OPTIMA 100

Mathematical Optimization Society Newsletter

MOS Chair's Column

May 15, 2016. In June 1980, the chair of the MPS Publications Committee, Michael Held, wrote "Thus the decision was made to establish a new Newsletter – OPTIMA." You can find the text of his article in *Optima* 1, available on the Web page www.mathopt.org/?nav=optima_details together with full copies of all 100 issues of *Optima*. It is great fun to click through the old copies – give it a try. And please join me in thanking the current and past Editors of *Optima*, Donald Hearn, issues 1–55 (!), Karen Aardal, issues 56–65, Jens Clausen, issues 66–72 , Andrea Lodi, issues 73–84, Katya Scheinberg, issues 85–93, and Volker Kaibel, issues 94–100. Their work has, of course, been strongly supported by a long line of Co-Editors, starting with Achim Bachem and continuing through the current team of Sam Burer and Jeff Linderoth. Many thanks to all of them.

The back issues of Optima provide great snapshots of the society's activities. Another set of snapshots are the ISMP program books. We have begun to collect pdf versions of the programs on www.mathopt.org/?nav=ismp. The plan is to eventually have documentation for all past symposia. In this effort, a big thanks goes to Art Geoffrion, for donating his collection of MOS material, including programs for 1967, 1973, 1976, and 1979. We will get these scanned during the coming weeks. If you have any old ISMP programs, particularly those before 1967, that you might be willing to donate or scan, please send me an email. A full set of programs on the MOS Web site would be a great resource for the society.

This summer brings to an end the current three-year cycle of the MOS. On July 18, 2016, Karen Aardal will take over as the society Chair and I will move to the role of Vice Chair. Also, Marina Epelman will take over as the MOS Treasurer, after a six-year run by Juan Meza. On the Publications Committee, Mike Juenger will take over as Chair, following the great work of Nick Gould. Finally, Sebastian Stiller will be the new Web Editor, following the long run of Marc Pfetsch. Thanks to our officers, both old and new.

Let me wrap up with a few quick remarks on the status of the society. We currently have 1,487 members. This follows our usual pattern: the membership goes up in the year following the ISMP, then declines in the following two years of the cycle. The society is in very good shape financially and we are well positioned to take on new activities. If you have something in mind, please send a note to any officer or member of the MOS Council.

As of March 31, 2016, the society has $794,610 in total assets, with $25,711 restricted for the Fulkerson Prize and $19,257 restricted for the Lagrange Prize. This is a substantial increase over our balance of $550,261 at the end of 2012. The MOS is a learned society, with no professional staff. Our main expense is an administrative fee paid to SIAM, to maintain our membership list, provide email service, and handle our financial matters. Over the past three years, the SIAM fee amounted to a total of $125,728. This was more than offset by our $155,966 royalty income over the three-year period, derived mainly from the publication of MPA, MPB, and MPC.

Thanks again to everyone for your support over the past three years. Let's all welcome Karen as the new chair of the society!

Bill Cook, University of Waterloo
bico@uwaterloo.ca

Contents of Issue 100 / May 2016

Frank Vallentin

Optimization in discrete geometry

Optimal structures in discrete geometry are fundamental to many areas in mathematics, physics, information theory, and materials science. Famous examples are densest packings of spheres, or minimal energy point configurations on Euclidean spheres.

Studying optimal structures we face two basic tasks: How to construct structures which are conjecturally optimal? How to prove that a given structure is indeed optimal?

For the first question researchers in mathematics and engineering have found many heuristics which often work well in practice. The second question is much harder, both from the mathematical as well as from the computational point of view. In this article a universal methodology is presented – based on a blend of tools coming from infinite-dimensional semidefinite optimization and harmonic analysis, together with computational techniques coming from real algebraic geometry and polynomial optimization – which frequently gives the best-known results.

1 Introduction

Many optimization problems in discrete geometry are concerned with the optimal distribution of finitely many points $X = \{x_1, \dots, x_N\}$ on a compact Riemannian manifold M. There are many possibilities to optimize the quality of such a geometric configuration X: One can maximize the packing density (or equivalently the packing radius), which is by far the best-studied example. Other important optimization problems in discrete geometry are minimizing potential energy, or minimizing covering density. In Section 2 we give concrete examples and explain in which areas they occur naturally.

These geometric optimization problems have the flavor of binary optimization problems, which occur frequently in classical combinatorial optimization: For every point $x \in M$ one has to make the binary decision whether x belongs to the finite set X or not.

On the one hand, the geometric setting is more difficult than the classical combinatorial setting, since the Riemannian manifold M contains infinitely many points. So one has to work with infinitely many binary decision variables and the optimization problems become infinite-dimensional. On the other hand, the geometric setting also has advantages: Usually the geometric structure of M is nice – it is smooth and it has many symmetries – and this one can exploit when performing the numerical optimization.

The goal of this article is to explain how one can generalize tools from finite-dimensional combinatorial optimization to this infinite-dimensional geometric setting. In Section 3 we define a hierarchy of increasingly stronger semidefinite programs which eventually solve the discrete geometry optimization problem, but which become more and more expensive to solve computationally. We will apply an infinite dimensional generalization of Lasserre's hierarchy introduced in [30] which, in the finite setting, exploits the duality between sum of squares of polynomials and the moment problem. In Section 4 we show how to solve the infinite-dimensional semidefinite programs. In particular we illustrate how the symmetry of the manifold M can be used to simplify the computation. We end with Section 5 where we report on recent results and pose open problems.

2 Optimization problems in discrete geometry

Before we explain our methods, we give a few concrete definitions of optimization problems in discrete geometry. Here we also show a few applications.

2.1 Packing

Let M be a compact metric space with metric d. Given a positive real number r, the *sphere packing problem* asks: How many pairwise nonoverlapping balls $B(x,r) = \{y \in M : d(x,y) \leq r\}$ of radius r can one fit into M. In other words, what is the constant

$$A(M,2r) = \max\{|X| : X \subseteq M, d(x,y) \geq 2r \text{ for } x,y \in X \text{ with } x \neq y\}\,?$$

In the theory of signal processing, especially when working with error correcting codes, the sphere packing problem in high dimensional manifolds is central. Already in his foundational paper from 1948 Claude E. Shannon established the close connection between dense sphere packings and good error correcting codes. In Shannon's model a message should be sent from a source to a destination via a communication channel. The channel is noisy: It may happen that the signal sent is corrupted by random noise. Error correcting codes help that in most cases this corruption can be repaired, so that error free communication is possible. The communication channel is modelled by a compact metric space M. Transmitter and receiver agree on a finite dictionary $X \subseteq M$ which the transmitter is allowed to send. When the transmitter sends $x \in X$ the channel corrupts x by random noise so that $x' \in M$ is received. Now it is natural for the receiver to find $x^* \in X$ so that $d(x',x^*) = \min\{d(x',y) : y \in X\}$. When X was a sphere packing with balls of radius r and when the noise was small ($d(x,x') < r$ holds), then one recovers the original message, i.e., one has $x^* = x$. Usually an estimate of the noise of the channel is known beforehand, meaning that one knows a bound for r. Now an important goal is to use the capacity of the channel. So finding a dense sphere packing X with $|X|$ as close as possible to $A(M,2r)$ is interesting.

There are different types of channels and they are modelled by different metric spaces M. For instance, the discrete binary symmetric channel can be modelled by $M = \mathbb{F}_2^n$, the n-dimensional vector space over the finite field with the two elements 0 and 1, and the Hamming distance $d(x,y) = |\{i \in \{1,\dots,n\} : x_i \neq y_i\}|$. The continuous symmetric channel with Gaussian white noise of given finite power can be modelled by the unit sphere S^{n-1} and error correcting codes in this model correspond to spherical codes. Constant weight codes are discrete analogs of spherical codes.

The innovation of new technological devices can make it necessary to consider "more exotic" manifolds. For example, when dealing with a multiple input multiple output scenario, we have m transmit antennas and m receive antennas, each of which can transmit or receive real vectors of length t. In this case, we have

$$x' = Hx + W \quad \text{with } x, x' \in \mathbb{R}^{m\times t}, H \in \mathbb{R}^{m\times m}, W \in \mathbb{R}^{m\times t},$$

where x is the transmitted signal, H is the random matrix where entry H_{ij} says how much of signal sent by transmit antenna j was received by antenna i, W is the random noise matrix, and x' is the received signal, see [56]. Then the row spaces of matrices x and Hx coincide with probability 1. So it makes sense to encode messages by using packings in the real Grassmannian manifold $\mathcal{G}_{m,t}$ where points represent m-dimensional subspaces in t-dimensional space $\mathbb{R}^t$. Distance between two subspaces P and $Q \in \mathcal{G}_{m,t}$ is measured by the chordal distance,

$$d_c(P,Q) = \left(\sum_{i=1}^{m} (\sin\theta_i)^2\right)^{1/2},$$

where $\theta_1,\dots,\theta_m \in [0,\pi/2]$ are the m principal angles defined by P and Q; θ_1 is the smallest angle between a line p_1 in P and a line q_1 in Q, θ_2 is the smallest angle between a line p_2 in the orthogonal complement of p_1 in P and a line q_2 in the orthogonal complement of q_1 in Q, θ_3 is the smallest angle between a line p_3 in the orthogonal complement of $p_1 + p_2$ in P and a line q_3 in the orthogonal complement of $q_1 + q_2$ in Q, etc.

Other geometric packing problems, which are non-compact, are translative packings of a centrally symmetric convex body $\mathcal{K}$. Here M is the n-dimensional Euclidean space $\mathbb{R}^n$, and the metric is given by the Minkowski functional of $\mathcal{K}$:

$$\|x\|_{\mathcal{K}} = \inf\{\lambda \in \mathbb{R}_{>0} : x \in \lambda\mathcal{K}\}.$$

When $\mathcal{K}$ is the n-dimensional unit ball, we speak about n-dimensional sphere packing.

Table 1 summarizes the packing problems we considered in this section.

2.2 Energy minimization

One can define the problem of minimizing the potential energy of a point configuration on any manifold. For the manifold being the unit sphere S^{n-1} the problem is defined as follows: Given a potential function $h : (0,4] \to \mathbb{R}$ (for example the Coulomb potential energy $t \mapsto 1/t^{n-2}$), the potential energy of N points $X = \{x_1,\dots,x_N\}$ on the unit sphere S^{n-1} is

$$\mathcal{E}_h(X) = \frac{1}{2}\sum_{i\neq j} h(\|x_i - x_j\|^2).$$

The goal is to arrange the N points in such a way that $\mathcal{E}_h(X)$ is minimized. In the limit, potential energy minimization specializes to maximizing packing density: If h is a hard-core potential, i.e. if it attains the value infinity from 0 up to a certain radius $2r$ and zero beyond it, then point configurations attaining minimal potential energy correspond to packings of spherical caps with radius r (measured in the Euclidean metric).

Potential energy minimization models physical particle systems, for instance colloidal particles in disperse media. A classical example of potential energy minimization is the Thomson problem which

asks for the minimal-energy configuration of N points on the unit sphere S^2 which interact via the Coulomb potential $1/r$ at Euclidean distance r.

Smale's 7th problem [47] deals with a logarithmic potential function, $h(t) = \frac{1}{\log \sqrt{t}}$. Smale asks whether one can efficiently find N points in $X \subseteq S^2$ so that

$$\mathcal{E}_h(X) - \min_{Y \subseteq S^2, |Y|=N} \mathcal{E}_h(Y) = O(\log N),$$

holds. The motivation for this question comes from finding good starting polynomials for homotopy algorithms for finding roots of complex polynomials.

It turns out that some configurations consisting of few points which are beautifully symmetric, like twelve points on S^2 forming the vertices of the regular icosahedron, are optimal for a huge class of natural potential functions. This was noted by Cohn and Kumar [7] who coined this phenomenon "universal optimality".

Definition 2.1. A point configuration $X \subseteq S^{n-1}$ is called *universally optimal* if it minimizes potential energy for all completely monotonic potential functions h, where a function $h : (0, 4] \to \mathbb{R}$ is called *completely monotonic* if its derivatives satisfy $(-1)^k h^{(k)}(x) \geq 0$ for all $x \in (0, 4]$ and $k \geq 0$, so that it is decreasing, convex, etc.

Energy minimizing point configurations on spheres have attracted mathematicians in fields such as approximation and coding theory, and biologists, chemists, and physicists in diverse fields such as viral morphology, crystallography, molecular structure and electrostatics. As engineers advance in gaining control of the microscopic and even nanoscopic world, energy minimization principles appear to become increasingly important for synthetic fabrication and design. Through the understanding of minimal energy configurations one can design nano-materials by a self-assembly process, shown for example by Whitesides, Kriebel, and Mayers [54]. In approximation theory, minimal energy points are used to discretize manifolds, see e.g. Hardin and Saff [22].

2.3 Covering

How can one distribute the N points X on the unit sphere S^{n-1} so that the *covering radius*

$$R(X) = \max_{y \in S^{n-1}} \min_{i=1,\dots,N} \|y - x_i\|$$

is minimized? This question of optimal coverings by spherical caps has applications in coding theory, for instance when quantizing n-dimensional Gaussian vectors with independent components (which with high probability lies near the surface of a sphere). Hardin, Sloane, and Smith [46] found point configurations which have a small covering radius by computer experiments. However, apart from asymptotic results for very large values of N (and besides the trivial volume bound), no method is known to find strong lower bounds for the covering radius.

3 Infinite-dimensional semidefinite optimization

The packing problems described above can be modeled as independent set problems in distance graphs. A *distance graph* $G = (V, E)$ is a graph where (V, d) is a metric space, and where there exists $D \subseteq (0, \infty)$ such that x and y are adjacent precisely when $d(x, y) \in D$. An *independent set* of an undirected graph $G = (V, E)$ is a subset of the vertex set which does not span an edge.

Now one is trying to find an independent set which is as large as possible. What "large" means depends on the situation. When the vertex set V is compact we can simply count and we use the *independence number*

$$\alpha(G) = \sup\{|I| : I \subseteq V, I \text{ is independent}\}.$$

In the non-compact translative body packing case one needs to use a density version of the independence number since maximal independent sets have infinite cardinality: The *(upper) point density* of an independent set $I \subset \mathbb{R}^n$ is

$$\delta(I) = \limsup_{R \to \infty} \frac{|I \cap [-R, R]^n|}{\mathrm{vol}([-R, R]^n)},$$

where $[-R, R]^n$ is the cube centered at the origin with side length $2R$. This measures the number of centers of bodies per unit volume. To determine the geometric density of the corresponding body packing we multiply $\delta(I)$ by the volume of the body $\mathcal{K}$.

Currently, these independent set problems have been solved only in a few special cases. One might expect that they will never be solved in full generality, for all parameters. Finding good lower bounds by constructions and good upper bounds by obstructions are both challenging tasks. Over the last years the best known results were achieved with computer assistance: Algorithms like the adaptive shrinking cell scheme of Torquato and Jiao [50] generate dense packings and give very good lower bounds. The combination of semidefinite programming and harmonic analysis often gives the best known upper bounds for these packing problems. This method originated from work of Hoffman [23], Delsarte [11], and Lovász [37].

3.1 Lasserre's hierarchy for finite graphs

Computing the independence number of a finite graph is an NP-hard problem as shown by Karp [26]. Approximating optimal solutions of NP-hard problems in combinatorial optimization with the help of linear and semidefinite optimization is a very wide and active area of research. The most popular semidefinite programming hierarchies for NP-hard combinatorial optimization problems are the Lovász-Schrijver hierarchy [38] (the N^+-operator) and the hierarchy of Lasserre [31]. Laurent [32] showed that Lasserre's hierarchy is stronger (at the same step) than the Lovász-Schrijver hierarchy.

We now give a formulation of Lasserre's hierarchy for computing the independence number of a finite graph. Here we follow Laurent [32].

Table 1. *Examples of packing problems*

Metric space M	Packing problem
$\mathbb{F}_2^n$, $d(x, y) = \lvert\{i : x_i \neq y_i\}\rvert$ Hamming distance	Binary codes
$\mathbb{F}_q^n$, d Hamming distance	q-ary codes
$\{x \in \mathbb{F}_q^n : d(0, x) = w\}$, d Hamming distance	Constant weight codes
$S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, $d(x, y) = \cos x \cdot y$ spherical distance	Spherical codes
$\mathcal{G}_{m,t}$, d_c chordal distance	Grassmannian codes
$\mathbb{R}P^{n-1} = \mathcal{G}_{1,n}$	Codes in real projective space
$\mathbb{R}^n$, $d(x, y) = \|x - y\|_{\mathcal{K}}$	Translative body packing

Definition 3.1. The *t-th step of Lasserre's hierarchy* of a finite graph $G = (V, E)$ is defined as

$$\mathrm{las}_t(G) = \max\Big\{ \sum_{x\in V} y_{\{x\}} : y \in \mathbb{R}^{I_{2t}}_{\geq 0},\ y_\emptyset = 1,\ M_t(y) \text{ is positive semidefinite}\Big\},$$

where I_t is the set of all independent sets with at most t elements and where $M_t(y) \in \mathbb{R}^{I_t \times I_t}$ is the moment matrix defined by the vector y: Its (J, J')-entry equals

$$(M_t(y))_{J,J'} = \begin{cases} y_{J\cup J'} & \text{if } J \cup J' \in I_{2t}, \\ 0 & \text{otherwise.} \end{cases}$$

One can show that the first step in Lasserre's hierarchy coincides with the ϑ'-number, the strengthened version of Lovász ϑ-number [37] which is due to Schrijver [43]; for a proof see for instance [44, Theorem 67.11].

Definition 3.2. The ϑ'-number of a finite graph $G = (V, E)$ is defined as

$$\vartheta'(G) = \max\Big\{ \sum_{x,y\in V} K(x,y) :\ K \in \mathbb{R}^{V\times V} \text{ is positive semidefinite, } \mathrm{Trace}(K) = 1,\ K(x,y) \geq 0 \text{ for } \{x,y\} \in E\Big\}$$

Furthermore, the hierarchy is a complete proof system, in the following sense:

Theorem 3.3. *The hierarchy converges to $\alpha(G)$ after at most $\alpha(G)$ steps:*

$$\vartheta'(G) = \mathrm{las}_1(G) \geq \mathrm{las}_2(G) \geq \ldots \geq \mathrm{las}_{\alpha(G)}(G) = \alpha(G).$$

Lasserre [31] showed that the hierachy converges in finitely many steps in the general setting of hierarchies for 0/1 polynomial optimization problems. For this he used Putinar's Positivstellensatz [40]. Laurent [32] gave an elementary proof based on combinatorial moment matrices and showed that the hierarchy converges in at most $\alpha(G)$ steps. She applied the fact that the cone of positive semidefinite moment matrices where rows and columns are indexed by the power set 2^V is a simplicial polyhedral cone; an observation due to Lindström [35] and Wilf [55]. More specifically using the inclusion-exclusion principle one can show:

Theorem 3.4. *The following equality holds:*

$$\Big\{ M \in \mathbb{R}^{2^V\times 2^V} : M \succeq 0,\ M \text{ is a moment matrix}\Big\} = \mathrm{cone}\{\chi_S \chi_S^{\mathsf{T}} : S \subseteq V\}, \quad (1)$$

where a moment matrix M is a matrix where the entry $M_{J,J'}$ only depends on the union $J \cup J'$ and where the vector $\chi_S \in \mathbb{R}^{2^V}$ is defined componentwise by

$$\chi_S(R) = \begin{cases} 1 & \text{if } R \subseteq S, \\ 0 & \text{otherwise.} \end{cases}$$

To set up a semidefinite programming hierarchy many variations are possible: For instance one can consider only "interesting" principal submatrices of the moment matrices to simplify the computation and one can also add more constraints coming from problem specific arguments. In fact, in the definition of $\mathrm{las}_t(G)$ we used the nonnegativity constraints $y_S \geq 0$ for $S \in I_{2t}$. Even without them, the convergence result holds, and the first step in the hierarchy coincides with the Lovász ϑ-number (without prime).

A rough classification for all these variations can be given in terms of *n-point bounds*. This refers to all variations which make use of variables y_S with $|S| \leq n$. An n-point bound is only capable of using obstructions coming from the local interaction of configurations having at most n points. For instance Lovász ϑ-number is a 2-point bound and the t-th step in Lasserre's hierarchy is a $2t$-point bound. The relation between n-point bounds and Lasserre's hierarchy was first made explicit by Laurent [33] in the case of bounds for binary codes.

3.2 Generalization of Lasserre's hierarchy to infinite graphs

In [29] de Laat and Vallentin generalized Lasserre's hierarchy to infinite graphs which arise in geometric packing problems. For this we consider topological packing graphs where vertices which are close are adjacent and where vertices which are adjacent will stay adjacent after slight perturbations:

Definition 3.5. A graph whose vertex set is a Hausdorff topological space is called a *topological packing graph* if each finite clique is contained in an open clique, where a *clique* is a subset of the vertices where every two vertices are adjacent.

For instance distance graphs G where D is open and D contains the interval $(0, \delta)$ for some $\delta > 0$ are topological packing graphs. That D contains an interval starting from 0 implies that vertices which are close are adjacent, and that D is open implies that adjacent vertices will stay adjacent after slight perturbations.

Now we introduce our generalization of Lasserre's hierarchy for compact topological packing graphs.

Before we go into some of the technical details we would like to comment on the choice of spaces in our generalization: In Lasserre's hierarchy for finite graphs the optimization variable y lies in the cone $\mathbb{R}^{I_{2t}}_{\geq 0}$. One might try to use the same cone when I_{2t} is uncountable, but then there are too many variables and it is impossible to express the objective function. At the other extreme one might try to restrict this cone to finitely (or countably) supported vectors, but then we do not know how to develop a duality theory. A duality theory is important for concrete computations: Minimization problems can be used to derive upper bounds rigorously, but in the convergence proof we use the maximization problem. We use the cone of Borel measures where we have one degree of freedom for every open set.

One can use the topology of V to equip the set I_t, consisting of the independent sets which have at most t elements, with a compact Hausdorff topology, see [29] for the technical details. Let $\mathcal{C}(I_{2t})$ be the set of continuous real-valued functions on I_{2t}. By the Riesz representation theorem (see e.g. [5, Chapter 2.2]) the topological dual of $\mathcal{C}(I_{2t})$, where the topology is defined by the supremum norm, can be identified with the space $\mathcal{M}(I_{2t})$ of signed Radon measures. A *signed Radon measure* is the difference of two Radon measures, where a *Radon measure* ν is a locally finite measure on the Borel algebra satisfying *inner regularity*: $\nu(B) = \sup\{\nu(C) : C \subseteq B,\ C \text{ compact}\}$ for each Borel set B. Nonnegative functions in $\mathcal{C}(I_{2t})$ form the cone $\mathcal{C}(I_{2t})_{\geq 0}$. Its conic dual is given by the cone of Radon measures

$$\begin{aligned}\mathcal{M}(I_{2t})_{\geq 0} &= (\mathcal{C}(I_{2t})_{\geq 0})^* \\ &= \{\lambda \in \mathcal{M}(I_{2t}) : \lambda(f) \geq 0 \text{ for all } f \in \mathcal{C}(I_{2t})_{\geq 0}\}.\end{aligned}$$

Denote by $\mathcal{C}(I_t \times I_t)_{\mathrm{sym}}$ the space of *symmetric kernels*, which are the continuous functions $K : I_t \times I_t \to \mathbb{R}$ such that

$$K(J, J') = K(J', J) \text{ for all } J, J' \in I_t.$$

We say that a symmetric kernel K is *positive* if

$$(K(J_i, J_j))_{i,j=1}^m \text{ is positive semidefinite for all } m \in \mathbb{N} \quad \text{and} \quad J_1, \dots, J_m \in I_t.$$

The positive kernels form a cone $\mathcal{C}(I_t \times I_t)_{\succeq 0}$. The dual of $\mathcal{C}(I_t \times I_t)_{\text{sym}}$ can be identified with the space of symmetric signed Radon measures $\mathcal{M}(I_t \times I_t)_{\text{sym}}$. Here a signed Radon measure $\mu \in \mathcal{M}(I_t \times I_t)$ is *symmetric* if

$$\mu(E \times E') = \mu(E' \times E) \text{ for all Borel sets } E \text{ and } E'.$$

We say that a measure $\mu \in \mathcal{M}(I_t \times I_t)_{\text{sym}}$ is *positive definite* if it lies in the dual cone $\mathcal{M}(I_t \times I_t)_{\succeq 0} = (\mathcal{C}(I_t \times I_t)_{\succeq 0})^*$.

Now we are ready to define our generalization:

– The optimization variable is $\lambda \in \mathcal{M}(I_{2t})_{\succeq 0}$.
– The objective function evaluates λ at $I_{=1}$, where in general,

$$I_{=t} = \{S \in I_t : |S| = t\},$$

and so when $t = 1$ we simply deal with all vertices, as singleton sets. This is similar to the objective function $\sum_{x \in V} y_{\{x\}}$ in Lasserre's hierarchy for finite graphs.
– The normalization condition reads $\lambda(\{\emptyset\}) = 1$.
– For generalizing the moment matrix condition "$M_t(y)$ is positive semidefinite" we use a dual approach. We define the operator $A_t \colon \mathcal{C}(I_t \times I_t)_{\text{sym}} \to \mathcal{C}(I_{2t})$ by

$$A_t K(S) = \sum_{J, J' \in I_t : J \cup J' = S} K(J, J').$$

which is bounded and hence continuous. Thus there exists the adjoint $A_t^* \colon \mathcal{M}(I_{2t}) \to \mathcal{M}(I_t \times I_t)_{\text{sym}}$ and the moment matrix condition reads $A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}$.

Definition 3.6. The t-th step of the generalized hierarchy is

$$\text{las}_t(G) = \sup \Big\{ \lambda(I_{=1}) : \lambda \in \mathcal{M}(I_{2t})_{\succeq 0},\ \lambda\big(\{\emptyset\}\big) = 1,\ A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0} \Big\}.$$

Clearly, we have a nonincreasing chain

$$\text{las}_1(G) \geq \text{las}_2(G) \geq \dots \geq \text{las}_{\alpha(G)-1}(G) \geq \text{las}_{\alpha(G)}(G) = \text{las}_{\alpha(G)+1}(G) = \dots, \quad (2)$$

which stabilizes after $\alpha(G)$ steps (note that $I_{\alpha(G)} = I_{\alpha(G)+1} = \dots$), and specializes to the original hierarchy if G is a finite graph. Each step gives an upper bound for $\alpha(G)$ because for every independent set S the measure

$$\lambda = \sum_{Q \in I_{2t} : Q \subseteq S} \delta_Q, \quad \text{where } \delta_Q \text{ is the delta measure at } Q,$$

is a feasible solution for $\text{las}_t(G)$ with objective value $|S|$. To see this we note that $\lambda(\{\emptyset\}) = 1$, and for any $K \in \mathcal{C}(I_t \times I_t)_{\succeq 0}$ we have

$$\langle K, A_t^* \lambda \rangle = \langle A_t K, \lambda \rangle = \sum_{R \in I_{2t} : R \subseteq S} \ \sum_{J, J' \in I_t : J \cup J' = R} K(J, J') = \sum_{J, J' \in I_t : J, J' \subseteq S} K(J, J') \geq 0.$$

The dual program of $\text{las}_t(G)$ is

$$\text{las}_t(G)^* = \inf \Big\{ K(\emptyset, \emptyset) : K \in \mathcal{C}(I_t \times I_t)_{\succeq 0},\ A_t K(S) \leq -1_{I_{=1}}(S) \quad \text{for} \quad S \in I_{2t} \setminus \{\emptyset\} \Big\},$$

and in [29] we showed that strong duality holds in every step:

Theorem 3.7. *Let G be a compact topological packing graph. For every $t \in \mathbb{N}$ we have $\text{las}_t(G) = \text{las}_t(G)^*$, and the optimum in $\text{las}_t(G)$ is attained.*

We show that the chain (2) converges to the independence number:

Theorem 3.8. *Let G be a compact topological packing graph. Then,*

$$\text{las}_{\alpha(G)}(G) = \alpha(G).$$

3.3 Explicit computations in the literature

Explicit computations of n-point bounds have been done in a variety of situations. Table 2 provides a guide to the literature.

For the first three packing problems in this table one can use Lasserre's hierarchy for finite graphs. For the last seven packing problems in this table our generalization can be used, where in the last four cases one has to perform a compactification of the vertex set first.

The convergence of the hierarchy shows that this approach is in theory capable of solving every given packing problem in discrete geometry. Computing higher steps in this hierarchy is computationally intractable but one attractive feature of the hierarchy is that already its first steps give strong upper bounds as one can see from the papers cited in the table above.

3.4 Generalization of Lasserre's hierarchy for energy minimization

De Laat [27] introduces an infinite-dimensional version of Lasserre's hierarchy for energy minimization problems. Consider a compact metric space (V, d) and a potential function $h \colon (0, \operatorname{diam} V] \to \mathbb{R}$ which should have the property that $h(s)$ goes to infinity when s goes to zero. Define a graph G with vertex set V where two distinct vertices are adjacent whenever $h(d(x, y)) \leq B$ where B is an upper bound on the minimum energy $\mathcal{E}_h(X)$ of N points $X = \{x_1, \dots, x_N\}$ in V. Define a continuous function $f \in \mathcal{C}(I_N)$ by

$$f(S) = \begin{cases} h(d(x, y)) & \text{if } S = \{x, y\} \text{ and } x \neq y, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 3.9. The t-th step of the generalized hierarchy for energy minimization is

$$E_t(G, N, h) = \inf \Big\{ \lambda(f) : \lambda \in \mathcal{M}(I_{2t})_{\succeq 0}, A_t^* \lambda \in \mathcal{M}(I_t \times I_t)_{\succeq 0}, \lambda(I_{=i}) = \binom{N}{i} \text{ for } 0 \leq i \leq 2t \Big\}.$$

Again, the hierarchy gives a sequence of increasingly stronger bounds which converges to the minimal energy, and again the infimum can be replaced by a minimum.

Theorem 3.10. *We have*

$$E_t(G, N, h) \leq E_{t+1}(G, N, h) \text{ for } t \in \mathbb{N},$$

and

$$E_N(G, N, h) = \min_{Y \subseteq V, |Y| = N} \mathcal{E}_h(Y).$$

For a proof of this theorem we refer to [27]. The bound $E_1(G, N, h)$ is a two-point bound and in the case of the unit sphere it is due to Yudin [53]. Yudin's bound is the principal tool in the study of universal optimal point configurations on the unit sphere by Cohn and Kumar [7].

4 Symmetry reduction

When a graph has infinitely many vertices, then computing any step in the semidefinite optimization hierarchies is an infinite-dimensional semidefinite program. In most cases, we do not know how to solve these optimization problems by analytic means. So one has to use a computer to determine an, at least approximate, optimal solution. Therefore a systematic approach to approximate the infinite-dimensional optimization problem by a sequence of finite-dimensional ones is needed.

One approach would be to discretize the graph and use the "classical" hierarchies. However, this is usually not a good idea, since by discretizing the graph one destroys the symmetry of the situation. It is a well-known fact that symmetries can be very beneficially exploited when solving convex optimization problems. Another approach, the one which we advocate here, is to first transform the semidefinite program at hand to its Fourier domain (e.g. we work with the space of Fourier coefficients) and then perform the discretization in the Fourier domain. Since in the Fourier domain the symmetries are particularly visible, the full symmetry of the situation can be exploited.

In this section we demonstrate this approach in a concrete setting, where the advantage of exploiting symmetry is especially apparent and nice. We will consider computing the first step $\mathrm{las}_1(\mathrm{Cay}(G,X))$ of the hierarchy for Cayley graphs defined on the compact Lie group $G = \mathrm{SO}(3)$, the group of three-dimensional rotation matrices.

Generally, a Cayley graph $\mathrm{Cay}(G,X)$ is defined by a group G and a subset $X \subseteq G$. The vertices of the Cayley graph $\mathrm{Cay}(G,X)$ are the elements of G and the neighborhood of the identity element e of G is the set X. Now this neighborhood is transported to all other group elements via the group action. This means that x and y in G are adjacent whenever $xy^{-1} \in X$. We want to work with undirected graphs so we additionally require that $X = X^{-1}$ holds.

We consider a Cayley graph whose vertices are the elements of the rotation group $\mathrm{SO}(3)$ which is given by all orthogonal matrices having determinant 1:

$$\mathrm{SO}(3) = \{A \in \mathbb{R}^{3\times 3} : A^{\mathsf{T}}A = I, \det(A) = 1\}.$$

Let $C \subseteq \mathbb{R}^3$ be a proper convex cone. Now we will study independent sets in the Cayley graph $\Gamma = \mathrm{Cay}(G,X)$ defined by

$$G = \mathrm{SO}(3) \quad \text{and} \quad X = \{A \in \mathrm{SO}(3) : C^\circ \cap AC^\circ \neq \emptyset\},$$

where C° denotes the topological interior of C. Independent sets in this Cayley graph exactly correspond to packings of rotations of the set $C \cap S^2$ on the unit sphere S^2. This means for instance that if the cone C has a regular k-gon as a base then we want to pack regular spherical k-gons on the unit sphere; when the cone C has a round disk as a base then we are packing spherical caps on S^2, or in other words, we are considering spherical codes.

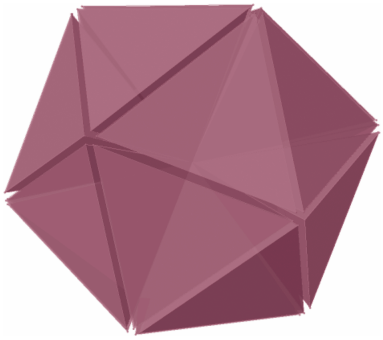

Figure 1. *Packing of twenty cones having an equilateral triangle as base*

As mentioned in the previous section, the first step of the hierarchy coincides with the ϑ'-number, a strengthened version of the Lovász ϑ-number. In our case, it has, after dualizing, the following form:

$$\begin{aligned} \vartheta'(\Gamma) = \inf \quad & M \\ & K - J \qquad \text{is a positive kernel,} \\ & K(x,x) = M \quad \text{for all } x \in G, \\ & K(x,y) \leq 0 \quad \text{for all } \{x,y\} \notin E \text{ where } x \neq y, \\ & M \in \mathbb{R},\ K \in \mathcal{C}(G \times G) \text{ is symmetric.} \end{aligned} \tag{3}$$

Here an element in the space $\mathcal{C}(G \times G)$ of real-valued continuous functions over $G \times G$ is called a *kernel*. A kernel K is *symmetric* if $K(x,y) = K(y,x)$ for all $x, y \in G$. It is *positive* if it is symmetric and if for any $m \in \mathbb{N}$ and for any $x_1, \ldots, x_m \in G$, the matrix $\big(K(x_i,x_j)\big)_{i,j=1}^m$ is positive semidefinite. The kernel $K - J$ is simply defined componentwise by $(K-J)(x,y) = K(x,y) - 1$ for $x, y \in G$.

Now we are ready to discuss how to exploit the symmetry to simplify the computation of $\vartheta'(\Gamma)$. The optimization problem (3) defininig $\vartheta'(\Gamma)$ is invariant under the action of G, i.e. if (M,K) is a feasible solution of (3) then so is (M,K') where

$$K'(x,y) = K(g^{-1}x, g^{-1}y) \quad \text{with} \quad g \in G.$$

Hence, we can symmetrize every feasible solution (M,K) by taking the group average and applying the Haar integral of the group G:

$$\overline{K}(x,y) = \int_G K(g^{-1}x, g^{-1}y)\, dg.$$

Then one can check that $(M,\overline{K})$ is again a feasible solution of (3) and that $\overline{K}$ is G-invariant, i.e.

$$\overline{K}(g^{-1}x, g^{-1}y) = K(x,y) \quad \text{for all} \quad x, y, g \in G.$$

Thus, instead of optimizing over all $K \in \mathcal{C}(G \times G)$ it suffices to optimize only over G-invariant kernels. In particular, these kernels only depend on one parameter in G and no longer on two parameters in $G \times G$ because $\overline{K}(x,y) = \overline{K}(e, x^{-1}y)$ holds for all $x, y \in G$. In harmonic analysis the cone of positive kernels which are invariant

Table 2. *Computation of n-point bounds*

Packing problem	2-point bound	3-point bound	4-point bound
Binary codes	Delsarte [11]	Schrijver [45]	Gijswijt, Mittelmann, Schrijver [20]
q-ary codes	Delsarte [11]	Gijswijt, Schrijver, Tanaka [19]	Litjens, Polak, Schrijver [36]
Constant weight codes	Delsarte [11]	Schrijver [45], Regts [41]	
Spherical codes	Delsarte, Goethals, Seidel [13]	Bachoc, Vallentin [4]	
Codes in $\mathbb{R}\mathrm{P}^{n-1}$	Kabatiansky, Levenshtein [25]	Cohn, Woo [10]	
Grassmannian codes	Bachoc [1]		
Sphere packings	Cohn, Elkies [6]		
Binary sphere and spherical cap packings	de Laat, Oliveira, Vallentin [28]		
Translative body packings	Dostert, Guzmán, Oliveira, Vallentin [14]		
Congruent copies of a convex body	Oliveira, Vallentin [39]		

under the group G is called the cone of *positive type functions* on G and it is denoted by $\mathcal{P}(G)$.

So we can simplify (3) by setting $\phi(x) = K(e, x)$:

$$\vartheta'(\Gamma) = \inf\{\phi(e) : \phi - 1 \in \mathcal{P}(G), \phi(x) \leq 0 \text{ for all } x \notin X\}. \quad (4)$$

The next step is to give an explicit parametrization of the cone $\mathcal{P}(G)$ and here harmonic analysis kicks in. For an excellent reference on harmonic analysis where the cone of positive type functions is discussed in detail we recommend the reader to consult the book [17] by Folland.

When the group G is compact, then the parametrization of $\mathcal{P}(G)$ is particularly nice (but in the following, for the sake of simplicity, we ignore all convergence issues, although they are not difficult). To write down the parametrization we have to have an explicit description of the irreducible unitary representations of the group G.

A few standard definitions from harmonic analysis: A (finite-dimensional) *unitary representation* of G is a group homomorphism $\pi\colon G \to \mathrm{U}(d_\pi)$ where $\mathrm{U}(d_\pi)$ is the group of unitary $d_\pi \times d_\pi$ matrices. A subspace M of $\mathbb{C}^{d_\pi}$ is *π-invariant* if $\pi(g)m \in M$ for all $g \in G$ and $m \in M$. The unitary representation π is said to be *irreducible* if $\{0\}$ and $\mathbb{C}^{d_\pi}$ are the only π-invariant subspaces of $\mathbb{C}^{d_\pi}$. Two unitary representations π and π' are (unitarily) *equivalent* if there is a unitary matrix T such that $T\pi(g) = \pi'(g)T$ for all $g \in G$. We can fix a set of mutually inequivalent irreducible unitary representations of G, so that each unitary equivalence class has a representative; call this set $\widehat{G}$. This allows us to define the *Fourier transform* of a function $f\colon G \to \mathbb{C}$:

$$\hat{f}(\pi) = \int_G f(g)\pi(g)\, dg,$$

where $\hat{f}(\pi)$ is a complex $d_\pi \times d_\pi$ matrix. The *Fourier inversion formula* says we can recover f from its Fourier transform:

$$f(g) = \sum_{\pi \in \widehat{G}} d_\pi \langle \hat{f}(\pi), \pi(g)\rangle.$$

The inner product used here is the trace inner product, defined as $\langle A, B\rangle = \mathrm{Trace}(B^*A)$ for square complex matrices A and B of the same dimension, where B^* denotes the conjugate-transpose of B.

Now the explicit parametrization of the cone of functions of positive type reads as follows (it is frequently called Bochner's theorem):

Theorem 4.1. *Let G be a compact group. The cone of positive type functions on G equals*

$$\mathcal{P}(G) = \Big\{ \sum_{\pi \in \widehat{G}} d_\pi \langle \hat{f}(\pi), \pi(g)\rangle : \hat{f}(\pi) \text{ is (Hermitian) positive semidefinite} \Big\}.$$

So the cone $\mathcal{P}(G)$ is an infinite-dimensional direct product of cones of finite-dimensional Hermitian positive semidefinite matrices. Using just finitely many cones of this direct product is a natural way to approximate the original infinite-dimensional optimization problem by finite-dimensional ones.

In the case of the rotation group $G = \mathrm{SO}(3)$ we can give explicit formulæ: We parametrize the elements of $\mathrm{SO}(3)$ by Euler angles. A triple of angles $(\varphi_1, \theta, \varphi_2)$, where $\varphi_1, \varphi_2 \in [0, 2\pi]$ and $\beta \in [0, \pi]$ corresponds to the rotation where we

(1) first rotate on the xy-plane, keeping z fixed, by an angle of φ_1;
(2) then rotate on the yz-plane, keeping x fixed, by an angle of θ;
(3) and finally rotate on the xy-plane again, keeping z fixed, by an angle of φ_2.

Thus we describe all elements of $\mathrm{SO}(3)$ as a product of the three corresponding matrices $g_{\varphi_1} g_\theta g_{\varphi_2}$ with

$$g_{\varphi_1} = \begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 & 0 \\ \sin\varphi_1 & \cos\varphi_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, g_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$
$$g_{\varphi_2} = \begin{pmatrix} \cos\varphi_2 & -\sin\varphi_2 & 0 \\ \sin\varphi_2 & \cos\varphi_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then a positive type function $f\colon \mathrm{SO}(3) \to \mathbb{R}$ is given by

$$f(\varphi_1, \theta, \varphi_2) = \sum_{l=0}^{\infty} (2l+1) \langle \hat{f}(l), \pi_l(\varphi_1, \theta, \varphi_2)\rangle,$$

where $\hat{f}(l)$ is a Hermitian positive semidefinite matrix of size $(2l+1) \times (2l+1)$ and the unitary representation π_l is given by

$$\pi_l(\varphi_1, \theta, \varphi_2) = \left(i^{m-n} e^{-i(m\varphi_1 + n\varphi_2)} P^l_{mn}(\cos\theta) \right)^l_{m,n=-l},$$

and where

$$P^l_{mn}(\cos\theta) = \left(\frac{(l-m)!(l+m)!}{(l-n)!(l+n)!} \right)^{1/2} \sin^{m-n}\frac{\theta}{2} \cos^{m+n}\frac{\theta}{2} P_l^{(m-n,m+n)}(\cos\theta),$$

and where $P_k^{\alpha,\beta}$ is a Jacobi polynomial of degree k which is orthogonal with respect to the measure $(1-x)^\alpha(1+x)^\beta\, dx$ on the interval $[-1, 1]$.

The nonnegative integers $l = 0, 1, \ldots$ determine the set $\widehat{G}$. In total, we have a very explicit parametrization of the cone. The book [18] for example contains an elementary way of deriving the formulæ above.

With this information one can compute approximations of $\vartheta'(\mathrm{Cay}(G, X))$ using standard SDP solvers: Instead of using all possible $f \in \mathcal{P}(G)$ one only optimizes over those f which can be written as finite sums

$$f(\varphi_1, \theta, \varphi_2) = \sum_{l=0}^{N} (2l+1) \langle \hat{f}(l), \pi_l(\varphi_1, \theta, \varphi_2)\rangle,$$

given some prescribed value of N. Usually, but this of course depends on X, already small values of N, like $N = 15$, give good approximations for $\vartheta'(\mathrm{Cay}(G, X))$.

Until now we only used the symmetry of the group G to simplify the problem but one can also use the symmetry of the set X, in our case the symmetry of the cone C which defines X, to further simplify the optimization problem.

Let $K \subseteq G$ be the symmetry group of cone C. We can assume that ϕ in (4) is K bi-invariant, i.e. that

$$\phi(k_1 x k_2) = \phi(x) \quad \text{for all } k_1, k_2 \in K, x \in G,$$

holds. As far as we know, a proof of this fact did not appear in the literature. So let's give it here, albeit in small print.

Indeed, let ϕ be a feasible solution of (4). Define

$$\psi(x) = \int_K \int_K \phi(k_1 x k_2)\, dk_1\, dk_2.$$

Then ψ is K bi-invariant and it is again a feasible solution of (4) with $\psi(e) \leq \phi(e)$. We have $\psi(x) \leq 0$ if $x \notin X$ because for $k_1, k_2 \in K$

$$k_1 x k_2 \notin X \iff C^\circ \cap k_1 x k_2 C^\circ = \emptyset \iff (k_1)^{-1} C^\circ \cap x k_2 C^\circ = \emptyset \iff C^\circ \cap x C^\circ = \emptyset$$

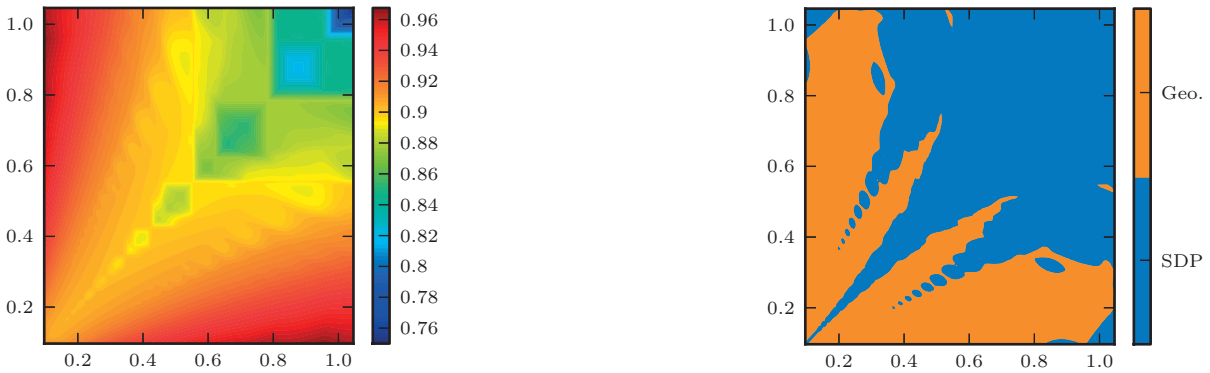

Figure 2. *Left: Two-point (SDP) bound for the packing of spherical caps of two different, given sizes on S^2. Right: Comparison of two-point bound with geometric bound.*

holds. We have $\psi - 1 \in \mathcal{P}(G)$ because we can write $\phi - 1 = f^* * f$ as a convolution (see [12, Proposition 9.4]) and because for $x_1, \dots, x_N \in G$ and $v \in \mathbb{R}^N$ we have (see [12, Proof of Proposition 12.4])

$$\begin{aligned}
& v^{\mathsf{T}} \left(\psi(x_i^{-1}x_j) - 1\right)_{i,j=1,\dots,N} v \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} v_i v_j \int_K \int_K (\phi(k_1 x_i^{-1} x_j k_2) - 1)\, dk_1\, dk_2 \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} v_i v_j \int_K \int_K \int_G \overline{f(z)} f(z k_1 x_i^{-1} x_j k_2)\, dz\, dk_1\, dk_2 \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} v_i v_j \int_K \int_K \int_G \overline{f(z x_i k_1^{-1})} f(z x_j k_2)\, dz\, dk_1\, dk_2 \\
&= \int_G \left| \sum_{i=1}^{N} v_i \int_K f(z x_i k)\, dk \right|^2 dz \\
&\geq 0.
\end{aligned}$$

When C is a convex cone having a regular k-gon as base, then K is the dihedral group, and we can work with K bi-invariant functions of the form

$$f(\varphi_1, \theta, \varphi_2) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{n=-l}^{l} \hat{f}^l_{mn} \pi^l_{mn}(\varphi_1, \theta, \varphi_2),$$

where we sum only over m and n which are multiples of k and where $\varphi_1, \varphi_2 \in [0, 2\pi/n]$, thus reducing the domain of f and reducing the number of Fourier coefficients we have to optimize over.

The more symmetry, the better: When C is a convex cone which has a round disk as base, then $K = \mathrm{SO}(2)$, and we can work with functions of the form

$$f(\varphi_1, \theta, \varphi_2) = \sum_{l=0}^{\infty} (2l+1) f^l_{00} P^l_{00}(\cos\theta), \text{ with } f^l_{00} \geq 0.$$

In fact, these functions no longer depend on the angles φ_1 and φ_2. In particular, this representation shows that (4) collapses from a semidefinite program to a linear program, as was realized in [3]. It is the classical linear programming bound of Delsarte, Goethals and Seidel [13], the first two-point bounds for non finite spaces which appeared in the literature.

5 Some recent results and open problems

We already emphasized that the successful computation of n-point bounds usually gives the strongest known bounds for geometric packing and energy minimization problems, see Table 2. In this final section we want to discuss some recent results and open problems.

5.1 Shape of two-point bounds for spherical codes

Even in the classical case of two-point point bounds for spherical codes in S^2 we do not understand the bound completely.

In the paper [28] de Laat, Oliveira and Vallentin computed two-point bounds for the packing of spherical caps of N different sizes on S^2, see Figure 2 for $N = 2$.

Florian [15, 16] provided 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 given angles $\alpha_1, \dots, \alpha_N \in (0, \pi/3]$ is at most

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

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 α_i, α_j, and α_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 2 we see that for $N = 2$ it depends on the angles whether the geometric or the two-point bound is sharper. In particular we see that near the diagonal the two-point bound is at least as good as the geometric bound.

We also used our programs to plot the upper bounds for $N = 1$, the classical linear programming bound of Delsarte, Goethals, and Seidel [13], see Figure 3. The plot seems to reveal interesting and new properties of the bound. For better orientation we show in the plot the packings where the two-point bound is sharp (cf. Levenshtein [34]; Cohn and Kumar [7] proved the much stronger statement that these packings provide point configurations which are universally optimal). The dotted line is the geometric bound, and since we know that both the geometric (cf. Florian [15]) and the two-point bounds are sharp for the given configurations, we know that at these peaks the bounds meet.

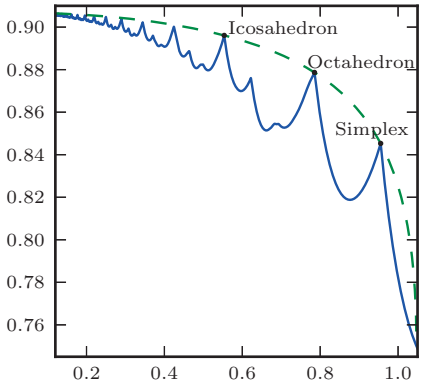

Figure 3. *Computation of two-point bounds for spherical codes in S^2 taken from [28]: The colored graph corresponds to the two-point bound, the dashed graph corresponds to the geometric bound of Florian.*

An interesting feature of the two-point bound seems to be that it has a periodic behavior: The numerical results suggest that the two-point bound and the Florian bound meet infinitely often as the angle decreases, and that between any two of these meeting points the two-point bound has a similar shape. Howerver, we do not know how to prove this.

5.2 *Translative packings of non-spherical shapes*

In the paper [14] Dostert, Guzmán, Oliveira, and Vallentin computed upper bounds for translative packing of three-dimensional superballs; unit balls of the l_3^p-norm, with $p \geq 1$:

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

Jiao, Stillinger, and Torquato [24] constructed by computer search the densest known packings of B_3^p for many values of p. Although they principally allow congruent packings in their computer simulations, the dense packings they find all have the structure of a lattice. They subdivide the range $p \in [1, \infty)$ into four different regimes

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

and give for every regime a family of lattices determining dense packings ($O1$, $O0$, $C0$, $C1$ as defined in [24]). There is a direct relation to materials science as three-dimensional superballs has been synthesized experimentally as colloids, see Rossi et al. [42].

As motivation for their study Jiao, Stillinger, and Torquato write:

> Understanding the organizing principles that lead to the densest packings of nonspherical particles that do not tile space is of great practical and fundamental interest. Clearly, the effect of asphericity is an important feature to include on the way to characterizing more fully real dense granular media.
>
> [...]
>
> On the theoretical side, no results exist that rigorously prove the densest packings of other congruent non-space-tiling particles in three dimension.

For computing our upper bounds we used the following theorem of Cohn and Elkies [6], which falls into the framework of two-point bounds.

Theorem 5.1. *Let $\mathcal{K}$ be a convex body in $\mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous L^1-function. Let*

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

denote the Fourier transform of f at u. Suppose f satisfies the following conditions

(i) $\widehat{f}(0) \geq 1$,

(ii) *f is of positive type, i.e. $\widehat{f}(u) \geq 0$ for every $u \in \mathbb{R}^n$,*

(iii) *$f(x) \leq 0$ whenever $\mathcal{K}^\circ \cap (x + \mathcal{K}^\circ) = \emptyset$.*

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

The Cohn–Elkies bound provides the basic framework for proving the best known upper bounds for the maximum density of sphere packing. For a long time it was conjectured to provide tight bounds in dimensions 8 and 24 and there was very strong numerical evidence to support this conjecture, see Cohn and Miller [9]. However, the only thing missing was a rigorous proof. Very recently, in March 2016, such a proof was found by Viazovska [52] for dimension 8 and a few days later, building on Viazovska's breakthrough result, by Cohn, Kumar, Miller, Radchenko, and Viazovska [8] for dimension 24. Here the construction of optimal functions f uses the theory of quasimodular forms from analytic number theory.

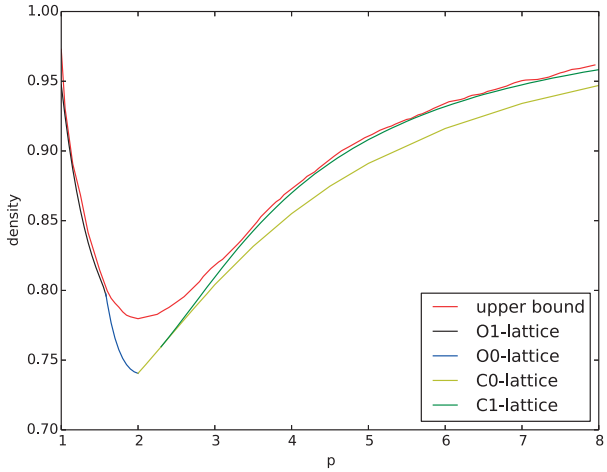

Figure 4. *Lower and upper bounds for superball packings in dimension* 3

To find good functions f for superball packings we used tools from polynomial optimization. In particular we used invariant theory of reflection groups G to give an explicit parametrization of G-invariant Hermitian symmetric polynomials which are complex versions of sums of squares. The upper bounds we obtained are given in Figure 4. In [14] we also computed new upper bounds for the optimal density of translative packing densities of several polytopes, like the regular tetrahedron.

For values $p \geq 3$ the upper bounds are remarkably close to the lower bounds. Maybe the two-point bounds are strong enough to prove optimality of some of the lattice packings of Jiao, Stillinger and Torquato? Being in a very optimistic mood one could imagine that the question: "For every dimension n there is a constant $P(n) < \infty$ so that the bound is tight for all values of $p \geq P(n)$?" has a positive answer. Even more optimistically: "Is it true that $P(8) \leq 2$ and that $P(24) \leq 2$?"

5.3 *Four-point bounds for energy minimization*

The last puzzling open problem we would like to mention concerns four-point bounds for the five-electron case of Thomson's problem. We want to distribute five points $x_1, \ldots, x_5$ on the unit sphere S^2 so that the Riesz s-potential energy

$$\sum_{1 \leq i < j \leq 5} \frac{1}{\|x_i - x_j\|^s}$$

is minimized. It is well known, but only experimentally, that there are two competing configurations: For the values $0 \leq s \leq 15.04\ldots$ the vertices of the triangular bipyramid (two antipodal points and three points forming an equilateral triangle on the equator) seem to be the unique minimizer, and for the other values $s \geq 15.04\ldots$ the vertices of a square pyramid (where the lattitude of the base depends on the specific values of s) are believed to be the unique minimizer. For the cases $s = 1$ and $s = 2$ Schwartz [48] proved this by essentially enumerating all possibilities. Recently, Schwartz [49] extended his result to the entire interval $0 \leq s \leq 6$ using an observation of Tumanov [51].

De Laat [27] computed (numerically) the four-point bound E_2 which we defined in Section 3 for the values $s = 1, 2, 4$. In all three cases the numerical results he obtained indicate that the four-point bound is sharp. It even might be the case that the bound E_2 is universally sharp for all values of s. This would be a fascinating example of the validity of a very general conjecture posed by Cohn and Woo [10]:

Conjecture 5.2. If there exists a completely monotonic potential function that is k-point sharp for N points in S^{n-1} and is not a

polynomial, then every completely monotonic potential function is k-point sharp for N points in S^{n-1}.

Acknowledgements

The author thanks Maria Dostert, Cristóbal Guzmán, David de Laat, and Jan Hendrik Rolfes for their comments on previous versions of this article.

Frank Vallentin, Mathematisches Institut, Universität zu Köln, Weyertal 86–90, 50931 Köln, Germany. frank.vallentin@uni-koeln.de

Bibliography

[1] C. Bachoc, *Linear programming bounds for codes in Grassmannian spaces*, IEEE Trans. Inf. Th. **52** (2006), 2111–2125. (http://arxiv.org/abs/math/0610812)

[2] C. Bachoc, D.C. Gijswijt, A. Schrijver, F. Vallentin, *Invariant semidefinite programs*, pages 219–269 in *Handbook on Semidefinite, Conic and Polynomial Optimization* (M.F. Anjos, J.B. Lasserre (ed.)), Springer 2012. (http://arxiv.org/abs/1007.2905)

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

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

[5] C. Berg, J.P.R. Christensen, P. Ressel, *Harmonic analysis on semigroups: theory of positive definite and related functions*, Springer-Verlag, 1984.

[6] H. Cohn, N.D. Elkies, *New upper bounds on sphere packings I*, Ann. of Math. (2) **157** (2003), 689–714. (http://arxiv.org/abs/math/0110009)

[7] H. Cohn, A. Kumar, *Universally optimal distribution of points on spheres*, J. Amer. Math. Soc. **20** (2007), 99–148. (http://arxiv.org/abs/math/0607446)

[8] H. Cohn, A. Kumar, S.D. Miller, D. Radchenko, M.S. Viazovska, *The sphere packing problem in dimension* 24, arXiv:1603.06518 [math.NT], 12 pages. (http://arxiv.org/abs/1603.06518)

[9] H. Cohn, S.D. Miller, *Some properties of optimal functions for sphere packing in dimensions* 8 *and* 24, arXiv:1603.04759 [math.MG], 23 pages. (http://arxiv.org/abs/1603.04759)

[10] H. Cohn, J. Woo, *Three-point bounds for energy minimization*, J. Amer. Math. Soc. **25** (2012), 929–958. (http://arxiv.org/abs/1103.0485)

[11] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. (1973), vi+97.

[12] P.E.B. DeCorte, *The Eigenvalue Method for Extremal Problems on Infinite Vertex-Transitive Graphs*, PhD thesis, TU Delft, 2015. http://repository.tudelft.nl/view/ir/uuid%3Ada16eb76-4031-4350-baf3-307aa01a337c/

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

[14] M. Dostert, C. Guzmán, F.M. de Oliveira Filho, F. Vallentin, *New upper bounds for the density of translative packings of three-dimensional convex bodies with tetrahedral symmetry*, arXiv:1510.02331 [math.MG], 29 pages. (http://arxiv.org/abs/1510.02331)

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

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

[17] G.B. Folland, *A course in abstract harmonic analysis*, Studies in Advanched Mathematics, 1995.

[18] I.M. Gel'fand, R.A. Minlos, Z.Ya. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications*, The MacMillan Company, New York, 1963.

[19] D.C. Gijswijt, A. Schrijver, H. Tanaka, *New upper bounds for nonbinary codes based on the Terwilliger algebra and semidefinite programming*, J. Comb. Theory, Ser. A **113** (2006), 1719–1731.

[20] D.C. Gijswijt, H.D. Mittelmann, A. Schrijver, *Semidefinite code bounds based on quadruple distances.*, IEEE Trans. Inf. Th. **58** (2012), 2697–2705. (http://arxiv.org/abs/1005.4959)

[21] T.C. Hales, *A proof of the Kepler conjecture*, Ann. of Math. (2) **162** (2005), 1065–1185.

[22] D.P. Hardin, E.B. Saff, *Discretizing manifolds via minimum energy points*, Notices Amer. Math. Soc. **51** (2004), 1186–1194.

[23] A.J. Hoffman, *On eigenvalues and colorings of graphs*, pp. 79–91 in (B. Harris, ed.) Graph Theory and its Applications, Academic Press, 1970.

[24] Y. Jiao, F.H. Stillinger, S. Torquato, *Optimal packings of superballs*, Phys. Rev. E **79** (2009), 041309, 12 pp. (http://arxiv.org/abs/0902.1504)

[25] G.A. Kabatiansky, V.I. Levenshtein, *On Bounds for Packings on a Sphere and in Space*, Probl. Peredachi Inf. **14** (1978), 3–25.

[26] R.M. Karp, Reducibility among combinatorial problems, in: *Complexity of Computer Computations* (Proceedings of a symposium on the Complexity of Computer Computations, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, 1972; R.E. Miller, J.W. Thatcher, eds.), Plenum Press, New York, 1972, pp. 85–103.

[27] D. de Laat, *Moment methods in extremal geometry*, PhD thesis, TU Delft, 2016. http://repository.tudelft.nl/view/ir/uuid%3Afce81f72-8261-484d-b9f4-d3d0c26f0473/

[28] D. de Laat, F.M. de Oliveira Filho, F. Vallentin, *Upper bounds for packings of spheres of several radii*, Forum Math. Sigma **2** (2014), e23 (42 pages). (http://arxiv.org/abs/1206.2608)

[29] D. de Laat, F. Vallentin, *A semidefinite programming hierarchy for packing problems in discrete geometry*, Math. Program., Ser. B **151** (2015), 529–553. (http://arxiv.org/abs/1311.3789)

[30] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. **11** (2001), 796–817.

[31] J.B. Lasserre, *An explicit equivalent positive semidefinite program for nonlinear 0-1 programs*, SIAM J. Optim. **12** (2002), 756–769.

[32] M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming*, Math. Oper. Res. **28** (2003), 470–496.

[33] M. Laurent, *Strengthened semidefinite programming bounds for codes*, Math. Program., Ser. B **109** (2007), 239–261.

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

[35] B. Lindström, *Determinants on semilattices*, Proc. Amer. Math. Soc. **20** (1969) 207–208.

[36] B. Litjens, S. Polak, A. Schrijver, *Semidefinite bounds for nonbinary codes based on quadruples*, arXiv:1602.02531 [math.CO], 13 pages. (http://arxiv.org/abs/1602.02531)

[37] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inf. Th. **25** (1979), 1–7.

[38] L. Lovász, A. Schrijver, *Cones of matrices and set-functions and 0-1 optimization*, SIAM J. Optim. **1** (1991), 166–190.

[39] F.M. de Oliveira Filho, F. Vallentin, *Computing upper bounds for packing densities of congruent copies of a convex body*, arXiv:1308.4893 [math.MG], 29 pages. (http://arxiv.org/abs/1308.4893)

[40] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Ind. Univ. Math. J. **42** (1993), 969–984.

[41] G. Regts, *Upper bounds for ternary constant weight codes from semidefinite programming and representation theory*, Master thesis, University of Amsterdam, 2009.

[42] L. Rossi, S. Sacanna, W.T.M. Irvine, P.M. Chaikin, D.J. Pine, A.P. Philipse, *Cubic crystals from cubic colloids*, Soft Matter **7** (2011) 4139–4142.

[43] A. Schrijver, *A comparison of the Delsarte and Lovász bounds*, IEEE Trans. Inf. Th. **25** (1979), 425–429.

[44] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Springer-Verlag, Berlin, 2003.

[45] A. Schrijver, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Transactions on Information Theory **51** (2005) 2859–2866.

[46] N.J.A. Sloane, *Spherical Coverings*, http://neilsloane.com/coverings/index.html

[47] S. Smale, *Mathematical problems for the next century*, Mathematical Intelligencer **20** (1998), 7–15.

[48] R.E. Schwartz, *The five-electron case of Thomson's problem*, Exp. Math. **22** (2013), 157–186. (http://arxiv.org/abs/1001.3702)

[49] R.E. Schwartz, *The triangular bi-pyramid minimizes a range of power law potentials*, arXiv:1512.04628 [math.MG], 57 pages. (http://arxiv.org/abs/1512.04628)

[50] S. Torquato, Y. Jiao, *Dense packings of the Platonic and Archimedean solids*, Nature **460** (2009), 876–879. (http://arxiv.org/abs/0908.4107)

[51] A. Tumanov, *Minimal biquadratic energy of 5 particles on 2-sphere*, Indiana Univ. Math Journal **62** (2013), 1717–1731. (http://arxiv.org/abs/1208.5044)

[52] M.S. Viazovska, *The sphere packing problem in dimension* 8, arXiv:1603.04246 [math.NT], 22 pages. (http://arxiv.org/abs/1603.04246)

[53] V.A. Yudin, *Minimum potential energy of a point system of charges*, Discrete Math. Appl. **3** (1993), 75–81.

[54] G.M. Whitesides, J.K. Kriebel, J.K., B.T. Mayers, *Self-Assembly and Nanostructured Materials*, in: *Nanoscale Assembly: Chemical Techniques*, (W.T.S. Huck, ed.) Springer, 2005, pp. 217–239.

[55] H.S. Wilf, *Hadamard determinants, Möbius functions, and the chromatic number of a graph*, Bull. Amer. Math. Soc. **74** (1968) 960–964.

[56] L. Zheng, D.N.C. Tse, *Communication on the Grassmann Manifold: A Geometric Approach to the Noncoherent Multiple-Antenna Channel*, IEEE Trans. Inf. Th. **48** (2002), 359–383.

Etienne de Klerk

Computer-assisted proofs and semidefinite programming

When asked to write a discussion column on the paper of Frank Vallentin that appears in this issue of OPTIMA, I recalled a recent interview with Roger Fletcher that also appeared in this newsletter [*Optima* 99 (12/2015)].

In this interview Roger Fletcher mentions some early work of his on semidefinite programming (SDP), in particular on SDP problems related to the educational testing problem. This was in the days before interior point methods, necessitating some novel algorithmic ideas on his part. Later in the interview, he continues (on the topic of applications):

> I think for us in numerical analysis – in optimization in particular – [...] the purpose [is] to optimize; it's not to produce pure mathematics in my view, differing from other people's view. But it's to solve problems.

The paper of Frank Vallentin sheds an interesting light on these remarks, since it reviews how computational optimization, and in particular SDP, can be useful in producing results in pure mathematics (some packing and covering problems in discrete geometry, to be precise).

At first glance this work both illustrates and contradicts the comments by Fletcher! In my view, a good way to reconcile these things is to recognize pure mathematics as an area of application of computational optimization.

Optimization for packing and covering problems

The most famous example of the use of computational optimization for proving a result in geometry must surely be the work by Thomas Hales [4] in using LP to prove the Kepler conjecture.

Vallentin and his co-authors have continued in this spirit by deriving new results for several problems, like bounds on kissing numbers in several dimensions; see Figure 1.

The basic plan of attack follows three steps:

1. Reformulate the packing/covering problem in question as a maximum stable set problem in an infinite graph; e.g., for the kissing number problem, the vertex set is the unit sphere, and two points on the unit sphere are deemed adjacent if one cannot place two unit balls to touch the sphere at these points without overlapping.
2. Generalize the Lasserre SDP hierarchy for the maximum stable set problem to infinite graphs to obtain a sequence of semi-infinite SDP problems;
3. Exploit symmetry via harmonic analysis to reduce the semi-infinite SDP problems to finite ones, and solve the resulting problems numerically to obtain bounds.

This simplification belies the mathematical sophistication of the approach, but already hints at the amount of theory involved. This in turn begs the question on whether there is not a simpler way to compute bounds of a similar quality. Indeed, for many (if not most) combinatorial optimization problems there are good computational alternatives to compute bounds for given instances, without invoking SDP.

For the packing and covering problems surveyed by Vallentin, however, it does seem that – at the moment at least – SDP often provides the best bounds. To give one example that is not explicitly mentioned in Vallentin's survey, Bachoc and Vallentin [1] have used the above methodology to show an upper bound of 45 on the kissing number in 5 dimensions (a lower bound of 40 is known by an explicit construction). This improved the previously best known upper bound, namely 46. Although this may seem like a small improvement,

Figure 1. *The kissing number in a given dimension is the largest number of Euclidean unit balls that can simultaneously touch a central ball. In* $\mathbb{R}^2$ *it is* 6 *(left picture) and in* $\mathbb{R}^3$ *it is* 12 *(right picture). (Pictures courtesy of Anja Vallentin)*

kissing numbers have been studied extensively, and any progress is deemed very significant. On the other hand, it would be nice if the approach surveyed by Vallentin could lead to the solution of a major open problem, like settling the kissing number in some dimension.

Cross-fertilization

If SDP solvers have proved useful for some discrete geometry problems, the reverse has also been true to some extent. The typical SDP instances that are solved are often very ill-conditioned numerically, while accurate dual solutions are needed to give formal proofs. This can lead to some tedious 'reverse engineering', where an approximately feasible dual solution is 'massaged' until it becomes feasible; see, e.g., the discussion in [3, Section 5.3].

Thus high precision solvers like SDPA-GMP [5] have been called on, and there has even been new work on interior point methods for SDP using exact arithmetic [2]. This kind of cross-fertilization is undoubtedly what Roger Fletcher was hinting at in his interview: solving concrete instances provides necessary feedback for further solver development.

Etienne de Klerk, TISEM, Econometrics and Operations Research, Tilburg University, 5000LE Tilburg, The Netherlands. e.deklerk@uvt.nl

Bibliography

[1] C. Bachoc and F. Vallentin, New upper bounds for kissing numbers from semidefinite programming. *Journal of the AMS*, **21**:909–924, 2008.

[2] E. de Klerk and F. Vallentin, On the Turing model complexity of interior point methods for semidefinite programming, arXiv:1507.03549, 2015.

[3] D. de Laat, F.M. de Oliveira Filho, and F. Vallentin, Upper bounds for packings of spheres of several radii, *Forum Math. Sigma* **2**, e23 (42 pages) 2014.

[4] T.C. Hales, A proof of the Kepler conjecture, *Ann. of Math.* **2**(162): 1065–1185, 2005.

[5] M. Yamashita, K. Fujisawa, M. Fukuda, K. Kobayashi, K. Nakata, and M. Nakata, *Latest developments in the SDPA family for solving large-scale SDPs*, pp. 687–713 in Handbook on Semidefinite, Conic and Polynomial Optimization (M.F. Anjos, J.B. Lasserre (ed.)), Springer, 2012.

Henry Cohn

When are semidefinite programming bounds sharp?

The aspect of Vallentin's paper I find most intriguing is the existence of sharp bounds in exceptional cases. We should not generally expect sharp bounds, and there seems to be no hope of solving most geometric optimization problems rigorously. For example, consider the Thomson problem for 113 electrically charged particles confined to the unit sphere: how can we choose unit vectors $x_1, \dots, x_{113}$ in $\mathbb{R}^3$ so as to minimize the Coulomb energy

$$\sum_{1 \le i < j \le 113} \frac{1}{|x_i - x_j|}$$

between them? In principle the search space of 113-particle configurations is large and complicated, but in practice the minimum is not

so difficult to approximate numerically, for example by the conjugate gradient algorithm. The minimal energy seems to be 5721.8249…, and it can be computed to many decimal places, but there is no apparent way to prove that we are not merely stuck in a local optimum. It is unlikely that the number 5721.8249… or the particle configuration has any simple description, and it is hard to believe that humanity will ever find a rigorous proof of optimality. Instead, the optimum is what it is, and we will probably never achieve a deep conceptual understanding of why.

Of course one explanation for the difficulty of the 113-particle Thomson problem is its complexity. If the answer is itself complicated, why should we expect to find a proof? By contrast, two-point bounds prove that the vertices of the regular icosahedron minimize Coulomb energy for twelve particles [1], and the sharpness of this bound is not so unreasonable. After all, the icosahedron is a much simpler and more tractable object.

On the other hand, there is more to this distinction than just counting particles or parameters. For example, rigorously resolving the generalized Thomson problem with five particles and an arbitrary inverse-power-law potential function is a notorious unsolved problem, despite the fact that everyone agrees on what the solution must be. Vallentin discusses this problem at the end of his article, and the numerical results of de Laat that he mentions are intriguing. It would be wonderful if four-point bounds can settle this question, but why should analyzing five particles be so much more difficult technically than analyzing twelve? Two-point bounds are much easier to apply and optimize, but they do not suffice to analyze five particles.

This is by no means an isolated occurrence. For example, the solution of the sphere packing problem in three dimensions [6] is far more complicated than in eight [8] or twenty-four dimensions [5]. How difficult it is to prove optimality sometimes reflects subtle aspects of the underlying symmetry, and from this perspective the icosahedron is much better behaved than any five-particle configuration.

Levenshtein [7] proposed a beautiful theory of optimality for two-point bounds in terms of combinatorial or geometric designs. Roughly, a point configuration will attain the two-point bound and thus be optimal if its design strength is sufficiently high compared with the number of distinct distances that occur between points in the configuration. Levenshtein showed that this condition implies the packing optimality of many beautiful and important structures, and the same is true for energy minimization [4]. It does not cover every case in which the two-point bounds are sharp, but it seemingly covers almost all of them.

Three- and four-point bounds are a different story, and far fewer cases are known in which they are sharp. Initially it was unclear whether they were ever sharp, but Bachoc and Vallentin found the first sharp case [2], and a number of others have since been found (see the references in Table 2 in Vallentin's paper). However, we lack any unifying pattern or theory. Instead, based on our current knowledge each case seems to arise for idiosyncratic reasons.

This mystery would not be so troubling if semidefinite programming bounds were less central to geometric optimization, but instead they encompass most of the field's insights into packing, coding, ground states, and related problems. The sharp cases highlight many of the most beautiful and important exceptional structures in mathematics [3], and any progress in producing more such objects would be valuable.

We need an optimality theory that offers conceptual insight into where to look for further examples and how to analyze them. What are the geometric principles behind sharp bounds? Levenshtein's theory for two-point bounds offers hope that an even broader theory could be possible, and constructing it is an important open problem.

Bibliography

[1] N. N. Andreev, *An extremal property of the icosahedron*, East J. Approx. **2** (1996), 459–462.

[2] C. Bachoc and F. Vallentin, *Optimality and uniqueness of the* (4, 10, 1/6) *spherical code*, J. Combin. Theory Ser. A **116** (2009), 195–204.

[3] H. Cohn, *Order and disorder in energy minimization*, Proceedings of the International Congress of Mathematicians. Volume IV, 2416–2443, Hindustan Book Agency, New Delhi, 2010, arXiv:1003.3053.

[4] H. Cohn and A. Kumar, *Universally optimal distribution of points on spheres*, J. Amer. Math. Soc. **20** (2007), 99–148.

[5] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, *The sphere packing problem in dimension* 24, preprint, 2016, arXiv:1603.06518.

[6] T. C. Hales, *A proof of the Kepler conjecture*, Ann. of Math. (2) **162** (2005), 1065–1185.

[7] V. I. Levenshtein, *Designs as maximum codes in polynomial metric spaces*, Acta Appl. Math. **29** (1992), 1–82.

[8] M. S. Viazovska, *The sphere packing problem in dimension* 8, preprint, 2016, arXiv:1603.04246.

Henry Cohn, Microsoft Research New England, One Memorial Drive, Cambridge, MA 02142, USA. cohn@microsoft.com

IMPRINT

Editor: Volker Kaibel, Institut für Mathematische Optimierung, Otto-von-Guericke Universität Magdeburg, Universitätsplatz 2, 39108 Magdeburg, Germany. kaibel@ovgu.de ■ **Co-Editors:** Samuel Burer, Department of Management Sciences, The University of Iowa, Iowa City, IA 52242-1994, USA. samuel-burer@uiowa.edu ■ Jeff Linderoth, Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI 53706-1572, USA. linderoth@wisc.edu ■ **Founding Editor:** Donald W. Hearn ■ Published by the Mathematical Optimization Society ■ Design and typesetting by Christoph Eyrich, Berlin, Germany. optima@0x45.de ■ Printed by Oktoberdruck AG, Berlin, Germany.