimedean tilings. A sequence of binary

packings is for instance obtained by taking the incircles of the prism tilings. These are the Archimedean tilings with vertex figure $(4, 4, m)$ for $m \geqslant 3$ (although strictly speaking for $m = 4$ this is a spherical Platonic tiling). The question then is whether the packing associated with the $m$-prism has maximal density among all packings with the same cap angles $\pi/m$ and $\pi/2 - \pi/m$, that is, whether the packing is *maximal*. The packing for $m = 3$ is not maximal while the one for $m = 4$ trivially is, since here there is only one cap size, and adding a ninth cap yields a density greater than 1.

Heppes and Kertész [**23**] showed that the configurations for $m \geqslant 6$ are maximal, and the remaining case $m = 5$ was later shown to maximal by Florian and Heppes [**16**]. Florian [**14**] showed that the geometric bound given above is in fact sharp for the cases where $m \geqslant 6$, and for the $m = 5$ case it is not sharp but is still good enough to prove maximality (notice that, given a finite number of cap angles, the set of obtainable densities is finite).

Now we illustrate that Theorem 1.2 gives a sharp bound for the density of the packing associated to the 5-prism, thus giving a simple proof of its maximality. The theorem also provides a sharp bound for $m = 4$, but whether it can provide sharp bounds for the cases $m \geqslant 6$ we do not know at the moment. The numerical results are not decisive.

We shall exhibit functions

$$f_{ij}(u) = \sum_{k=0}^{4} f_{ij,k} P_k^n(u)$$

which satisfy the conditions of Theorem 1.2 with $f_{11}(1) = f_{22}(1) = 5w(\alpha_1) + 2w(\alpha_2)$, where

$$\alpha_1 = \frac{\pi}{5}, \quad \alpha_2 = \frac{3\pi}{10}, \quad w(\alpha_1) = \frac{1}{2}\left(1 - \cos\frac{\pi}{5}\right), \quad w(\alpha_2) = \frac{1}{2}\left(1 - \cos\frac{3\pi}{10}\right).$$

For a sharp solution, the inequalities in the proof of Theorem 1.2 must be equalities, so the $f_{ij,k}$ have to satisfy the following linear conditions.

$$0 = f_{11}\left(\cos\frac{2\pi}{5}\right) = f_{11}\left(\cos\frac{4\pi}{5}\right) = f_{11}'\left(\cos\frac{4\pi}{5}\right) = f_{12}(0) = f_{22}(-1).$$

The product

$$\begin{pmatrix} f_{11,0} & f_{12,0} \\ f_{12,0} & f_{22,0} \end{pmatrix} \begin{pmatrix} 25w(\alpha_1) & 10\sqrt{w(\alpha_1)w(\alpha_2)} \\ 10\sqrt{w(\alpha_1)w(\alpha_2)} & 4w(\alpha_2) \end{pmatrix}$$

equals

$$\begin{pmatrix} 25w(\alpha_1)^2 + 10w(\alpha_1)w(\alpha_2) & \sqrt{w(\alpha_1)w(\alpha_2)}(10w(\alpha_1) + 4w(\alpha_2)) \\ \sqrt{w(\alpha_1)w(\alpha_2)}(25w(\alpha_1) + 10w(\alpha_2)) & 10w(\alpha_1)w(\alpha_2) + 4w(\alpha_2)^2 \end{pmatrix}.$$

For $k = 1, \ldots, 4$ the product of the two matrices $\begin{pmatrix} f_{11,k} & f_{12,k} \\ f_{12,k} & f_{22,k} \end{pmatrix}$ and

$$\begin{pmatrix} w(\alpha_1)\left(5P_k(1) + 10P_k\left(\cos\frac{2\pi}{5}\right) + 10P_k\left(\cos\frac{2\pi}{4}\right)\right) & \sqrt{w(\alpha_1)w(\alpha_2)}10P_k(0) \\ \sqrt{w(\alpha_1)w(\alpha_2)}10P_k(0) & w(\alpha_2)(2P_k(1) + 2P_k(-1)) \end{pmatrix}$$

equals zero. This linear system together with the additional assumptions

$$0 = f_{11}(-1) = f_{12}\left(-\frac{95}{100}\right) = f'_{12}\left(-\frac{95}{100}\right)$$

has a one-dimensional space of solutions from which it is easy to select one which fulfills all requirements of Theorem 1.2.

For the remaining 13 Archimedean solids in dimension $n = 3$, we are only able to show maximality of the packing associated to the truncated octahedron, the Archimedean solid with vertex figure $(6, 6, 5)$. Its density is $0.9056\ldots$, the geometric bound shows that the density is at most $0.9088\ldots$, and using the semidefinite program we get $0.9079\ldots$ as an upper bound. The first packing with caps of angles $\arcsin(1/3)$ and $\arcsin(1/\sqrt{3})$ which would be denser is obtained by taking 19 of the smaller caps and 4 of the bigger caps, and it has density $0.9103\ldots$. The upper bounds show however that it is not possible to obtain this dense a packing, thus showing that the truncated octahedron packing is maximal.

We also used our programs to plot the upper bounds for $N = 1$, the classical linear programming bound of Delsarte *et al.* [12], for dimensions $n = 3, 4$, and 5 in Figure 2. To the best of our knowledge these kinds of plots have not been made before, and they seem to reveal interesting properties of the bound. For better orientation, we show in the plots the packings where the linear programming bound is sharp (see Levenshtein [29]; Cohn and Kumar [9] proved the much stronger statement that these packings provide point configurations which are universally optimal). The dotted line in the plot for $n = 3$ is the geometric bound, and, since we know that both the geometric (see Florian [14]) and the semidefinite programming bounds are sharp for the given configurations, we know that at these peaks the bounds meet.

An interesting feature of the upper bound seems to be that it has some periodic behavior. Indeed, the numerical results suggest that, for $n = 3$, the two bounds in fact meet infinitely often as the angle decreases, and that between
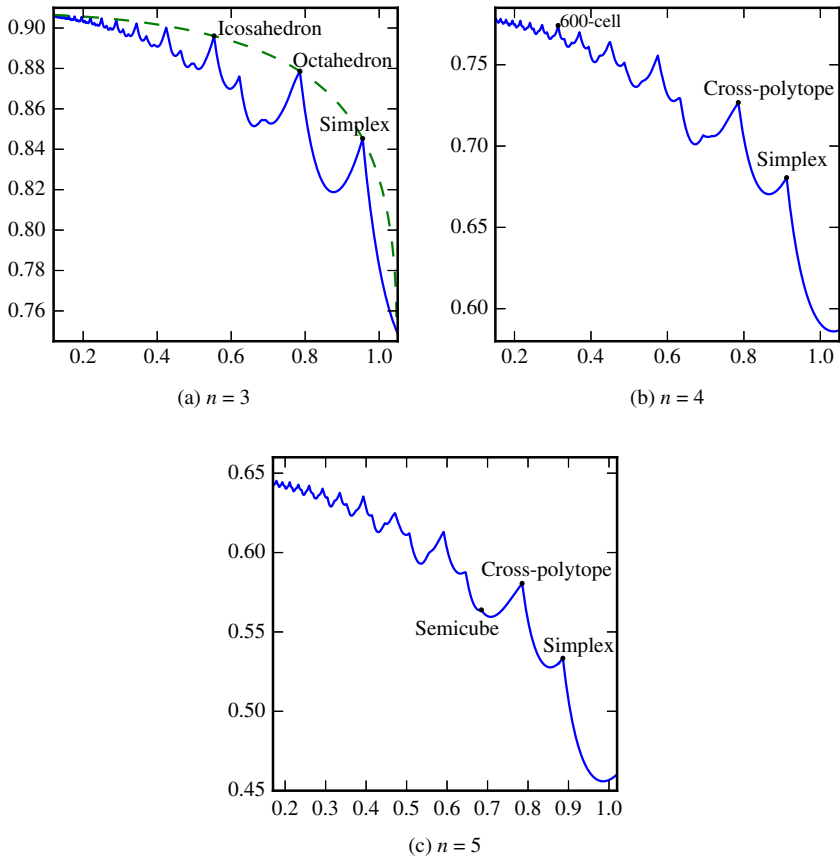
(a) *n* = 3

(b) *n* = 4

(c) *n* = 5

Figure 2. Upper bounds on the packing density for $N = 1$. The horizontal axis represents the spherical cap angle and the vertical axis the packing density.

any two of these meeting points the semidefinite programming bound has a similar shape. Although in higher dimensions we do not have a geometric bound, the semidefinite programming bound seems to admit the same kind of periodic behavior.

**1.3. Computational results for binary sphere packings.** We applied Theorem 1.3 to compute upper bounds for the densities of binary sphere packings. The results we obtained are summarized in the plot of Figure 3, where we show bounds computed for dimensions 2, . . . , 5. A detailed account of our
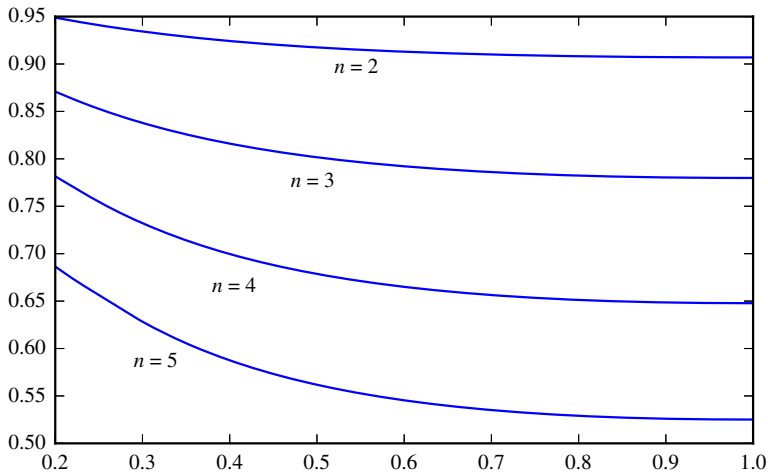
Figure 3. The horizontal axis represents the ratio between the radii of the small and the large spheres. The vertical axis represents our upper bound. Our bounds for dimensions 2, ..., 5 are shown together.

approach is given in Section 5. We now briefly discuss the bounds presented in Figure 3.

*Dimension* 2. Only in dimension 2 have binary sphere (that is, circle) packings been studied in depth. We refer to the introduction in the paper of Heppes [22], which surveys the known results about binary circle packings in the plane.

Currently, one of the best-known upper bounds for the maximum density of a binary circle packing is due to Florian [13]. Florian's bound states that a packing of circles in which the ratio between the radii of the smallest and largest circles is $r$ has density at most

$$\frac{\pi r^2 + 2(1 - r^2) \arcsin(r/(1 + r))}{2r\sqrt{2r + 1}},$$

and that this bound is achieved exactly for $r = 1$ (that is, for classical circle packings) and for $r = 0$ in the limit.

The question arises of which bound is better, our bound or Florian's bound. From our experiments, it seems that our bound is worse than Florian's bound, at least for $r < 1$. For instance, for $r = 1/2$ we obtain the upper bound $0.9174426\ldots$, whereas Florian's bound is $0.9158118\ldots$. Whether this really means that the bound of Theorem 1.3 is worse than Florian's bound, or just that the computational approach of Section 5 is too restrictive to attain his bound, we do not know.

It is interesting to note that, for $r = 1$, that is, for packings of circles of one size, our bound clearly coincides with that of Cohn and Elkies [8]. This bound seems to be equal to $\pi/\sqrt{12}$, but no proof of this is known.

*Dimension* 3. Much less is known in dimension 3. In fact we do not know about other attempts to find upper bounds for the densities of binary sphere packings in dimensions 3 and higher.

Let us compare our upper bound with the lower bound by Hopkins *et al.* [26]. The record holder for $r \geqslant 0.2$ in terms of highest density occurs for $r = 0.224744\ldots$, and its density is $0.824539\ldots$. Our computations show that there cannot be a packing with this $r$ having density more than $0.8617125\ldots$, so this leaves a margin of 5%.

Another interesting case is $r = \sqrt{2}-1 = 0.414\ldots$. Here the best-known lower bound of $0.793\ldots$ comes from the crystall structure of sodium chloride NaCl. The large spheres are centered at a face centered cubic lattice and the small spheres are centered at a translated copy of the face centered cubic lattice so that they form a jammed packing. Our upper bound for $r = \sqrt{2} - 1$ is $0.813\ldots$, less than 3% away from the lower bound. Therefore, we believe that proving optimality of NaCl might be doable.

*Dimension* 4 *and beyond.* In higher dimensions, even less is known about binary sphere packings. We observed from Figure 3 that it seems that the upper bound is decreasing: as the radius of the small sphere increases from 0.2 to 1, the bound seems to decrease. This suggests that the bound given by Theorem 1.3 is decreasing in this sense, but we do not know a proof of this.

We also do not know the limit behavior of our bound when $r$ approaches 0. Due to numerical instabilities we could not perform numerical calculations in this regime of $r$.

## 1.4. Improving the Cohn–Elkies bounds.

We now present a theorem that can be used to find better upper bounds for the densities of monodisperse sphere packings than those provided by Cohn and Elkies [8]; our theorem is a strengthening of theirs.

Fix $\varepsilon > 0$. Given a packing of spheres of radius $1/2$, we consider its $\varepsilon$-*tangency graph*, a graph whose vertices are the spheres in the packing, and in which two vertices are adjacent if the distance between the centers of the respective spheres lies in the interval $[1, 1 + \varepsilon]$.

Let $M(\varepsilon)$ be the least upper bound on the average degree of the $\varepsilon$-tangency graph of any sphere packing. Our theorem is the following.

THEOREM 1.4. *Take $0 = \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_m$, and let $f : \mathbb{R}^n \to \mathbb{R}$ be a Schwartz function such that the following hold.*

(i)  $\hat{f}(0) \geqslant \mathrm{vol}\, B$, *where $B$ is the ball of radius $1/2$.*

(ii)  $\hat{f}(u) \geqslant 0$ *for all $u \in \mathbb{R}^n \setminus \{0\}$.*

(iii)  $f(x) \leqslant 0$ *whenever $\|x\| \geqslant 1 + \varepsilon_m$.*

(iv)  $f(x) \leqslant \eta_k$ *whenever $\|x\| \in [1 + \varepsilon_{k-1}, 1 + \varepsilon_k)$ with $\eta_k \geqslant 0$, for $k = 1, \ldots, m$.*

*Then the density of a sphere packing is at most the optimal value of the following linear programming problem in variables $A_1, \ldots, A_m$:*

$$\begin{aligned}
\max \quad & f(0) + \eta_1 A_1 + \cdots + \eta_m A_m \\
& A_1 + \cdots + A_k \leqslant U(\varepsilon_k) \qquad \text{for } k = 1, \ldots, m, \\
& A_i \geqslant 0 \qquad\qquad\qquad\quad \text{for } i = 1, \ldots, m,
\end{aligned} \qquad (1.2)$$

*where $U(\varepsilon_k) \geqslant M(\varepsilon_k)$ for $k = 1, \ldots, m$.*

In Section 6, we give a proof of Theorem 1.4, and we show how to compute upper bounds for $M(\varepsilon)$ using the semidefinite programming bounds of Bachoc and Vallentin [4] for the sizes of spherical codes. There we also show how to use semidefinite programming and the same ideas we employ in the computations for binary sphere packings (see Section 5) to compute better upper bounds for the densities of sphere packings.

In Table 1, we show the upper bounds obtained through our application of Theorem 1.4. To better compare our bounds with those of Cohn and Elkies, in Table 1 we show bounds for the center density of a packing, the *center density* of a packing of unit spheres being equal to $\Delta / \mathrm{vol}\, B$, where $\Delta$ is the density of the packing, and $B$ is a unit ball.

We omit dimension 8 because for this dimension it is already believed that the Cohn–Elkies bound is itself optimal, and therefore as is to be expected we did not manage to obtain any improvement over their bound. We also note that the bounds by Cohn and Elkies are the best-known upper bounds in all other dimensions shown.

In dimension 3, the Cohn–Elkies bound is 0.18616, whereas the optimal sphere packing has center density 0.17678. We can improve the Cohn–Elkies bound to 0.184559, which is also better than the upper bound 0.1847 due to Rogers [35].

## 2.  Multiple-size spherical cap packings

In this section, we prove Theorem 1.2 and discuss its relation to an extension of the weighted theta prime number for the spherical cap packing graph.

Table 1. For each dimension, we show the best lower bound known, the bound by Cohn and Elkies [8], and the upper bound coming from Theorem 1.4.

| Dimension | Lower bound | Cohn–Elkies bound | New upper bound |
|:---:|:---:|:---:|:---:|
| 4 | 0.12500 | 0.13126 | 0.130587 |
| 5 | 0.08839 | 0.09975 | 0.099408 |
| 6 | 0.07217 | 0.08084 | 0.080618 |
| 7 | 0.06250 | 0.06933 | 0.069193 |
| 9 | 0.04419 | 0.05900 | 0.058951 |

**2.1.   Proof of Theorem 1.2.**   Let $x_1, \ldots, x_m \in S^{n-1}$ and $r : \{1, \ldots, m\} \to \{1, \ldots, N\}$ be such that

$$\bigcup_{i=1}^{m} C(x_i, \alpha_{r(i)})$$

is a packing of spherical caps on $S^{n-1}$.

Consider the sum

$$\sum_{i,j=1}^{m} w(\alpha_{r(i)})^{1/2} w(\alpha_{r(j)})^{1/2} f_{r(i)r(j)}(x_i \cdot x_j). \qquad (2.1)$$

By expanding $f_{r(i)r(j)}(x_i \cdot x_j)$ according to (1.1), this sum is equal to

$$\sum_{k=0}^{\infty} \sum_{i,j=1}^{m} w(\alpha_{r(i)})^{1/2} w(\alpha_{r(j)})^{1/2} f_{r(i)r(j),k} P_k^n(x_i \cdot x_j).$$

By the addition formula (see for example Section 9.6 of Andrews *et al.* [1]) for the Jacobi polynomials $P_k^n$, the matrix $\left(P_k^n(x_i \cdot x_j)\right)_{i,j=1}^{m}$ is positive semidefinite. From condition (ii) of the theorem, we also know that the matrix $\left(f_{r(i)r(j),k}\right)_{i,j=1}^{m}$ is positive semidefinite for $k \geqslant 1$. So the inner sum above is nonnegative for $k \geqslant 1$. If we then consider only the summand for $k = 0$, we see that (2.1) is at least

$$\sum_{i,j=1}^{m} w(\alpha_{r(i)})^{1/2} w(\alpha_{r(j)})^{1/2} f_{r(i)r(j),0} P_0^n(x_i \cdot x_j) \geqslant \left(\sum_{i=1}^{m} w(\alpha_i)\right)^2, \qquad (2.2)$$

where the inequality follows from condition (i) of the theorem.

Now, notice that, whenever $i \neq j$, the caps $C(x_i, \alpha_{r(i)})$ and $C(x_j, \alpha_{r(j)})$ have disjoint interiors. Condition (iii) then implies that $f_{r(i)r(j)}(x_i \cdot x_j) \leqslant 0$. So we see

that (2.1) is at most

$$\sum_{i=1}^{m} w(\alpha_i) f_{r(i)r(i)}(1) \leqslant \max\{f_{ii}(1) : i = 1, \ldots, N\} \sum_{i=1}^{m} w(\alpha_i). \qquad (2.3)$$

So (2.1) is at least (2.2) and at most (2.3), yielding

$$\sum_{i=1}^{m} w(\alpha_i) \leqslant \max\{f_{ii}(1) : i = 1, \ldots, N\}. \qquad \square$$

### 2.2. Theorem 1.2 and the Lovász theta number.

We now briefly discuss a generalization of $\vartheta'_w$ to infinite graphs and its relation to the bound of Theorem 1.2. Similar ideas were developed by Bachoc *et al.* [3].

Let $G = (V, E)$ be a graph, where $V$ is a compact space, and let $w : V \to \mathbb{R}_{\geqslant 0}$ be a continuous weight function. An element in the space $\mathcal{C}(V \times V)$ of real-valued continuous functions over $V \times V$ is called a *kernel*. A kernel $K$ is *symmetric* if $K(x, y) = K(y, x)$ for all $x, y \in V$. It is *positive* if it is symmetric and if, for any $m \in \mathbb{N}$ and for any $x_1, \ldots, x_m \in V$, the matrix $\big(K(x_i, x_j)\big)_{i,j=1}^{m}$ is positive semidefinite. The *weighted theta prime number* of $G$ is defined as

$$
\begin{aligned}
\vartheta'_w(G) = \inf \quad & M \\
& K - w^{1/2} \otimes (w^{1/2})^* && \text{is a positive kernel,} \\
& K(x, x) \leqslant M && \text{for all } x \in V, \\
& K(x, y) \leqslant 0 && \text{for all } \{x, y\} \notin E \text{ where } x \neq y, \\
& M \in \mathbb{R}, \; K \in \mathcal{C}(V \times V) && \text{is symmetric.}
\end{aligned}
\qquad (2.4)
$$

One may show, mimicking the proof of Theorem 1.1, that $\vartheta'_w(G) \geqslant \alpha_w(G)$.

Let $G = (V, E)$ be the spherical cap packing graph as defined in Section 1.1. We will use the symmetry of this graph to show that (2.4) gives the sharpest bound obtainable by Theorem 1.2.

The orthogonal group $\mathrm{O}(n)$ acts on $S^{n-1}$, and this defines the action of $\mathrm{O}(n)$ on the vertex set $V = S^{n-1} \times \{1, \ldots, N\}$ by $A(x, i) = (Ax, i)$ for $A \in \mathrm{O}(n)$. The group average of a kernel $K \in \mathcal{C}(V \times V)$ is given by

$$\overline{K}((x, i), (y, j)) = \int_{\mathrm{O}(n)} K(A(x, i), A(y, j)) \, d\mu(A),$$

where $\mu$ is the Haar measure on $\mathrm{O}(n)$ normalized so that $\mu(\mathrm{O}(n)) = 1$. If $(K, M)$ is feasible for (2.4), then $(\overline{K}, M)$ is feasible too. This follows since, for each $A \in \mathrm{O}(n)$, a point $(x, i)$ has the same weight as $A(x, i)$, and two points $(x, i)$ and $(y, j)$ are adjacent if and only if $A(x, i)$ and $A(y, j)$ are adjacent. Since $(K, M)$

and $(\overline{K}, M)$ have the same objective value $M$, and since $\overline{K}$ is invariant under the action of $O(n)$, we may restrict consideration to $O(n)$-invariant kernels (that is, kernels $K$ such that $K(Au, Av) = K(u, v)$ for all $A \in O(n)$ and $u, v \in V$) in finding the infimum of (2.4).

Schoenberg [36] showed that a symmetric kernel $K \in \mathcal{C}(S^{n-1} \times S^{n-1})$ is positive and $O(n)$-invariant if and only if it lies in the cone spanned by the kernels $(x, y) \mapsto P_k^n(x \cdot y)$. We will use the following generalization for kernels over $V \times V$.

THEOREM 2.1. *A symmetric kernel $K \in \mathcal{C}(V \times V)$, with $V = S^{n-1} \times \{1, \dots, N\}$, is positive and $O(n)$-invariant if and only if*

$$K((x, i), (y, j)) = f_{ij}(x \cdot y) \tag{2.5}$$

*with*

$$f_{ij}(u) = \sum_{k=0}^{\infty} f_{ij,k} P_k^n(u),$$

*where $(f_{ij,k})_{i,j=1}^{N}$ is positive semidefinite for all $k \geqslant 0$ and $\sum_{k=0}^{\infty} |f_{ij,k}| < \infty$ for all $i, j = 1, \dots, N$, implying in particular that we have uniform convergence above.*

Before we prove the theorem, we apply it to simplify problem (2.4). If $K$ is an $O(n)$-invariant feasible solution of (2.4), then $K - w^{1/2} \otimes (w^{1/2})^*$ is a positive $O(n)$-invariant kernel, and hence it can be written in the form (2.5). Using in addition that $P_0^n = 1$, problem (2.4) reduces to

$$\vartheta'_w(G) = \inf \quad M$$
$$f_{ii}(0) + w(\alpha_i) \leqslant M \qquad \text{for all } 1 \leqslant i \leqslant N,$$
$$f_{ij}(u) + (w(\alpha_i)w(\alpha_j))^{1/2} \leqslant 0 \quad \text{when } -1 \leqslant u \leqslant \cos(\alpha_i + \alpha_j),$$
$$M \in \mathbb{R} \text{ and } (f_{ij,k})_{i,j=1}^{N} \text{ positive semidefinite for all } k \geqslant 0.$$

By substituting $f_{ij,0} - (w(\alpha_i)w(\alpha_j))^{1/2}$ for $f_{ij,0}$, we see that the solution to this problem indeed equals the sharpest bound given by Theorem 1.2.

*Proof of Theorem 2.1.* If we endow the space $\mathcal{C}(S^{n-1})$ of real-valued continuous function on the unit sphere $S^{n-1}$ with the usual $L^2$ inner product, then, for $f$, $g \in \mathcal{C}(V)$,

$$\langle f, g \rangle = \sum_{i=1}^{N} \int_{S^{n-1}} f(x, i) g(x, i) \, d\omega(x)$$

gives an inner product on $\mathcal{C}(V)$. The space $\mathcal{C}(S^{n-1})$ decomposes orthogonally as

$$\mathcal{C}(S^{n-1}) = \bigoplus_{k=0}^{\infty} H_k,$$

where $H_k$ is the space of homogeneous harmonic polynomials of degree $k$ restricted to $S^{n-1}$. With

$$H_{k,i} = \{f \in \mathcal{C}(V) : \text{there is a } g \in H_k \text{ such that } f(\cdot, j) = \delta_{ij} g(\cdot)\},$$

it follows that $\mathcal{C}(V)$ decomposes orthogonally as

$$\mathcal{C}(V) = \bigoplus_{k=0}^{\infty} \bigoplus_{i=1}^{N} H_{k,i}.$$

Given the action of $O(n)$ on $V$, we have the natural unitary representation on $\mathcal{C}(V)$ given by $(Af)(x, i) = f(A^{-1}x, i)$ for $A \in O(n)$ and $f \in \mathcal{C}(V)$. It follows that each space $H_{k,i}$ is $O(n)$-irreducible and that two spaces $H_{k,i}$ and $H_{k',i'}$ are $O(n)$-equivalent if and only if $k = k'$. Let

$$\{e_{k,i,l} : k \geqslant 0, 1 \leqslant i \leqslant N, \text{ and } 1 \leqslant l \leqslant h_k\}$$

be a complete orthonormal system of $\mathcal{C}(V)$ such that $e_{k,i,1}, \ldots, e_{k,i,h_k}$ is a basis of $H_{k,i}$. By Bochner's characterization [5], a kernel $K \in \mathcal{C}(V \times V)$ is positive and $O(n)$-invariant if and only if

$$K((x, i), (y, j)) = \sum_{k=0}^{\infty} \sum_{i',j'=1}^{N} f_{ij,k} \sum_{l=1}^{h_k} e_{k,i',l}(x, i) e_{k,j',l}(y, j), \qquad (2.6)$$

where each $\left(f_{ij,k}\right)_{i,j=1}^{N}$ is positive semidefinite and $\sum_{k=0}^{\infty} |f_{ij,k}| < \infty$ for all $i, j$.

By the addition formula (see Ch. 9.6 of Andrews *et al.* [1]), we have

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

for any orthonormal basis $e_{k,1}, \ldots, e_{k,h_k}$ of $H_k$. It follows that

$$\sum_{l=1}^{h_k} e_{k,i',l}(x, i) e_{k,j',l}(y, j) = \delta_{ii'} \delta_{jj'} \frac{h_k}{\omega_n(S^{n-1})} P_k^n(x \cdot y),$$

and substituting this into (2.6) completes the proof. $\qquad \square$

Bochner's characterization for the kernel $K$, which we used above, usually assumes that the spaces under consideration are homogeneous, so that the decompositions into isotypic irreducible spaces are guaranteed to be finite. This finiteness is then used to conclude uniform convergence. Since the action of $\mathrm{O}(n)$ on $V$ is not transitive, we do not immediately have this guarantee. We can still use the characterization, however, since irreducible subspaces of $\mathcal{C}(V)$ have finite multiplicity.

## 3. Translational packings of bodies and multiple-size sphere packings

Before giving a proof of Theorem 1.3, we briefly present some technical considerations regarding density. Here, we follow closely Appendix A of Cohn and Elkies [8].

Let $\mathcal{K}_1, \ldots, \mathcal{K}_N$ be convex bodies, and let $\mathcal{P}$ be a packing of translated copies of $\mathcal{K}_1, \ldots, \mathcal{K}_N$; that is, $\mathcal{P}$ is a union of translated copies of the bodies, any two copies having disjoint interiors. We say that the *density* of $\mathcal{P}$ is $\Delta$ if for all $p \in \mathbb{R}^n$ we have

$$\Delta = \lim_{r \to \infty} \frac{\mathrm{vol}(B(p, r) \cap \mathcal{P})}{\mathrm{vol}\, B(p, r)},$$

where $B(p, r)$ is the ball of radius $r$ centered at $p$. Not every packing has a density, but every packing has an *upper density* given by

$$\limsup_{r \to \infty} \sup_{p \in \mathbb{R}^n} \frac{\mathrm{vol}(B(p, r) \cap \mathcal{P})}{\mathrm{vol}\, B(p, r)}.$$

We say that a packing $\mathcal{P}$ is *periodic* if there is a lattice $L \subseteq \mathbb{R}^n$ that leaves $\mathcal{P}$ invariant, that is, which is such that $\mathcal{P} = x + \mathcal{P}$ for all $x \in L$. In other words, a periodic packing consists of some translated copies of the bodies $\mathcal{K}_1, \ldots, \mathcal{K}_N$ arranged inside the fundamental parallelotope of $L$, and this arrangement repeats itself at each copy of the fundamental parallelotope translated by vectors of the lattice.

It is easy to see that a periodic packing has a density. This is particularly interesting for us, since in computing upper bounds for the maximum possible density of a packing we may restrict ourselves to periodic packings, as it is known (and also easy to see) that the supremum of the upper densities of packings can be approximated arbitrarily well by periodic packings (see Appendix A in Cohn and Elkies [8]).

To provide a proof of the theorem, we need another fact from harmonic analysis, the Poisson summation formula. Let $f \colon \mathbb{R}^n \to \mathbb{C}$ be a Schwartz function, and let $L \subseteq \mathbb{R}^n$ be a lattice. The *Poisson summation formula* states that,

for every $x \in \mathbb{R}^n$,

$$\sum_{v \in L} f(x + v) = \frac{1}{\text{vol}(\mathbb{R}^n/L)} \sum_{u \in L^*} \hat{f}(u) e^{2\pi i u \cdot x},$$

where $L^* = \{u \in \mathbb{R}^n : u \cdot x \in \mathbb{Z} \text{ for all } x \in L\}$ is the *dual lattice* of $L$ and where $\text{vol}(\mathbb{R}^n/L)$ is the volume of a fundamental domain of the lattice $L$.

*Proof of Theorem 1.3.* As observed above, we may restrict ourselves to periodic packings. Let then $L \subseteq \mathbb{R}^n$ be a lattice, and let $x_1, \ldots, x_m \in \mathbb{R}^n$ and $r : \{1, \ldots, m\} \to \{1, \ldots, N\}$ be such that

$$\mathcal{P} = \bigcup_{v \in L} \bigcup_{i=1}^{m} v + x_i + \mathcal{K}_{r(i)}$$

is a packing. This means that, whenever $i \neq j$ or $v \neq 0$, bodies $x_i + \mathcal{K}_{r(i)}$ and $v + x_j + \mathcal{K}_{r(j)}$ have disjoint interiors. This packing is periodic and therefore has a well-defined density, which equals

$$\frac{1}{\text{vol}(\mathbb{R}^n/L)} \sum_{i=1}^{m} \text{vol}\,\mathcal{K}_{r(i)}.$$

Consider the sum

$$\sum_{v \in L} \sum_{i,j=1}^{m} (\text{vol}\,\mathcal{K}_{r(i)})^{1/2} (\text{vol}\,\mathcal{K}_{r(j)})^{1/2} f_{r(i)r(j)}(v + x_j - x_i). \tag{3.1}$$

Applying the Poisson summation formula, we may express (3.1) in terms of Fourier transform of $f$, obtaining

$$\frac{1}{\text{vol}(\mathbb{R}^n/L)} \sum_{u \in L^*} \sum_{i,j=1}^{m} (\text{vol}\,K_{r(i)})^{1/2} (\text{vol}\,K_{r(j)})^{1/2} \hat{f}_{r(i)r(j)}(u) e^{2\pi i u \cdot (x_j - x_i)},$$

where $L^*$ is the dual lattice of $L$.

Since $f$ satisfies condition (ii) of the theorem, matrix $\left(\hat{f}_{r(i)r(j)}(u)\right)_{i,j=1}^{m}$ is positive semidefinite for every $u \in \mathbb{R}^n$. So the inner sum above is always nonnegative. If we then consider only the summand for $u = 0$, we see that (3.1)

is at least

$$\frac{1}{\text{vol}(\mathbb{R}^n/L)} \sum_{i,j=1}^{m} (\text{vol}\,\mathcal{K}_{r(i)})^{1/2} (\text{vol}\,\mathcal{K}_{r(j)})^{1/2} \hat{f}_{r(i)r(j)}(0)$$

$$\geqslant \frac{1}{\text{vol}(\mathbb{R}^n/L)} \sum_{i,j=1}^{m} \text{vol}\,\mathcal{K}_{r(i)} \,\text{vol}\,\mathcal{K}_{r(j)} \tag{3.2}$$

$$= \frac{1}{\text{vol}(\mathbb{R}^n/L)} \left( \sum_{i=1}^{m} \text{vol}\,\mathcal{K}_{r(i)} \right)^2,$$

where the inequality comes from condition (i) of the theorem.

Now, notice that, whenever $v \neq 0$ or $i \neq j$, one has $f_{r(i)r(j)}(v + x_j - x_i) \leqslant 0$. Indeed, since $\mathcal{P}$ is a packing, if $v \neq 0$ or $i \neq j$, then the bodies $x_i + \mathcal{K}_{r(i)}$ and $v + x_j + \mathcal{K}_{r(j)}$ have disjoint interiors. But then also $\mathcal{K}_{r(i)}$ and $v + x_j - x_i + \mathcal{K}_{r(j)}$ have disjoint interiors, and then from (iii) we see that $f_{r(i)r(j)}(v + x_j - x_i) \leqslant 0$.

From this observation, we see immediately that (3.1) is at most

$$\sum_{i=1}^{m} \text{vol}\,\mathcal{K}_{r(i)} f_{r(i)r(i)}(0) \leqslant \max\{f_{ii}(0) : i = 1, \dots, N\} \sum_{i=1}^{m} \text{vol}\,\mathcal{K}_{r(i)}. \tag{3.3}$$

So (3.1) is at least (3.2) and at most (3.3). Putting this all together, we get that

$$\frac{1}{\text{vol}(\mathbb{R}^n/L)} \sum_{i=1}^{m} \text{vol}\,\mathcal{K}_{r(i)} \leqslant \max\{f_{ii}(0) : i = 1, \dots, N\},$$

proving the theorem. $\square$

We mentioned in the beginning of the section that Theorem 1.3 is an analog of the weighted theta prime number for a certain infinite graph. The connection will become clearer after we present a slightly more general version of Theorem 1.3.

An $L^\infty$ function $f \colon \mathbb{R}^n \to \mathbb{C}^{N \times N}$ is said to be of *positive type* if $f(x) = f(-x)^*$ for all $x \in \mathbb{R}^n$, and if for all $L^1$ functions $\rho \colon \mathbb{R}^n \to \mathbb{C}^N$ we have

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \rho(y)^* f(x - y) \rho(x) \, dx \, dy \geqslant 0.$$

When $N = 1$, we have the classical theory of functions of positive type (see for example the book by Folland [17] for background). Many useful properties of such functions can be extended to the matrix-valued case (that is, to the $N > 1$ case) via a simple observation: a function $f \colon \mathbb{R}^n \to \mathbb{C}^{N \times N}$ is of positive type if and only if for all $p \in \mathbb{C}^N$ the function $g_p \colon \mathbb{R}^n \to \mathbb{C}$ such that

$$g_p(x) = p^* f(x) p$$

is of positive type.

From this observation, two useful classical characterizations of functions of positive type can be extended to the matrix-valued case. The first one is useful when dealing with continuous functions of positive type. It states that a continuous and bounded function $f\colon \mathbb{R}^n \to \mathbb{C}^{N\times N}$ is of positive type if and only if, for every choice $x_1, \ldots, x_m$ of finitely many points in $\mathbb{R}^n$, the block matrix $\left(f(x_i - x_j)\right)_{i,j=1}^{m}$ is positive semidefinite.

The second characterization is given in terms of the Fourier transform. It states that an $L^1$ function $f\colon \mathbb{R}^n \to \mathbb{C}^{N\times N}$ is of positive type if and only if the matrix $\left(\hat{f}_{ij}(u)\right)_{i,j=1}^{N}$ is positive semidefinite for all $u \in \mathbb{R}^n$. So, in the statement of Theorem 1.3, for instance, one could replace condition (i) by the equivalent condition that $f$ be a function of positive type.

When $N = 1$, the previous two characterizations of functions of positive type date back to Bochner [6].

With this, we may give an alternative and more general version of Theorem 1.3.

THEOREM 3.1. *Let $\mathcal{K}_1, \ldots, \mathcal{K}_N$ be convex bodies in $\mathbb{R}^n$, and let $f\colon \mathbb{R}^n \to \mathbb{R}^{N\times N}$ be a continuous and $L^1$ function. Suppose that $f$ satisfies the following conditions.*

(i) *The matrix $\left(\hat{f}_{ij}(0) - (\mathrm{vol}\,\mathcal{K}_i)^{1/2}(\mathrm{vol}\,\mathcal{K}_j)^{1/2}\right)_{i,j=1}^{N}$ is positive semidefinite.*

(ii) *$f$ is of positive type.*

(iii) *$f_{ij}(x) \leqslant 0$ whenever $\mathcal{K}_i^{\circ} \cap (x + \mathcal{K}_j^{\circ}) = \emptyset$.*

*Then the density of every packing of translates of $\mathcal{K}_1, \ldots, \mathcal{K}_N$ in the Euclidean space $\mathbb{R}^n$ is at most $\max\{f_{ii}(0) : i = 1, \ldots, N\}$.*

Let $V = \mathbb{R}^n \times \{1, \ldots, N\}$. Notice that the kernel $K\colon V \times V \to \mathbb{R}$ such that

$$K((x, i), (y, j)) = f_{ij}(x - y),$$

implicitly defined by the function $f$, plays the same role as the matrix $K$ from the definition of the theta prime number (see Section 2.2). For instance, this is a positive kernel, since $f$ is of positive type, and hence for any $L^1$ function $\rho\colon V \to \mathbb{R}$ we have that

$$\int_V \int_V K((x, i), (y, j))\rho(x, i)\rho(y, j)\,d(x, i)\,d(y, j) \geqslant 0.$$

Theorem 3.1 can then be seen as an analog of the weighted theta prime number for the packing graph with vertex set $V$ that we consider.

When one reads through the proof of Theorem 1.3, the one step that fails when $f$ is $L^1$ instead of Schwartz is the use of the Poisson summation formula. Indeed, sum (3.1) is no longer well defined in such a situation. The summation formula also holds, however, under somewhat different conditions that are just what we need to make the proof go through. The proof of the following lemma makes use of the well-known interpretation of the Poisson summation formula as a trace formula, which for instance is explained by Terras [**39**, Ch. 1.3].

LEMMA 3.2. *Let* $f\colon \mathbb{R}^n \to \mathbb{C}^{N\times N}$ *be a continuous function of bounded support and positive type. Then, for every lattice* $L \subseteq \mathbb{R}^n$, *every* $x \in \mathbb{R}^n$, *and all* $i$, $j = 1$, $\ldots, N$, *we have*

$$\sum_{v\in L} f_{ij}(x+v) = \frac{1}{\mathrm{vol}(\mathbb{R}^n/L)} \sum_{u\in L^*} \hat{f}_{ij}(u) e^{2\pi i u\cdot x}.$$

*Proof.* Since each function $f_{ij}$ is continuous and of bounded support, the functions $g_{ij}\colon \mathbb{R}^n/L \to \mathbb{C}$ such that

$$g_{ij}(x) = \sum_{v\in L} f_{ij}(x+v)$$

are continuous. Indeed, the sum above is well defined, being in fact a finite sum (since $f_{ij}$ has bounded support), and therefore $g_{ij}$ can be seen locally as a sum of finitely many continuous functions.

Let us now compute the Fourier transform of $g_{ij}$. For $u \in L^*$, we have that

$$\begin{aligned}
\hat{g}_{ij}(u) &= \int_{\mathbb{R}^n/L} g_{ij}(x) e^{-2\pi i u\cdot x}\, dx \\
&= \int_{\mathbb{R}^n/L} \sum_{v\in L} f_{ij}(x+v) e^{-2\pi i u\cdot x}\, dx \\
&= \int_{\mathbb{R}^n} f_{ij}(x) e^{-2\pi i u\cdot x}\, dx \\
&= \hat{f}_{ij}(u).
\end{aligned}$$

So we know that

$$g_{ij}(x) = \frac{1}{\mathrm{vol}(\mathbb{R}^n/L)} \sum_{u\in L^*} \hat{f}_{ij}(u) e^{2\pi i u\cdot x} \tag{3.4}$$

in the sense of $L^2$ convergence. Our goal is to prove that pointwise convergence also holds above.

To this end, we consider for $i = 1, \ldots, N$ the kernel $K_i \colon (\mathbb{R}^n/L) \times (\mathbb{R}^n/L) \to \mathbb{C}$ such that

$$K_i(x, y) = \sum_{v \in L} f_{ii}(v + x - y).$$

Since each function $f_{ii}$ is of bounded support and continuous, each kernel $K_i$ is continuous. Since for each $i$ we have that $f_{ii}(x) = \overline{f_{ii}(-x)}$ for all $x \in \mathbb{R}^n$ (since $f$ is of positive type), each kernel $K_i$ is self-adjoint. Notice that the functions $x \mapsto (\mathrm{vol}(\mathbb{R}^n/L))^{-1/2} e^{2\pi i u \cdot x}$, for $u \in L^*$, form a complete orthonormal system of $L^2(\mathbb{R}^n/L)$. Each such function is also an eigenfunction of $K_i$, with eigenvalue $\hat{f}_{ii}(u)$. Indeed, we have

$$\int_{\mathbb{R}^n/L} K_i(x, y)(\mathrm{vol}(\mathbb{R}^n/L))^{-1/2} e^{2\pi i u \cdot y} \, dy$$

$$= (\mathrm{vol}(\mathbb{R}^n/L))^{-1/2} \int_{\mathbb{R}^n/L} \sum_{v \in L} f_{ii}(v + x - y) e^{2\pi i u \cdot y} \, dy$$

$$= (\mathrm{vol}(\mathbb{R}^n/L))^{-1/2} \int_{\mathbb{R}^n} f_{ii}(x - y) e^{2\pi i u \cdot y} \, dy$$

$$= (\mathrm{vol}(\mathbb{R}^n/L))^{-1/2} \int_{\mathbb{R}^n} f_{ii}(y) e^{2\pi i u \cdot (x-y)} \, dy$$

$$= \hat{f}_{ii}(u)(\mathrm{vol}(\mathbb{R}^n/L))^{-1/2} e^{2\pi i u \cdot x}.$$

Since $f$ is of positive type, the matrices of Fourier transforms $\left(\hat{f}_{ij}(u)\right)_{i,j=1}^N$, for $u \in \mathbb{R}^n$, are all positive semidefinite. In particular, this implies that the Fourier transforms of $f_{ii}$, for $i = 1, \ldots, N$, are nonnegative. So we see that each $K_i$ is a continuous and positive kernel. Mercer's theorem (see for instance Courant and Hilbert [11]) then implies that $K_i$ is trace-class, its trace being the sum of all its eigenvalues. So, for each $i = 1, \ldots, N$, the series

$$\sum_{u \in L^*} \hat{f}_{ii}(u) \tag{3.5}$$

converges, and, since each summand is nonnegative, it converges absolutely.

Suppose now that $i, j = 1, \ldots, N$ are such that $i \neq j$. Since the matrices of Fourier transforms are nonnegative, for all $u \in \mathbb{R}^n$, we have that the matrix

$$\begin{pmatrix} \hat{f}_{ii}(u) & \hat{f}_{ij}(u) \\ \overline{\hat{f}_{ij}(u)} & \hat{f}_{jj}(u) \end{pmatrix}$$

is positive semidefinite, and this in turn implies that $|\hat{f}_{ij}(u)|^2 \leqslant \hat{f}_{ii}(u) \hat{f}_{jj}(u)$ for all $u \in \mathbb{R}^n$. Using then the convergence of the series (3.5) and the

Cauchy–Schwarz inequality, one gets

$$\sum_{u \in L^*} |\hat{f}_{ij}(u)| \leqslant \sum_{u \in L^*} (\hat{f}_{ii}(u)\hat{f}_{jj}(u))^{1/2} \leqslant \left(\sum_{u \in L^*} \hat{f}_{ii}(u)\right)^{1/2} \left(\sum_{u \in L^*} \hat{f}_{jj}(u)\right)^{1/2},$$

and we see that, in fact, for all $i, j = 1, \ldots, N$ the series

$$\sum_{u \in L^*} \hat{f}_{ij}(u)$$

converges absolutely.

This convergence result shows that the sum in (3.4) converges absolutely and uniformly for all $x \in \mathbb{R}^n/L$. This means that the function defined by this sum is a continuous function, and, since $g_{ij}$ is also a continuous function, and in (3.4) we have convergence in the $L^2$ sense, we must also then have pointwise convergence, as we aimed to establish.                                                      □

With this we may give a proof of Theorem 3.1.

*Proof of Theorem 3.1.* Using Lemma 3.2, we may repeat the proof of Theorem 1.3 given before, proving the theorem for continuous functions of bounded support. To extend the proof also to continuous $L^1$ functions, we use the following trick.

Let $f: \mathbb{R}^n \to \mathbb{R}^{N \times N}$ be a continuous and $L^1$ function satisfying the hypothesis of the theorem. For each $T > 0$, consider the function $g^T: \mathbb{R}^n \to \mathbb{R}^{N \times N}$ defined such that

$$g^T(x) = \frac{\mathrm{vol}(B(0, T) \cap B(x, T))}{\mathrm{vol}\, B(0, T)} f(x),$$

where $B(p, T)$ is the ball of radius $T$ centered at $p$.

It is easy to see that $g^T$ is a continuous function of bounded support. It is also clear that it satisfies condition (iii) from the statement of the theorem. We now show that $g^T$ is a function of positive type; that is, it satisfies condition (ii).

For this, pick any points $x_1, \ldots, x_m \in \mathbb{R}^n$. Let $\chi_i: \mathbb{R}^n \to \{0, 1\}$ be the characteristic function of $B(x_i, T)$, and denote by $\langle \phi, \psi \rangle$ the standard inner product between functions $\phi$ and $\psi$ in the Hilbert space $L^2(\mathbb{R}^n)$. Then

$$\begin{aligned}
g^T(x_i - x_j) &= \frac{\mathrm{vol}(B(0, T) \cap B(x_i - x_j, T))}{\mathrm{vol}\, B(0, T)} f(x_i - x_j) \\
&= \frac{\mathrm{vol}(B(x_i, T) \cap B(x_j, T))}{\mathrm{vol}\, B(0, T)} f(x_i - x_j) \\
&= \frac{\langle \chi_i, \chi_j \rangle}{\mathrm{vol}\, B(0, T)} f(x_i - x_j).
\end{aligned}$$

This shows that the matrix $\left(g^T(x_i - x_j)\right)_{i,j=1}^m$ is positive semidefinite, being the Hadamard product, i.e. entrywise product, of two positive semidefinite matrices. We therefore have that $g^T$ is of positive type.

Now, $g^T$ is a continuous function of positive type and bounded support, satisfying condition (iii). It is very possible, however, that $g^T$ does not satisfy condition (i), and so the conclusion of the theorem may not apply to $g^T$. Let us now fix this problem.

Notice that $g_{ij}^T$ converges pointwise to $f_{ij}$ as $T \to \infty$. Moreover, for all $T > 0$ we have $|g_{ij}^T(x)| \leqslant |f_{ij}(x)|$. It then follows from Lebesgue's dominated convergence theorem that $\hat{g}_{ij}^T(0) \to \hat{f}_{ij}(0)$ as $T \to \infty$. This means that there exists a number $T_0 > 0$ such that for each $T \geqslant T_0$ we may pick a number $\alpha(T) \geqslant 1$ so that the function $h^T : \mathbb{R}^n \to \mathbb{C}^{N \times N}$ such that

$$
\begin{aligned}
h_{ii}^T(x) &= \alpha(T) g_{ii}^T(x) && \text{for } i = 1, \ldots, N, \\
h_{ij}^T(x) &= g_{ij}^T(x) && \text{for } i, j = 1, \ldots, N \quad \text{with } i \neq j
\end{aligned}
$$

for all $x \in \mathbb{R}^n$ satisfies condition (i). We may moreover pick the numbers $\alpha(T)$ in such a way that $\lim_{T \to \infty} \alpha(T) = 1$.

It is also easy to see that each function $h^T$ is of positive type and bounded support and satisfies condition (iii). Hence the conclusion of the theorem applies for each $h^T$, and so for every $T \geqslant T_0$ we see that

$$
M_T = \max\{h_{ii}^T(0) : i = 1, \ldots, N\}
$$

is an upper bound for the density of any packing of translated copies of $\mathcal{K}_1, \ldots, \mathcal{K}_N$. But then, since $g_{ii}^T(0) = f_{ii}(0)$ for all $T \geqslant 0$, and since $\lim_{T \to \infty} \alpha(T) = 1$, we see that

$$
\max\{f_{ii}(0) : i = 1, \ldots, N\} = \lim_{T \to \infty} M_T,
$$

finishing the proof. $\qquad\square$

## 4. Computations for binary spherical cap packings

In this section and in the next section, we describe how we obtained the numerical results of Sections 1.2 and 1.3. Our approach is computational: to apply Theorems 1.2 and 1.3, we use techniques from semidefinite programming and polynomial optimization.

We start by briefly discussing the case of binary spherical cap packings. Next, we will discuss the more computationally challenging case of binary sphere packings.

It is a classical result of Lukács (see, for example, Theorem 1.21.1 in Szegö [**38**]) that a real univariate polynomial $p$ of degree $2d$ is nonnegative on the

interval $[a, b]$ if and only if there are real polynomials $q$ and $r$ such that $p(x) = (q(x))^2 + (x - a)(b - x)(r(x))^2$. This characterization is useful when we combine it with the elementary but powerful observation (discovered independently by several authors; see Laurent [**28**]) that a real univariate polynomial $p$ of degree $2d$ is a sum of squares of polynomials if and only if $p(x) = v(x)^\mathsf{T} Q v(x)$ for some positive semidefinite matrix $Q$, where $v(x) = (1, x, \ldots, x^d)$ is a vector whose components are the monomial basis.

Let $\alpha_1, \ldots, \alpha_N \in (0, \pi]$ be angles, and let $d$ be an integer. Write $v_0(x) = (1, x, \ldots, x^d)$ and $v_1(x) = (1, x, \ldots, x^{d-1})$. Using this characterization together with Theorem 1.2, we see that the optimal value of the following optimization problem gives an upper bound for the density of a packing of spherical caps with angles $\alpha_1, \ldots, \alpha_N$.

**Problem A.** For $k = 0, \ldots, 2d$, find positive semidefinite matrices $\left(f_{ij,k}\right)_{i,j=1}^N$, and, for $i, j = 1, \ldots, N$, find $(d+1) \times (d+1)$ positive semidefinite matrices $Q_{ij}$ and $d \times d$ positive semidefinite matrices $R_{ij}$ that minimize

$$\max\left\{ \sum_{k=0}^{2d} f_{ii,k} : i = 1, \ldots, N \right\}$$

and are such that

$$\left( f_{ij,0} - w(\alpha_i)^{1/2} w(\alpha_j)^{1/2} \right)_{i,j=1}^N$$

is positive semidefinite and the polynomial identities

$$\sum_{k=0}^{2d} f_{ij,k} P_k^n(u) + \langle Q_{ij}, v_0(u) v_0(u)^\mathsf{T} \rangle$$
$$+ \langle R_{ij}, (u+1)(\cos(\alpha_i + \alpha_j) - u) v_1(u) v_1(u)^\mathsf{T} \rangle = 0 \quad (4.1)$$

are satisfied for $i, j = 1, \ldots, N$. ◁

Above, $\langle A, B \rangle$ denotes the trace inner product between matrices $A$ and $B$. Problem A is a semidefinite programming problem, as the polynomial identities (4.1) can each be expressed as $2d + 1$ linear constraints on the entries of the matrices involved. Indeed, to check that a polynomial is identically zero, it suffices to check that the coefficient of each monomial $1, x, \ldots, x^{2d}$ is zero, and for each such monomial we get a linear constraint.

In the above, we work with the standard monomial basis $1, x, \ldots, x^{2d}$, but we could use any other basis of the space of polynomials of degree at most $2d$, both to define the vectors $v_0$ and $v_1$ and to check the polynomial identity (4.1). Such a change of basis does not change the problem from a formal point of view, but it

can drastically improve the performance of the solvers used. In our computations for binary spherical cap packings it was enough to use the standard monomial basis. We will see in the next section, when we present our computations for the Euclidean space, that a different choice of basis is essential.

We reported in Section 1.2 on our calculations for $N = 1$, and 2 and $n = 3$, 4, and 5. The bounds, for the angles under consideration, do not seem to improve beyond $d = 25$, so we use this value for $d$ in all computations. To obtain these bounds, we used the solver SDPA-QD, which works with quadruple precision floating-point numbers, from the SDPA family [**19**]. To generate the input for the solver, we wrote a SAGE [**37**] program using SDPSL [**34**]. Our programs are contained in the examples folder in the source distribution of SDPSL; version 0.85 is available as a supplementary file (available at http://dx.doi.org/10.1017/fms.20 14.24) to this article, and the latest version can be obtained from the SDPSL [**34**] website.

## 5. Computations for binary sphere packings

In this section, we discuss our computational approach to find upper bounds for the density of binary sphere packings using Theorem 1.3. This is a more difficult application of semidefinite programming and polynomial optimization techniques than the one described in Section 4.

It is often the case in applications of sum of squares techniques that, if one formulates the problems carelessly, high numerical instability invalidates the final results, or even numerical results cannot easily be obtained. This raises questions of how to improve the formulations used and the precision of the computations, so that we may provide *rigorous* bounds. We also address these questions and, since the techniques we use and develop might be of interest to the reader who wants to perform computations in polynomial optimization, we include some details.

**5.1. Theorem 1.3 for multiple-size sphere packings.** In the case of multiple-size sphere packings, Theorem 1.3 can be simplified. The key observation here is that, when all the bodies $\mathcal{K}_i$ are spheres, then condition (iii) depends only on the norm of the vector $x$. More specifically, if each $\mathcal{K}_i$ is a sphere of radius $r_i$, then $\mathcal{K}_i^\circ \cap (x + \mathcal{K}_j^\circ) = \emptyset$ if and only if $\|x\| \geqslant r_i + r_j$.

So in Theorem 1.3 one can choose to restrict oneself to radial functions. A function $f : \mathbb{R}^n \to \mathbb{C}$ is *radial* if the value of $f(x)$ depends only on the norm of $x$. If $f : \mathbb{R}^n \to \mathbb{C}$ is radial, for $t \geqslant 0$ we denote by $f(t)$ the common value of $f$ for vectors of norm $t$.

The Fourier transform $\hat{f}(u)$ of a radial function $f$ also depends only on the norm of $u$; in other words, the Fourier transform of a radial function is also radial.

By restricting ourselves to radial functions, we obtain the following version of Theorem 1.3.

THEOREM 5.1. *Let* $r_1, \ldots, r_N > 0$, *and let* $f \colon \mathbb{R}^n \to \mathbb{R}^{N \times N}$ *be a matrix-valued function whose every component* $f_{ij}$ *is a radial Schwartz function. Suppose that* $f$ *satisfies the following conditions.*

(i) *The matrix* $\left( \hat{f}_{ij}(0) - (\mathrm{vol}\, B(r_i))^{1/2} (\mathrm{vol}\, B(r_j))^{1/2} \right)_{i,j=1}^{N}$ *is positive semidefinite, where* $B(r)$ *is the ball of radius* $r$ *centered at the origin.*

(ii) *The matrix of Fourier transforms* $\left( \hat{f}_{ij}(t) \right)_{i,j=1}^{N}$ *is positive semidefinite for every* $t > 0$.

(iii) $f_{ij}(w) \leqslant 0$ *if* $w \geqslant r_i + r_j$, *for* $i, j = 1, \ldots, N$.

*Then the density of any packing of spheres of radii* $r_1, \ldots, r_N$ *in the Euclidean space* $\mathbb{R}^n$ *is at most* $\max\{f_{ii}(0) : i = 1, \ldots, N\}$.

One might ask whether the restriction to radial functions worsens the bound of Theorem 1.3. For spheres, this is not the case. Indeed, suppose that each body $\mathcal{K}_i$ is a sphere. If $f \colon \mathbb{R}^n \to \mathbb{R}^{N \times N}$ is a function satisfying the conditions of the theorem, then its radialized version, the function

$$\overline{f}(x) = \int_{S^{n-1}} f(\|x\|\xi)\, d\omega_n(\xi),$$

also satisfies the conditions of the theorem, and it provides the same upper bound. This shows in particular that, for the case of multiple-size sphere packings, Theorem 5.1 is equivalent to Theorem 1.3.

**5.2. A semidefinite programming formulation.** To simplify notation, and because it is the case of our main interest, we now take $N = 2$. Everything in the following also goes through for arbitrary $N$, with obvious modifications.

To find a function $f$ satisfying the conditions of Theorem 5.1, we specify $f$ via its Fourier transform. Let $d \geqslant 0$ be an odd integer, and consider the even function $\varphi \colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}^{2 \times 2}$ such that

$$\varphi_{ij}(t) = \sum_{k=0}^{d} a_{ij,k} t^{2k},$$

where each $a_{ij,k}$ is a real number and $a_{ij,k} = a_{ji,k}$ for all $k$. We set the Fourier transform of $f$ to be

$$\hat{f}_{ij}(u) = \varphi_{ij}(\|u\|) e^{-\pi \|u\|^2}.$$

Notice that each $\hat{f}_{ij}$ is a Schwartz function, so its Fourier inverse is also Schwartz.

The reason why we choose this form for the Fourier transform of $f$ is that it makes it simple to compute $f$ from its Fourier transform by using the following result.

LEMMA 5.2. We have that

$$\int_{\mathbb{R}^n} \|u\|^{2k} e^{-\pi\|u\|^2} e^{2\pi i u \cdot x} \, du = k!\, \pi^{-k} e^{-\pi\|x\|^2} L_k^{n/2-1}(\pi\|x\|^2), \qquad (5.1)$$

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

For background on Laguerre polynomials, we refer the reader to the book by Andrews *et al.* [1].

*Proof.* With $f(u) = \|u\|^{2k} e^{-\pi\|u\|^2}$, the left-hand side of (5.1) is equal to $\hat{f}(-x)$. By [1, Theorem 9.10.3], we have

$$\hat{f}(-x) = 2\pi\|x\|^{1-n/2} \int_0^\infty s^{2k} e^{-\pi s^2} J_{n/2-1}(2\pi s\|x\|) s^{n/2} \, ds,$$

where $J_{n/2-1}$ is the Bessel function of the first kind with parameter $n/2 - 1$. Using [1, Corollary 4.11.8], we see that this is equal to

$$\pi^{-k} \frac{\Gamma(k + n/2)}{\Gamma(n/2)} e^{-\pi\|x\|^2} {}_1F_1\binom{-k}{n/2}; \pi\|x\|^2\big), \qquad (5.2)$$

where ${}_1F_1$ is a hypergeometric series.

By [1, (6.2.2)], we have

$${}_1F_1\binom{-k}{n/2}; \pi\|x\|^2\big) = \frac{k!}{(n/2)_k} L_k^{n/2-1}(\pi\|x\|^2),$$

where $(n/2)_k = (n/2)(1 + n/2) \cdots (k - 1 + n/2)$.

By substituting this in (5.2), and using the property that $\Gamma(x+1) = x\Gamma(x)$ for all $x \neq 0, -1, -2, \ldots$, we obtain the right-hand side of (5.1), as desired. $\square$

So we have

$$f_{ij}(x) = \int_{\mathbb{R}^n} \varphi_{ij}(\|u\|) e^{-\pi\|u\|^2} e^{2\pi i u \cdot x} \, du = \sum_{k=0}^d a_{ij,k}\, k!\, \pi^{-k} e^{-\pi\|x\|^2} L_k^{n/2-1}(\pi\|x\|^2).$$

Notice that it becomes clear that $f_{ij}$ is indeed real valued, as required by the theorem.

Consider the polynomial

$$p(t) = \sum_{k=0}^{d} a_k t^{2k}.$$

According to Lemma 5.2, if $g(x)$ is the Fourier inverse of $\hat{g}(u) = p(\|u\|)e^{-\pi\|u\|^2}$, then $g(\|x\|) = q(\|x\|)e^{-\pi\|x\|^2}$, where

$$q(w) = \sum_{k=0}^{d} a_k \, k! \, \pi^{-k} L_k^{n/2-1}(\pi w^2)$$

is a univariate polynomial. We denote the polynomial $q$ above by $\mathcal{F}^{-1}[p]$. Notice that $\mathcal{F}^{-1}[p]$ is obtained from $p$ via a linear transformation; that is, its coefficients are linear combinations of the coefficients of $p$. With this notation, we have

$$f_{ij}(x) = \mathcal{F}^{-1}[\varphi_{ij}](\|x\|)e^{-\pi\|x\|^2}.$$

Let

$$\sigma(t, y_1, y_2) = \sum_{i,j=1}^{2} \sum_{k=0}^{d} a_{ij,k} t^{2k} y_i y_j. \tag{5.3}$$

If this polynomial is a sum of squares, then it is nonnegative everywhere, and hence the matrices $\big(\varphi_{ij}(t)\big)_{i,j=1}^{2}$ are positive semidefinite for all $t \geqslant 0$. This implies that $f$ satisfies condition (ii) of Theorem 5.1. (The converse is also true: that, if the matrices $\big(\varphi_{ij}(t)\big)_{i,j=1}^{2}$ are positive semidefinite for all $t \geqslant 0$, then $\sigma$ is a sum of squares; For a proof see Choi *et al.* [7]. This fact is related to the Kalman–Yakubovich–Popov lemma in systems and control; see the discussion in Aylward *et al.* [2].)

Moreover, we may recover $\varphi$, and hence $\hat{f}$, from $\sigma$. Indeed, we have

$$\varphi_{11}(t) = \sigma(t, 1, 0),$$
$$\varphi_{22}(t) = \sigma(t, 0, 1), \quad \text{and}$$
$$\varphi_{12}(t) = (1/2)(\sigma(t, 1, 1) - \sigma(t, 1, 0) - \sigma(t, 0, 1)).$$

So we can express condition (i) of Theorem 5.1 in terms of $\sigma$. We may also express condition (iii) in terms of $\sigma$, since it can be translated as

$$\mathcal{F}^{-1}[\varphi_{ij}](w) \leqslant 0 \quad \text{for all } w \geqslant r_i + r_j \quad \text{and} \quad i, j = 1, 2 \quad \text{with } i \leqslant j. \tag{5.4}$$

If we find a polynomial $\sigma$ of the form (5.3) that is a sum of squares, is such that

$$\big(\varphi_{ij}(0) - (\operatorname{vol} B(r_i))^{1/2}(\operatorname{vol} B(r_j))^{1/2}\big)_{i,j=1}^{2} \tag{5.5}$$

is positive semidefinite, and satisfies (5.4), then the density of a packing of spheres of radii $r_1$ and $r_2$ is upper bounded by

$$\max\{\mathcal{F}^{-1}[\varphi_{11}](0), \mathcal{F}^{-1}[\varphi_{22}](0)\}.$$

We may encode conditions (5.4) in terms of sums of squares polynomials (see Section 4), and therefore we may encode the problem of finding a $\sigma$ as above as a semidefinite programming problem, as we show now.

Let $P_0, P_1, \ldots$ be a sequence of univariate polynomials, where polynomial $P_k$ has degree $k$. Consider the vector of polynomials $v$, which has entries indexed by $\{0, \ldots, \lfloor d/2 \rfloor\}$ given by

$$v(t)_k = P_k(t^2)$$

for $k = 0, \ldots, \lfloor d/2 \rfloor$. We also write $V(t) = v(t)v(t)^{\mathsf{T}}$.

Consider also the vector of polynomials $m$ with entries indexed by $\{1, 2\} \times \{0, \ldots, \lfloor d/2 \rfloor\}$ given by

$$m(t, y_1, y_2)_{i,k} = P_k(t^2)y_i$$

for $i, j = 1, 2$, and $k = 0, \ldots, \lfloor d/2 \rfloor$.

Since $\sigma$ is an even polynomial, it is a sum of squares if and only if there are positive semidefinite matrices $S_0, S_1 \in \mathbb{R}^{(d+1)\times(d+1)}$ such that

$$\sigma(t, y_1, y_2) = \langle S_0, m(t, y_1, y_2)m(t, y_1, y_2)^{\mathsf{T}} \rangle + \langle S_1, t^2 m(t, y_1, y_2)m(t, y_1, y_2)^{\mathsf{T}} \rangle.$$

From the matrices $S_0$ and $S_1$, we may then recover $\varphi_{ij}$ and also $\mathcal{F}^{-1}[\varphi_{ij}]$. A more convenient way for expressing $\varphi_{ij}$ in terms of $S_0$ and $S_1$ is as follows. Consider the matrices

$$Y_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Y_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad Y_{12} = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}.$$

Then

$$\varphi_{ij}(t) = \langle S_0, V(t) \otimes Y_{ij} \rangle + \langle S_1, t^2 V(t) \otimes Y_{ij} \rangle$$

and

$$\mathcal{F}^{-1}[\varphi_{ij}](w) = \langle S_0, \mathcal{F}^{-1}[V(t)](w) \otimes Y_{ij} \rangle + \langle S_1, \mathcal{F}^{-1}[t^2 V(t)](w) \otimes Y_{ij} \rangle,$$

where $\mathcal{F}^{-1}$, when applied to a matrix, is applied to each entry individually.

With this, we may consider the following semidefinite programming problem for finding a polynomial $\sigma$ satisfying the conditions we need.

**Problem B.** Find $(d+1) \times (d+1)$ real positive semidefinite matrices $S_0$ and $S_1$, and $(\lfloor d/2 \rfloor + 1) \times (\lfloor d/2 \rfloor + 1)$ real positive semidefinite matrices $Q_{11}$, $Q_{22}$,

and $Q_{12}$ that minimize

$$\max\{\langle S_0, \mathcal{F}^{-1}[V(t)](0) \otimes Y_{11}\rangle + \langle S_1, \mathcal{F}^{-1}[t^2 V(t)](0) \otimes Y_{11}\rangle,$$
$$\langle S_0, \mathcal{F}^{-1}[V(t)](0) \otimes Y_{22}\rangle + \langle S_1, \mathcal{F}^{-1}[t^2 V(t)](0) \otimes Y_{22}\rangle\}$$

and are such that

$$\left(\langle S_0, V(0) \otimes Y_{ij}\rangle - (\operatorname{vol} B(r_i))^{1/2}(\operatorname{vol}(B(r_j))^{1/2}\right)_{i,j=1}^2$$

is positive definite and the polynomial identities

$$\langle S_0, \mathcal{F}^{-1}[V(t)](w) \otimes Y_{ij}\rangle + \langle S_1, \mathcal{F}^{-1}[t^2 V(t)](w) \otimes Y_{ij}\rangle$$
$$+ \langle Q_{ij}, (w^2 - (r_i + r_j)^2)V(w)\rangle = 0 \tag{5.6}$$

are satisfied for $i, j = 1, 2$ and $i \leqslant j$. ◁

Any solution to this problem gives us a polynomial $\sigma$ of the shape (5.3) which is a sum of squares and satisfies conditions (5.4) and (5.5), and so the optimal value is an upper bound for the density of any packing of spheres of radius $r_1$ and $r_2$. There might be, however, polynomials $\sigma$ satisfying these conditions that cannot be obtained as feasible solutions to Problem B, since condition (5.6) is potentially more restrictive than condition (5.4) (compare Problem B above with the result of Lukács mentioned in Section 4). In our practical computations, this restriction was not problematic, and we found very good functions.

Observe also that Problem B is really a semidefinite programming problem. Indeed, the polynomial identities in (5.6) can each be represented as $d + 1$ linear constraints in the entries of the matrices $S_i$ and $Q_{ij}$. This is the case because testing whether a polynomial is identically zero is the same as testing whether each monomial has a zero coefficient and so, since all our polynomials are even and of degree $2d$, we need only check if the coefficients of the monomials $x^{2k}$ are zero for $k = 0, \ldots, d$.

## 5.3. Numerical results.

When solving Problem B, we need to choose a sequence $P_0, P_1, \ldots$ of polynomials. A choice which works well in practice is

$$P_k(t) = \mu_k^{-1} L_k^{n/2-1}(2\pi t),$$

where $\mu_k$ is the absolute value of the coefficient of $L_k^{n/2-1}(2\pi t)$ with largest absolute value. We observed in practice that the standard monomial basis performs poorly.

To represent the polynomial identities in (5.6) as linear constraints, we may check that each monomial $x^{2k}$ of the resulting polynomial has coefficient zero. We may use, however, any basis of the space of even polynomials of degree at most $2d$ to represent such identities. Given such a basis, we expand each polynomial in it and check that the expansion has only zero coefficients. The basis we use to represent the identities is $P_0(t^2)$, $P_1(t^2)$, ..., $P_d(t^2)$, which we observed to work much better than $t^0$, $t^2$, ..., $t^{2d}$. Notice that no extra variables are necessary if we use a different basis to represent the identities. We need only keep, for each polynomial in the matrices $\mathcal{F}^{-1}[V(t)](w)$, $\mathcal{F}^{-1}[t^2 V(t)](w)$, $w^2 V(w)$, and $V(w)$, its expansion in the basis we want to use.

The plot in Figure 3 was generated by solving Problem B with $d = 31$ using the solver SDPA-GMP from the SDPA family [19]. To generate the solver input, we wrote a SAGE [37] program using SDPSL [34] working with floating-point arithmetic and precision of 256 bits; see the examples in the source distribution of SDPSL for the source code. For each dimension 2, ..., 5, we solved Problem B with $r_1 = r/1000$ and $r_2 = 1$ for $r = 200, 201, \ldots, 1000$; the reason we start with $r = 200$ is that for smaller values of $r$ the solver runs into numerical stability problems. We also note that the solver has failed to solve some of the problems, and these points have been ignored when generating the plot. The number of problems that could not be solved was small though: for $n = 2$, all problems could be solved; for $n = 3$, there were 6 failures; for $n = 4$, we had 18 failures; and finally, for $n = 5$, the solver failed for 137 problems.

With our methods we can achieve higher values for $d$, but we noticed that the bound does not improve much after $d = 31$. For instance, in dimension 2 for $r_1 = 1/2$ and $r_2 = 1$, we obtain the bound $0.9174466\ldots$ for $d = 31$ and the bound $0.9174426\ldots$ for $d = 51$.

<center>* * *</center>

In the previous account of how the plot in Figure 3 was generated, we swept under the rug all precision issues. We generate the data for the solver using floating-point arithmetic, and the solver also uses floating-point arithmetic. We cannot therefore be sure that the optimal value found by the solver gives a valid bound at all.

If we knew *a priori* that Problem B is strictly feasible (that is, that it admits a solution in which the matrices $S_i$ and $Q_{ij}$ are positive definite), and if we had some control over the dual solutions, then we could use semidefinite programming duality to argue that the bounds we compute are rigorous; see, for instance, Gijswijt [18, Ch. 7.2] for an application of this approach in coding theory. However, we do not know that Problem B is strictly feasible; neither do we have knowledge about the dual solutions. In fact, most of our approach to provide rigorous bounds consists in finding a strictly feasible solution.

A naive idea to turn the bound returned by the solver into a rigorous bound would be to simply project a solution returned by the solver onto the subspace given by the constraints in (5.6). If the original solution is of good quality, then this would yield a feasible solution.

There are two problems with this approach, though. The first problem is that the matrices returned by the solver will have eigenvalues too close to zero, and therefore after the projection they might not be positive semidefinite anymore. We discuss how to handle this issue below.

The second problem is that to obtain a rigorous bound one would need to perform the projection using symbolic computations and rational arithmetic, and the computational cost is just too big. For instance, we failed to do so even for $d = 7$.

Our approach avoids projecting the solution using symbolic computations. Here is an outline of our method.

(1) Obtain a solution to the problem with objective value close to the optimal value returned by the solver, but in which every matrix $S_i$ and $Q_{ij}$ is positive definite by a good margin and the maximum violation of the linear constraints is very small.

(2) Approximate matrices $S_i$ and $Q_{ij}$ by rational positive semidefinite matrices $\bar{S}_i$ and $\bar{Q}_{ij}$ having minimum eigenvalues at least $\lambda_i$ and $\mu_{ij}$, respectively.

(3) Compute a bound on how much constraints (5.6) are violated by $\bar{S}_i$ and $\bar{Q}_{ij}$ using rational arithmetic. If the maximum violation of the constraints is small compared to the bounds $\lambda_i$ and $\mu_{ij}$ on the minimum eigenvalues, then we may be sure that the solution can be changed into a feasible solution without changing its objective value too much.

We now explain how each step above can be accomplished.

First, most likely the matrices $S_i$, $Q_{ij}$ returned by the solver will have eigenvalues very close to zero, or even slightly negative due to the numerical method which might allow infeasible steps.

To obtain a solution with positive definite matrices, we may use the following trick (see Löfberg [31]). We solve Problem B to find its optimal value, say $z^*$. Then we solve a feasibility version of Problem B in which the objective function is absent, but we add a constraint to ensure that

$$\max\{\langle S_0, \mathcal{F}^{-1}[V(t)](0) \otimes Y_{11}\rangle + \langle S_1, \mathcal{F}^{-1}[t^2 V(t)](0) \otimes Y_{11}\rangle,$$
$$\langle S_0, \mathcal{F}^{-1}[V(t)](0) \otimes Y_{22}\rangle + \langle S_1, \mathcal{F}^{-1}[t^2 V(t)](0) \otimes Y_{22}\rangle\} \leqslant z^* + \eta,$$

where $\eta > 0$ should be small enough so that we do not jeopardize the objective value of the solution, but not too small so that a good strictly feasible solution exists. (We take $\eta = 10^{-5}$, which works well for the purpose of making a plot.) The trick here is that most semidefinite programming solvers, when solving a feasibility problem, will return a positive definite solution — the analytical center — if one can be found.

This partially addresses step (1), because, though the solution we find will be positive definite, it might violate the linear constraints too much. To quickly obtain a solution that violates the constraints only slightly, we may project our original solution onto the subspace given by constraints (5.6) using floating-point arithmetic of high enough precision. If the solution returned by the solver had good precision to begin with, then the projected solution will still be positive definite.

As an example, for our problems with $d = 31$, SDPA-GMP returns solutions that violate the constraints by at most $10^{-30}$. By doing a projection using floating-point arithmetic with 256 bits of precision in SAGE, we can bring the violation down to about $10^{-70}$ without affecting much the eigenvalues of the matrices.

So we have addressed step (1). For step (2), we observe that simply converting the floating-point matrices $S_i$, $Q_{ij}$ to rational matrices would work, but then we would be in trouble to estimate the minimum eigenvalues of the resulting rational matrices in a rigorous way. Another idea of how to make the conversion is as follows.

Say we want to approximate floating-point matrix $A$ by a rational matrix $\bar{A}$. We start by computing numerically an approximation to the least eigenvalue of $A$. Say $\tilde{\lambda}$ is this approximation. We then use binary search in the interval $[\tilde{\lambda}/2, \tilde{\lambda}]$ to find the largest $\lambda$ so that the matrix $A - \lambda I$ has a Cholesky decomposition; this we do using floating-point arithmetic of high enough precision. If we have this largest $\lambda$, then

$$A = LL^{\mathsf{T}} + \lambda I,$$

where $L$ is the Cholesky factor of $A - \lambda I$. Then we approximate $L$ by a rational matrix $\bar{L}$, and we set

$$\bar{A} = \bar{L}\bar{L}^{\mathsf{T}} + \lambda I,$$

obtaining thus a rational approximation of $A$ and a bound on its minimum eigenvalue.

Our idea for step (3) is to compare the maximum violation of constraints (5.6) with the minimum eigenvalues of the matrices. To formalize this idea, suppose that constraints (5.6) are slightly violated by $\bar{S}_i$, $\bar{Q}_{ij}$. So, for instance, we have

$$\langle \bar{S}_0, \mathcal{F}^{-1}[V(t)](w) \otimes Y_{11} \rangle + \langle \bar{S}_1, \mathcal{F}^{-1}[t^2 V(t)](w) \otimes Y_{11} \rangle$$
$$+ \langle \bar{Q}_{11}, (w^2 - (2r_1)^2)V(w) \rangle = p, \tag{5.7}$$

where $p$ is an even polynomial of degree at most $2d$. Notice that we may compute an upper bound on the absolute values of the coefficients of $p$ using rational arithmetic.

To fix this constraint, we may distribute the coefficients of $p$ in the matrices $\bar{S}_0$ and $\bar{Q}_{11}$ (a very similar idea was presented by Löfberg [30]). To make things precise, for $k = 1, \ldots, d$, write

$$i(k) = \min\{\lfloor d/2 \rfloor, k - 1\},$$
$$j(k) = k - 1 - i(k).$$

Pairs $(i(k), j(k))$ correspond to entries of the matrix $V(w)$. Notice that the polynomial $(w^2 - (2r_1)^2)V(w)_{i(k)j(k)}$ has degree $2k$.

So the polynomials

$$R_0 = \mathcal{F}^{-1}[V(t)_{00}](w),$$
$$R_1 = (w^2 - (2r_1)^2)V(w)_{i(1)j(1)},$$
$$\vdots$$
$$R_d = (w^2 - (2r_1)^2)V(w)_{i(d)j(d)}$$

form a basis of the space of even polynomials of degree at most $2d$. We may then express our polynomial $p$ in this basis as

$$p = \alpha_0 R_0 + \cdots + \alpha_d R_d.$$

Now, we subtract $\alpha_0$ from $(\bar{S}_0)_{(1,0),(1,0)}$ and $\alpha_k$ from $(\bar{Q}_{11})_{i(k)j(k)}$, for $k = 1, \ldots, d$. The resulting matrices satisfy constraint (5.7), and, as long as the $\alpha_k$ are small enough, they should remain positive semidefinite. More precisely, it suffices to require that $d \, \|(\alpha_1, \ldots, \alpha_d)\|_\infty \leqslant \mu_{11}$ and $|\alpha_0| \leqslant \lambda_0$.

There are two issues to note in our approach. The first one is that it has to be applied again twice to fix the other two constraints in (5.6). The applications do not conflict with each other: in each one we change a different matrix $\bar{Q}_{ij}$ and different entries of $\bar{S}_0$. We have to be careful though that we consider the changes to $\bar{S}_0$ simultaneously in order to check that it remains positive semidefinite.

The second issue is how to find bounds on the coefficients $\alpha_k$, since computing them explicitly using symbolic computation is infeasible. One way to do this is to consider the basis change matrix between the bases $x^{2k}$, for $k = 0, \ldots, d$, and $R_0, \ldots, R_d$, which we denote by $U$. Then we know that

$$\|(\alpha_0, \ldots, \alpha_d)\|_\infty \leqslant \|U^{-1}\|_\infty \|p\|_\infty,$$

where $\|p\|_\infty$ is the $\infty$-norm of the vector of coefficients of $p$ in the basis $x^{2k}$.

So, if we have an upper bound for $\|U^{-1}\|_\infty$, we are done. To quickly find such an upper bound, we use an algorithm of Higham [**25**] (see also Higham [**24**]) which works for triangular matrices, like $U$. This bound proved to be good enough for our purposes.

## 6. Improving sphere packing bounds

We now prove Theorem 1.4 and show how to use it in order to compute the bounds presented in Table 1.

*Proof of Theorem 1.4.* Let $x_1, \ldots, x_N \in \mathbb{R}^n$ and $L \subseteq \mathbb{R}^n$ be a lattice such that

$$\bigcup_{v \in L} \bigcup_{i=1}^{N} v + x_i + B$$

is a sphere packing, where $B$ is the ball of radius $1/2$ centered at the origin. We may assume that, if $i \neq j$ and $v \neq 0$, then the distance between the centers of $v + x_i + B$ and $x_j + B$ is greater than $1 + \varepsilon_m$. Indeed, we could discard all $x_i$ that lie at distance less than $1 + \varepsilon_m$ from the boundary of the fundamental parallelotope of $L$. If the fundamental parallelotope is big enough (and if it is not, we may consider a dilated version of $L$ instead), this will only slightly alter the density of the packing, and the resulting packing will have the desired property.

Consider the sum

$$\sum_{i,j=1}^{N} \sum_{v \in L} f(v + x_i - x_j). \tag{6.1}$$

Using the Poisson summation formula, we may rewrite it as

$$\frac{1}{\mathrm{vol}(\mathbb{R}^n/L)} \sum_{i,j=1}^{N} \sum_{u \in L^*} \hat{f}(u) e^{2\pi i u \cdot (x_i - x_j)}.$$

By discarding all summands in the inner sum above except the one for $u = 0$, we see that (6.1) is at least

$$\frac{N^2 \, \mathrm{vol} \, B}{\mathrm{vol}(\mathbb{R}^n/L)}.$$

For $k = 1, \ldots, m$, write $F_k = \{(i, j) : \|x_i - x_j\| \in [1 + \varepsilon_{k-1}, 1 + \varepsilon_k)\}$. Then we see that (6.1) is at most

$$N f(0) + \eta_1 |F_1| + \cdots + \eta_m |F_m|.$$

So we see that

$$\frac{N \operatorname{vol} B}{\operatorname{vol}(\mathbb{R}^n/L)} \leqslant f(0) + \eta_1 \frac{|F_1|}{N} + \cdots + \eta_m \frac{|F_m|}{N}.$$

Notice that the left-hand side above is exactly the density of our packing. Now, from the definition of $M(\varepsilon)$, it is clear that, for $k = 1, \ldots, m$, we have

$$\frac{|F_1|}{N} + \cdots + \frac{|F_k|}{N} \leqslant M(\varepsilon_k),$$

and the theorem follows. $\square$

To find good functions $f$ satisfying the conditions required by Theorem 1.4, we use the approach from Section 5. We fix an odd positive integer $d$ and specify $f$ via its Fourier transform, writing

$$\varphi(t) = \sum_{k=0}^{d} a_k t^{2k}$$

and setting

$$\hat{f}(u) = \varphi(\|u\|)e^{-\pi \|u\|^2}.$$

Using Lemma 5.2, we then have that

$$f(x) = \mathcal{F}^{-1}[\varphi](\|x\|)e^{-\pi \|x\|^2},$$

where

$$\mathcal{F}^{-1}[\varphi](w) = \sum_{k=0}^{d} a_k k! \, \pi^{-k} L_k^{n/2-1}(\pi w^2)$$

is a polynomial obtained as a linear transformation of $\varphi$.

Constraint (ii), requiring that $\hat{f}(u) \geqslant 0$ for all $u \in \mathbb{R}^n$, can be equivalently expressed as requiring that the polynomial $\varphi$ should be a sum of squares.

Recalling the result of Lukács mentioned in Section 4, one may also express constraint (iii) in terms of sums of squares: one simply has to require that there exist polynomials $p_0(w)$ and $q_0(w)$ such that

$$\mathcal{F}^{-1}[\varphi](w) = -(p_0(w))^2 - (w^2 - (1 + \varepsilon_m)^2)(q_0(w))^2.$$

In a similar way, one may express constraints (iv). For instance, for a given $k$, we require that there should exist polynomials $p_k(w)$ and $q_k(w)$ such that

$$\mathcal{F}^{-1}[\varphi](w)e^{-\pi(1+\varepsilon_{k-1})^2} - \eta_1 = -(p_k(w))^2 - (w - (1+\varepsilon_{k-1}))((1+\varepsilon_k) - w)(q_k(w))^2,$$

and this implies (iv).

So we may represent the constraints on $f$ in terms of sums of squares, and therefore also in terms of semidefinite programming, as we did in Sections 4 and 5. There is only the issue that now we want to find a function $f$ that satisfies constraints (i)–(iv) of the theorem and that minimizes the maximum in (1.2). This does not look like a linear objective function, but, since by linear programming duality this maximum is equal to

$$
\begin{aligned}
\min \quad & f(0) + y_1 U(\varepsilon_1) + \cdots + y_m U(\varepsilon_m) \\
& y_i + \cdots + y_m \geqslant \eta_i && \text{for } i = 1, \ldots, m, \\
& y_k \geqslant 0 && \text{for } k = 1, \ldots, m,
\end{aligned}
$$

we may transform our original problem into a single minimization semidefinite programming problem, the optimal value of which provides an upper bound for the densities of sphere packings.

It is still a question how to compute upper bounds for $M(\varepsilon)$. For this, we use upper bounds on the sizes of spherical codes. A *spherical code* with *minimum angular distance* $0 < \theta \leqslant \pi$ is a set $C \subseteq S^{n-1}$ such that the angle between any two distinct points in $C$ is at least $\theta$. In other words, a spherical code with minimum angular distance $\theta$ gives a packing of spherical caps with angle $\theta/2$. We denote by $A(n, \theta)$ the maximum cardinality of any spherical code in $S^{n-1}$ with minimum angular distance $\theta$.

For $\epsilon \leqslant (\sqrt{5} - 1)/2$, we have

$$
M(\varepsilon) \leqslant A(n, \arccos t(\varepsilon)), \quad \text{where } t(\varepsilon) = 1 - \frac{1}{2(1 + \varepsilon)^2}.
$$

This follows from the proof of [10, Lemma 4.1], which we replicate here for the convenience of the reader: suppose that $x, y \in \mathbb{R}^n$ are such that $\|x\|, \|y\| \in [1, 1 + \varepsilon]$ and $\|x - y\| \geqslant 1$. Then, by the law of cosines,

$$
\cos \angle(x, y) = \frac{\|x\|^2 + \|y\|^2 - \|x - y\|^2}{2\|x\| \|y\|} \leqslant \frac{\|x\|^2 + \|y\|^2 - 1}{2\|x\| \|y\|}.
$$

The right-hand side is convex as a function of $\|x\|$ or $\|y\|$ individually, so it is upper bounded by the maximal value at the four vertices of the square $[1, 1 + \epsilon]^2$. Since $\epsilon \leqslant (\sqrt{5} - 1)/2$, the maximum occurs for $\|x\| = \|y\| = 1 + \epsilon$, which gives the desired bound.

For the bounds of Table 1, we took $d = 31$. To compute upper bounds for $A(n, \theta)$, we used the semidefinite programming bound of Bachoc and Vallentin [4]. The bounds we used for computing Table 1 are given in Table 2.

Table 2. For each dimension considered in Table 1, we show here the sequence $\varepsilon_1 < \cdots < \varepsilon_m$ and the upper bounds $U(\varepsilon_k)$ used in our application of Theorem 1.4.

| Dimension | $(\varepsilon, U(\varepsilon))$ pairs |
|---|---|
| 3 | (0.022753, 12), (0.054092, 13), (0.082109, 14), (0.113864, 15) |
| 4 | (0.008097, 24), (0.017446, 25), (0.025978, 26), (0.036951, 27) |
| 5 | (0.003013, 45), (0.008097, 46), (0.013259, 47), (0.017446, 48) |
| 6 | (0.002006, 79), (0.004024, 80), (0.006054, 81), (0.008097, 82) |
| 7 | (0.001001, 136), (0.002006, 137), (0.003013, 138), (0.004024, 139), (0.005037, 140) |
| 9 | (0.003013, 373), (0.029233, 457), (0.030325, 459), (0.031421, 464), (0.032520, 468), (0.033622, 473) |

Finally, we mention that all numerical issues discussed in Section 5 also happen with the approach we sketched in this section. In particular, the choices of bases are important for the stability of the semidefinite programming problems involved. We use the same bases as described in Section 5 though, so we skip a detailed discussion here. Notice moreover that our bounds are rigorous, having been checked with the same approach as that described in Section 5.

We refrained from performing similar calculations for higher dimensions for two reasons. First, we expect that the improvements are only minor. Second, the computations of the upper bounds for $M(\varepsilon)$ in higher dimensions require substantially more time, as one needs to solve the semidefinite programs with a high accuracy solver; see Mittelmann and Vallentin [33].

## Supplementary material

Supplementary material is available at http://dx.doi.org/10.1017/fms.2014.24.

# References

[1] G. E. Andrews, R. Askey and R. Roy, 'Special functions', Encyclopedia of Mathematics and its Applications, 71 (Cambridge University Press, Cambridge, 1999).

[2] E. Aylward, S. Itani and P. A. Parrilo, 'Explicit SOS decompositions of univariate polynomial matrices and the Kalman–Yakubovich–Popov lemma', *Proceedings of the 46th IEEE Conference on Decision and Control* (2007).

[3] C. Bachoc, G. Nebe, F. M. de Oliveira Filho and F. Vallentin, 'Lower bounds for measurable chromatic numbers', *Geom. Funct. Anal.* **19** (2009), 645–661. arXiv:0801.1059.

[4] C. Bachoc and F. Vallentin, 'New upper bounds for kissing numbers from semidefinite programming', *J. Amer. Math. Soc.* **21** (2008), 909–924. arXiv:math/0608426.

[5] S. Bochner, 'Hilbert distances and positive definite functions', *Ann. of Math. (3)* **42** (1941), 647–656.

[6] S. Bochner, *Vorlesungen über Fouriersche Integrale* (Akademische Verlagsgesellschaft, Leipzig, 1932).

[7] M. D. Choi, T. Y. Lam and B. Reznick, 'Real zeros of positive semidefinite forms I', *Math. Z.* **171** (1980), 1–26.

[8] H. Cohn and N. D. Elkies, 'New upper bounds on sphere packings I', *Ann. of Math. (2)* **157** (2003), 689–714. arXiv:math/0110009.

[9] H. Cohn and A. Kumar, 'Universally optimal distribution of points on spheres', *J. Amer. Math. Soc.* **20** (2007), 99–148. arXiv:math/0607446.

[10] H. Cohn and A. Kumar, 'Optimality and uniqueness of the Leech lattice among lattices', *Ann. of Math.* (2) **170** (2009), 1003–1050. arXiv:math/0403263.

[11] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, 1953).

[12] P. Delsarte, J. M. Goethals and J. J. Seidel, 'Spherical codes and designs', *Geom. Dedicata* **6** (1977), 363–388.

[13] A. Florian, 'Ausfüllung der Ebene durch Kreise', *Rend. Circ. Mat. Palermo* **9** (1960), 300–312.

[14] A. Florian, 'Packing of incongruent circles on the sphere', *Monatsh. Math.* **133** (2001), 111–129.

[15] A. Florian, 'Remarks on my paper: packing of incongruent circles on the sphere', *Monatsh. Math.* **152** (2007), 39–43.

[16] A. Florian and A. Heppes, 'Packing circles of two different sizes on the sphere II', *Period. Math. Hungar.* **39** (1999), 125–127.

[17] G. B. Folland, 'A course in abstract harmonic analysis', Studies in Advanced Mathematics (CRC Press, Boca Raton, 1995).

[18] D. C. Gijswijt, 'Matrix algebras and semidefinite programming techniques for codes', PhD thesis, University of Amsterdam, 2005, arXiv:1007.0906.

[19] K. Fujisawa, M. Fukuda, K. Kobayashi, M. Kojima, K. Nakata, M. Nakata and M. Yamashita, SDPA (SemiDefinite Programming Algorithm) Users Manual — Version 7.0.5, Research Report B-448, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, 2008, http://sdpa.sourceforge.net.

[20] M. Grötschel, L. Lovász and A. Schrijver, 'The ellipsoid method and its consequences in combinatorial optimization', *Combinatorica* **1** (1981), 169–197.

[21] T. C. Hales, 'A proof of the Kepler conjecture', *Ann. of Math. (2)* **162** (2005), 1065–1185.

[22] A. Heppes, 'Some densest two-size disc packings in the plane', *Discrete Comput. Geom.* **30** (2003), 241–262.

[23]  A. Heppes and G. Kertész, 'Packing circles of two different sizes on the sphere: intuitive geometry', *Bolyai Soc. Math. Stud.* **6** (1997), 357–365.

[24]  N. J. Higham, 'A survey of condition number estimation for triangular matrices', *SIAM Rev.* **29** (1987), 575–595.

[25]  N. J. Higham, 'Upper bounds for the condition number of a triangular matrix', Numerical Analysis Report No. 86, University of Manchester, Manchester, 1983.

[26]  A. B. Hopkins, Y. Jiao, F. H. Stillinger and S. Torquato, 'Phase diagram and structural diversity of the densest binary sphere packings', *Phys. Rev. Lett.* **107** 125501 (2011), 5pp. arXiv:1108.2210.

[27]  A. B. Hopkins, F. H. Stillinger and S. Torquato, 'Densest binary sphere packings', *Phys. Rev. E* **85** 021130 (2012), 19pp. arXiv:1111.4917.

[28]  M. Laurent, 'Sums of squares, moment matrices and optimization over polynomials', in *Emerging Applications of Algebraic Geometry*, (eds M. Putinar and S. Sullivant), IMA Volumes in Mathematics and its Applications, 149 (Springer, New York, 2009), 157–270.

[29]  V. I. Levenshtein, 'Universal bounds for codes and designs', in *Handbook of Coding Theory,* Vol. I (North-Holland, Amsterdam, 1998), 499–648.

[30]  J. Löfberg, 'Pre- and post-processing sums-of-squares programs in practice', *IEEE Trans. Automat. Control* **54** (2009), 1007–1011.

[31]  J. Löfberg, 'Strictly feasible sums-of-squares solutions, post in YALMIP Wiki', 2011.

[32]  B. Masnick and J. Wolf, 'On linear unequal error protection codes', *IEEE Trans. Inform. Theory* **13** (1967), 600–607.

[33]  H. D. Mittelmann and F. Vallentin, 'High accuracy semidefinite programming bounds for kissing numbers', *Experiment. Math.* **19** (2010), 174–178. arXiv:0902.1105.

[34]  F. M. de Oliveira Filho, 'SDPSL: A semidefinite programming specification library', http://www.ime.usp.br/∼fmario/sdpsl.html.

[35]  C. A. Rogers, 'The packing of equal spheres', *Proc. London Math. Soc.* **8** (1958), 609–620.

[36]  I. J. Schoenberg, 'Positive definite functions on spheres', *Duke Math. J.* **9** (1942), 96–108.

[37]  W. A. Stein et al., 'Sage mathematics software (version 4.8), the sage development team', 2012, http://www.sagemath.org.

[38]  G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, XXIII (American Mathematical Society, Providence, 1975).

[39]  A. Terras, *Harmonic Analysis on Symmetric Spaces and Applications I* (Springer, Berlin, Heidelberg, New York, 1985).

[40]  S. Torquato, *Random Heterogeneous Materials, Microstructure and Macroscopic Properties* (Springer, New York, 2002).