Mathematisches Forschungsinstitut Oberwolfach

Report No. 20/2022

DOI: 10.4171/OWR/2022/20

Conic Linear Optimization for Computer-Assisted Proofs

Organized by
Etienne de Klerk, Tilburg
Didier Henrion, Toulouse/Prague
Frank Vallentin, Cologne
Angelika Wiegele, Klagenfurt

10 April – 16 April 2022

Abstract. From a mathematical perspective, optimization is the science of proving inequalities. In this sense, computational optimization is a method for computer-assisted proofs.

Conic (linear) optimization is the problem of minimizing a linear functional over the intersection of a convex cone with an affine subspace of a topological vector space. For many cones this problem is computationally tractable, and as a result there is a growing number of computer-assisted proofs using conic optimization in discrete geometry, (extremal) graph theory, numerical analysis, and other fields, the most famous example perhaps being the proof of the Kepler Conjecture.

The aim of this workshop was to bring researchers from these diverse fields together to work towards expanding the current scope of conic optimization as a method of generating proofs, and to identify problems and challenges to work on together.

Mathematics Subject Classification (2020): 90Cxx, 90C22, 52-xx, 14Pxx, 65-xx, 05Cxx.

Introduction by the Organizers

The workshop *Conic linear optimization for computer-assisted proofs*, organized by Etienne de Klerk, Didier Henrion, Frank Vallentin, and Angelika Wiegele, was attended by 25 participants, including one online participant. The specializations of the participants included mathematical optimization, control theory, real algebraic geometry, computer algebra, representation theory, and discrete geometry.

Here we give a short overview of the talks given at the workshop, grouped by area. We first describe the talks on recent developments in conic optimization and symmetry reduction techniques, followed by applications in the areas of graph

theory and combinatorial optimization, extremal combinatorics (via flag algebras), numerical analysis, discrete geometry, and dynamical systems.

Conic optimization. Monique Laurent (CWI Amsterdam) gave a featured (review) talk on *Recent developments in polynomial optimization*, and discussed lower and upper bounds for semidefinite programming (SDP) hierarchies using techniques from approximation theory and orthogonal polynomials.

Victor Magron (LAAS-CNRS Toulouse) gave the second featured talk on *Conic programming for certified polynomial optimization*, and described algorithms with good complexity which can be used in formal proof systems.

Georgina Hall (INSEAD Fontainebleau) discussed certificates of nonnegativity for sums of separable and quadratic polynomials, where the main results are that deciding nonnegativity of such polynomials is NP-hard, but deciding convexity is tractable.

Simone Naldi (Univ Limoges) discussed the projective geometry of conic feasibility problems.

Valentin Dannenberg (Univ Rostock) demonstrated algorithms for obtaining rational certificates of complete positivity of matrices, inspired by the classical theory of Voronoi of perfect forms (finding all critical points for lattice packings).

Symmetry reduction and representation theory. David de Laat (TU Delft) gave the featured talk on *Exploiting symmetry in conic optimization*, and discussed applications to spherical codes, the Lovász theta function, the Delsarte-Goethals-Seidel 2-point linear programming bound, and Bochner's theorem applied to the Bachoc-Vallentin 3-point SDP bound. He also described a new toolbox for extracting exact SDP bounds and computer-assisted proofs for the optimality and uniqueness of point configurations.

Mohab Safey El Din (Sorbonne Univ Paris) described an exact algorithm for symmetric polynomial optimization problems.

Philippe Moustrou (IMT Univ Toulouse) discussed symmetry reduction techniques in the context of AM/GM optimization.

Dmitrii Pasechnik (Univ Oxford) showed new results in exact computations with group representations.

Graph theory and combinatorial optimization. Elisabeth Gaar (Johannes Kepler Univ Linz) gave the talk *Towards a Computer-Assisted Proof for a Conjecture from Graph Theory*, where an algebraic reformulation of Vizing's conjecture on dominating sets in graphs is used to certify Vizing's conjecture via positivity of a class of polynomials on a class of varieties.

Renata Sotirov (Tilburg Univ) characterized the Chvátal-Gomory closure of spectrahedra, and applied this theory to a branch-and-cut framework using Chvátal-Gomory cuts for integer SDP problems.

Konstantin Golubev (ETH Zurich) extended the Hoffman eigenvalue bound from graphs to hypergraphs and considered applications in extremal combinatorics (Erdős-Ko-Rado theorem for intersecting families, their p-biased versions, Mantel's theorem, and Frankl's problem on triangle free families).

Fernando de Oliveira (TU Delft) introduced an extension of the Lovász theta number from graphs to r-uniform hypergraphs, deriving fundamental properties of the recursive theta body, relations to the Hoffman eigenvalue bound for hypergraphs, and applications in extremal graph theory.

Extremal combinatorics: flag algebras. Fernando de Oliveira (TU Delft) gave a featured talk on *The use of flag algebras in proofs*, reviewing Razborov's flag algebra approach for problems in extremal graph theory (presented examples: Caccetta-Häggkvist conjecture, Mantel's theorem), and relations to Lasserre's hierarchy and to the theory of graph limits by Lovász and Szegedy.

Daniel Brosch (Tilburg Univ) spoke on the symmetries of flag algebras, and introduced a different and simplified version of Razborov's flag algebra.

Numerical analysis. Adrien Taylor (INRIA Paris) gave the featured talk on the *Analysis of worst-case performance of iterative methods via conic optimization*. He reviewed the use of SDPs to prove estimates for performance measures of first-order methods, and described the development of a toolbox which can be used for the analysis of many different first-order methods.

Felix Kirschner (Tilburg Univ) spoke about the construction of multivariate polynomial approximation kernels via semidefinite programming, and showed how kernels of the Jackson-type may be extended to the multivariate case.

Discrete geometry. Andreas Spomer (Univ Cologne) described semidefinite programming bounds for the kissing number problem for regular tetrahedra and for more general packings of regular n-gons on the unit sphere S^2.

Maria Dostert (KTH Stockholm) considered semidefinite programming bounds for the average kissing number, and described explicit computation of new bounds by SDP, including rigorous certification by rounding the SDP.

David de Laat (TU Delft) spoke on the computation of three-point bounds for sphere packings in Euclidean space $\mathbb{R}^n$ using SDP, and showed improved bounds in dimension $n = 4, 5, 6, 7, 9$. He made a link to recent developments of the conformal bootstrap program in theoretical physics.

Control and dynamical systems. Jared Miller (Northeastern Univ Boston) discussed bounding the minimum distance to unsafe sets in control theory using the moment hierarchy.

Milan Korda (LAAS-CNRS Toulouse) showed how to find Lyapunov functions (by SDP) to certify stability of dynamical systems controlled by neural networks.

Acknowledgement: The organizers gratefully acknowledge that three participants received travel support through the Oberwolfach Leibniz Graduate program.

Workshop: Conic Linear Optimization for Computer-Assisted Proofs

Table of Contents

Abstracts

Recent developments on convergence analysis of sum-of-squares hierarchies in polynomial optimization

Monique Laurent

We consider polynomial optimization problems of the form

$$f_{\min} = \min\{f(x) : x \in K\}, \quad \text{where } K = \{x \in \mathbb{R}^n : g_j(x) \geq 0 \ (j \in [m]\}$$

with $f, g_j \in \mathbb{R}[x]$ multivariate polynomials. Such problems are in general nonlinear, nonconvex hard problems, already for simple sets K like the ball, the unit sphere, the hypercube or the simplex, since they can capture NP-hard combinatorial problems like max-cut and maximum stable set problems in graphs. Hierarchies of upper and lower bounds can be constructed using the cone Σ of sums of squares of polynomials. For an integer $r \in \mathbb{N}$ define the parameters

$$(1) \qquad f^{(r)} = \inf\Big\{ \int_K f(x)\sigma(x)d\mu(x) : \int_K \sigma(x)d\mu(x) = 1,\ \sigma \in \Sigma,\ \deg(\sigma) \leq 2r\Big\},$$

$$(2) \qquad f_{(r)} = \sup\Big\{\lambda : f - \lambda = \sum_{j=0}^{m} \sigma_j g_j,\ \sigma_j \in \Sigma,\ \deg(\sigma_j g_j) \leq 2r\Big\},$$

$$(3) \quad f_{(r),sch} = \sup\Big\{\lambda : f - \lambda = \sum_{J \subseteq [m]} \sigma_J g_J,\ \sigma_J \in \Sigma,\ \deg(\sigma_J g_J) \leq 2r\Big\}.$$

Here μ is a positive Borel measure with support K and we set $g_0 = 1$, $g_J = \prod_{j \in J} g_j$. We have $f_{(r)} \leq f_{(r),sch} \leq f_{\min} \leq f^{(r)}$ and each parameter $f^{(r)}$, $f_{(r)}$, $f_{(r),sch}$ can be expressed as a semidefinite program. When K is compact the bounds $f^{(r)}$ and $f_{(r),sch}$ are known to converge asymptotically to $f_{\min}$, as well as $f_{(r)}$ under an Archimedean condition [5, 6]. In this lecture we discuss the state-of-the-art results concerning the quality of the upper and lower bounds, namely how fast the error ranges $f^{(r)} - f_{\min}$, $f_{\min} - f_{(r),sch}$ and $f_{\min} - f_{(r)}$ tend to 0 as a function of the order r of the relaxation. In both cases the results can be divided into two categories: for general sets K, and for special sets K such as the ball, the unit sphere, the hypercube or the simplex. Different techniques are used for each category and, naturally, stronger results can be shown for special sets. In addition there is an intimate link between the analysis for both upper and lower bounds.

Convergence analysis of the upper bounds. The most general result is when the set K is a convex body, or a semialgebraic set with a dense interior (selecting the Lebesgue measure for μ); then one can show $f^{(r)} - f_{\min} = O\big(\frac{\log^2 r}{r^2}\big)$ [10]. The analysis has two key steps: (1) reduce the search to a univariate sum of squares s and then set $\sigma(x) = s(f(x))$, (2) select the univariate sum of squares s as a tight approximation of the Dirac delta at an extremity of an interval (using so-called needle polynomials).

A stronger analysis in $O\left(\frac{1}{r^2}\right)$ can be shown for special sets: for the interval $[-1,1]$ (with $d\mu(x) = (1-x^2)^{-1/2}dx$ in [3], with $d\mu(x) = (1-x^2)^{\lambda}$ and $\lambda \geq -1/2$ in [10]), for the ball and the simplex in [10], and for the sphere in [4]. The starting point is expressing the parameter $f^{(r)}$ as the smallest eigenvalue of the matrix $A_f = (\int_K f(x)p_\alpha(x)p_\beta(x)d\mu(x))_{|\alpha|,|\beta|\leq r}$, where $\{p_\alpha\}$ are orthonormal polynomials with respect to the inner product provided by the measure μ on K. In the univariate case, when $K = [-1,1]$ and $f(x) = x$, the matrix A_x coincides with the (tridiagonal) Jacobi matrix, whose entries are given by the three-term recurrence satisfied by the polynomials p_k, and whose eigenvalues are the roots of the polynomial p_{r+1}. Hence, in this case, $f^{(r)}$ coincides with the least root of p_{r+1}, known to be in the order $-1 + \Omega\left(\frac{1}{r^2}\right)$ for Jacobi type measures $(1-x^2)^{\lambda}$ ($\lambda > -1$). This is the key ingredient used in [3] for the analysis in the case when $K = [-1,1]^n$. Additional reductions are needed to extend the analysis to other sets. This includes an integration trick in order to extend the analysis to the sphere in [4] and looking at the local shape around a global minimizer to extend the analysis to the ball, the simplex and 'round' convex bodies in [10].

Convergence analysis of the lower bounds. The best general result shows a convergence rate in $O\left(\frac{1}{r^c}\right)$ for the bounds $f_{(r)}$ [1], improving an earlier result in [8] (with a logarithmic dependence on r). The analysis combines a variety of techniques, including the Lojasiewicz inequality and approximation theory tools, and a reduction to the analysis for the case of the hypercube $[-1,1]^n$ (in [7]).

For special sets such as the hypercube $[-1,1]^n$, the ball, the sphere and the simplex, a better performance analysis in $O\left(\frac{1}{r^2}\right)$ can be shown for the bounds $f_{(r),sch}$ (that involve richer sum of squares decompositions). A key ingredient here is using the *polynomial kernel method* to construct suitable sum of squares representations. Here is a general 'recipe'. Assume we have an invertible linear operator $\mathbf{K}_r : \mathbb{R}[x]_r \to \mathbb{R}[x]_r$ that satisfies the following properties:

- (P1) $\mathbf{K}_r$ preserves the constant polynomial: $\mathbf{K}_r 1 = 1$,
- (P2) if $p \geq 0$ on K and $\deg(p) \leq r$, then $\mathbf{K}_r p = \sum_J \sigma_J g_J$, where $\sigma_J \in \Sigma$ and $\deg(\sigma_J g_J) \leq 2r$,
- (P3) $\|\mathbf{K}_r^{-1} f - f\|_\infty \leq \epsilon$.

It is not difficult to see that one can conclude $f_{\min} - f_{(r),sch} \leq \epsilon$. The challenge then is constructing a linear operator $\mathbf{K}_r$ having these properties. Property (P3) says (roughly) that $\mathbf{K}_r$ should be close to the identity operator when acting on polynomials of degree at most $d = \deg(f)$. Following the polynomial kernel method, one can select a polynomial kernel $K_r(x,y) \in \mathbb{R}[x,y]$ and then define the corresponding linear operator that acts by convolution: for $p \in \mathbb{R}[x]_r$, define $\mathbf{K}_r p(x) = \int_K p(y)K_r(x,y)d\mu(y)$.

Consider first the case of the hypercube $K = [-1,1]^n$, equipped with the Chebyshev product measure $d\mu(x) = \prod_{i=1}^n (1-x_i^2)^{-1/2}dx_i$. Consider the multivariate Chebyshev polynomials $T_\alpha = \prod_{i=1}^n T_{\alpha_i}$ where T_k are the univariate Chebyshev polynomials. Then the T_α are orthogonal w.r.t. the inner product given by μ. Define the multivariate polynomial kernel $K_r(x,y) = \prod_{i=1}^n K_r(x_i,y_i)$, where the

univariate kernel is of the form $K_r(x_i, y_i) = 1+2\sum_{k=1}^r \lambda_k^r T_k(x_i)T_k(y_i)$ (also known as the Jackson kernel), and the scalars λ_k^r are selected so that $0 \leq 1-\lambda_k^r \leq \frac{\pi^2 d^2}{(r+2)^2}$ for $0 \leq k \leq d$ and $K_r(x_i, y_i) \geq 0$ on $[-1,1]^2$. By construction, the corresponding operator $\mathbf{K}_r$ satisfies (P1) and, using Markov-Lukacz, it also satisfies (P2). Finally, the selection of the scalars λ_k^r ensures that (P3) holds with $\epsilon = O\big(\frac{1}{r^2}\big)$, thus giving the analysis in $O\big(\frac{1}{r^2}\big)$ (as in [7]).

The same analysis in $O\big(\frac{1}{r^2}\big)$ is shown for the sphere in [2], the ball and the simplex in [9]. It however relies on a different class of polynomial kernels, constructed by exploiting the symmetry structure of the sets under consideration. The idea (pioneered in [2]) is construct the kernel $K_r(x, y)$ as a suitable perturbation of the Christoffel-Darboux kernel, as $K_r(x, y) = \sum_{k=0}^r \lambda_k \sum_{|\alpha|=k} p_\alpha(x)p_\alpha(y)$, for appropriate scalars λ_k. Setting $\lambda_0 = 1$ ensures (P1) holds. The key (harmonic analysis) fact is that the sum $\sum_{|\alpha|=k} p_\alpha(x)p_\alpha(y)$ can be expressed in terms of an associated univariate polynomial, when dealing with the sphere, the ball or the simplex.

In the case of the unit sphere, this univariate polynomial is a Gegenbauer polynomial G_k evaluated at $x^T y$. Then, one selects $K_r(x, y) = s(x^T y)$, where $s(t) = \sum_{k=0}^r \lambda_k G_k(t)$, and one searches for scalars λ_k so that s is a sum of squares (which ensures (P2) holds) and (P3) holds. Interestingly, the search for this univariate s boils down to an instance of the upper bounds $g^{(r)}$ for a suitably defined univariate g on the interval $[-1, 1]$. As a consequence, it follows from the earlier results on the upper bounds that a suitable s can be found, which gives a suitable polynomial kernel $K_r(x, y)$ and thus a suitable kernel operator $\mathbf{K}_r$.

For the ball and the simplex, a similar approach can be followed, however with more technical details as the summation formulas for $\sum_{|\alpha|=k} p_\alpha(x)p_\alpha(y)$ (given in [11, 12]) are more involved. The full details can be found in [9].

References

[1] L. Baldi and B. Mourrain. *On Moment Approximation and the Effective Putinar's Positivstellensatz*, arXiv:2111.11258 (2021).

[2] K. Fang and Fawzi. *The sum-of-squares hierarchy on the sphere, and applications in quantum information theory.* Mathematical Programming, 2020.

[3] E. de Klerk and M. Laurent. *Worst-case examples for Lasserre's measure–based hierarchy for polynomial optimization on the hypercube*, Mathematics of Operations Research, **45(1)** (2020) 86-98.

[4] E. de Klerk and M. Laurent. *Convergence analysis of a Lasserre hierarchy of upper bounds for polynomial minimization on the sphere*, Mathematical Programming (2020).

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

[6] J.B. Lasserre. *A New Look at Nonnegativity on Closed Sets and Polynomial Optimization*, SIAM Journal on Optimization, **21(3)** (2010), 864–885.

[7] M. Laurent and L. Slot. An effective version of Schmüdgen's Positivstellensatz for the hypercube. arXiv:2109.09528 (2020).

[8] J. Nie and M. Schweighofer. *On the complexity of Putinar?s Positivstellensatz.* Journal of Complexity **23(1)** (2007) 135–150.

[9] L. Slot. Sum-of-squares hierarchies for polynomial optimization and the Christoffel-Darboux kernel. arXiv:2111.04610 (2021).

[10] L. Slot and M. Laurent. *Improved convergence analysis of Lasserre's measure–based upper bounds for polynomial minimization on compact sets*, Mathematical Programming (2020)
[11] Yuan Xu. *Summability of Fourier orthogonal series for Jacobi weight functions on the simplex in* $\mathbb{R}^d$. Proc. Amer. Math. Soc. **126(10)** (1998) 3027–3036.
[12] Yuan Xu. *Summability of Fourier orthogonal series for Jacobi weight on a ball in* $\mathbb{R}^d$. Trans. Amer. Math. Soc. **351(6)** (1999) 2439–2458.

Conic programming for certified polynomial optimization

Victor Magron
(joint work with Mohab Safey El Din, Markus Schweighofer, Trung-Hieu Vu, Timo de Wolff, Heining Seidler, Jie Wang)

In general, certified algorithms provide a way to ensure the safety of several systems in engineering sciences, program analysis as well as cyber-physical critical components. Since these systems often involve nonlinear functions, such as polynomials, it is highly desirable to design certified polynomial optimization schemes and to be able to interpret the behaviors of numerical solvers implementing these schemes.

In this lecture talk, we describe, analyze and compare both from the theoretical and practical points of view, several algorithms computing nonnegativity certificates with rational coefficients.

We start to focus on the case of nonnegative univariate polynomials with rational coefficients and provide two algorithms computing weighted sums of squares (SOS) decompositions. The first algorithm, due to Schweighofer, relies on real root isolation, quadratic approximations of positive polynomials and square-free decomposition. We provide bit complexity estimates, both on runtime and output size of this algorithm. They are exponential in the degree of the input univariate polynomial and linear in the maximum bitsize of its complexity. This analysis is obtained using quantifier elimination and root isolation bounds. The second algorithm, due to Chevillard, Harrison, Joldes and Lauter, relies on complex root isolation and square-free decomposition and has been introduced for certifying positiveness of polynomials in the context of computer arithmetics. We provide bit complexity estimates, both on runtime and output size of this algorithm, which are polynomial in the degree of the input polynomial and linear in the maximum bitsize of its complexity. This analysis is obtained using Vieta's formula and root isolation bounds. We extend this framework to the case of trigonometric polynomials with Gaussian integer coefficients, i.e., with real and imaginary parts being integers.

Then, we consider the problem of computing exact SOS decompositions for certain classes of non-negative multivariate polynomials, relying on semidefinite programming (SDP) solvers. We provide a hybrid numeric-symbolic algorithm computing exact rational SOS decompositions with rational coefficients for polynomials lying in the interior of the SOS cone. The first step of this algorithm computes an approximate SOS decomposition for a perturbation of the input polynomial with an arbitrary-precision SDP solver. Next, an exact SOS decomposition

is obtained thanks to the perturbation terms and a compensation phenomenon. We prove that bit complexity estimates on output size and runtime are both singly exponential in the cardinality of the Newton polytope (or doubly exponential in the number of variables). Next, we apply this algorithm to compute exact Reznick, Hilbert-Artin's representation and Putinar's representations respectively for positive definite forms and positive polynomials over basic compact semi-algebraic sets. We also report on practical experiments done with the implementation of these algorithms and existing alternatives such as the critical point method and cylindrical algebraic decomposition. Then we extend this framework to nonnegative polynomials with rational coefficients, assuming that they reach their infimum and have associated gradient ideal being zero-dimensional radical. We prove that, under these assumptions, nonnegativity is equivalent to be an SOS of rational polynomials modulo the gradient ideal. We provide exact algorithms to compute them, and show that their bit complexity is singly exponential in the number of variables and polynomial in the degree.

Next, we provide two hybrid numeric-symbolic optimization algorithms, computing exact sums of nonnegative circuits (SONC) and sums of arithmetic-geometric-exponentials (SAGE) decompositions. Moreover, we provide a hybrid numeric-symbolic decision algorithm for polynomials lying in the interior of the SAGE cone. Each framework, inspired by previous contributions of Parrilo and Peyrl, is a rounding-projection procedure. For a polynomial lying in the interior of the SAGE cone, we prove that the decision algorithm terminates within a number of arithmetic operations, which is polynomial in the number of terms of the input, and linear in the distance to the boundary of the cone. We also provide experimental comparisons regarding the implementation of the two optimization algorithms.

Eventually, we focus on computing SONC decompositions with second-order cone (SOC) programming. We prove constructively that the cone of sums of nonnegative circuits (SONC) admits a SOC representation. Based on this, we give a new algorithm for unconstrained polynomial optimization via SOC programming. We also provide a hybrid numeric-symbolic scheme which combines the numerical procedure with a rounding-projection algorithm to obtain exact nonnegativity certificates.

Sums of Separable Plus Quadratic Polynomials

Georgina Hall
(joint work with Amir Ali Ahmadi, Cemil Dibek)

In this talk, we considered *separable plus quadratic* (SPQ) polynomials, i.e., polynomials that are the sum of univariate polynomials in different variables and a quadratic polynomial.

Motivated by the fact that nonnegative separable and nonnegative quadratic polynomials are sums of squares, we first study whether nonnegative SPQ polynomials are the sum of a nonnegative separable and a nonnegative quadratic polynomial,

which would imply that any nonnegative SPQ polynomial is actually a sum of squares. We establish that, in fact, the answer to this question is positive for univariate plus quadratic polynomials and for convex SPQ polynomials only, and negative already for bivariate quartic SPQ polynomials. We further use our decomposition result for convex SPQ polynomials to show that convex SPQ polynomial optimization problems can be solved by semidefinite programs which are much smaller than those obtained via the first level of the Lasserre hierarchy.

We then investigate whether nonnegative SPQ polynomials are sums of squares and provide a complete characterization of when this is the case, based on the degree and number of variables of the SPQ polynomial. Interestingly, the cases where nonnegative SPQ polynomials are sums of squares are exactly the same as those established by Hilbert for non-SPQ polynomials. The examples we produce of SPQ polynomials which are nonnegative but not sum of squares are obtained in an automated fashion. Furthermore, the proofs that these polynomials are indeed nonnegative but not sum of squares are computer assisted.

Though nonnegative SPQ polynomials are not sums of squares in general, this does not necessarily preclude the existence of a polynomial-time algorithm for testing nonnegativity of SPQ polynomials. We thus investigate this question next and show that testing nonnegativity of SPQ polynomials is already NP-hard when the degree is at least four, via a reduction from Partition.

Finally, we conclude with a generalization of Newton's method which leverages SPQ polynomials. Rather than approximating the function to minimize by a quadratic polynomial, we instead approximate the function to minimize by an SPQ polynomial. This method gives rise experimentally to larger basins of attraction when compared to Newton's method.

On the projective geometry of conic feasibility problems

Simone Naldi
(joint work with Rainer Sinn)

This extended abstract is based on the article [7] to which we refer for a more complete list of contributions on this topic.

Let $\mathbf{K} \subset U$ be a closed convex pointed cone with non-empty interior in a real vector space U, and let $L \subset U$ be an affine subspace. The *conic feasibility problem* is the algorithmic question of deciding whether $\mathbf{K} \cap L = \emptyset$ or $\mathbf{K} \cap L \neq \emptyset$.

Several classical questions can be cast as instances of conic feasibility problems, let us briefly mention a few. The existence of solutions to a system of affine inequalities

$$a_{i1}x_1 + \cdots + a_{in}x_n \geq b_i, \quad i = 1, \ldots, m$$

is the feasibility problem of *linear programming*, that is, the question whether an affine space $L \subset U = \mathbb{R}^n$ intersects the n-dimensional nonnegative orthant

$\mathbf{K} = \mathbb{R}^n_+$. The solvability of a linear matrix inequality

$$A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0$$

for given real symmetric matrices $A_0, \ldots, A_n$, is the feasibility problem of *semidefinite programming*: in this case $\mathbf{K} = \mathbb{S}^d_+$, the cone of positive semidefinite real symmetric matrices, and L is an affine space in $U = \mathbb{S}^d$. Other important classes arise in combinatorial optimization (*e.g.* copositive, completely positive matrices), real algebraic geometry (moment cones, cone of positive polynomials, SOS cone) and in several other areas.

One classically makes the distinction between different shades, or *types*, of feasibility and infeasibility. A feasible conic program is called *strongly feasible* (resp. *weakly feasible*) if the affine space L contains (resp. does not contain) an interior point of $\mathbf{K}$. On the other side, an infeasible conic program is called *strongly infeasible* if the Euclidean distance $d(\mathbf{K}, L)$ is strictly positive, and it is called *weakly infeasible* if $d(\mathbf{K}, L) = 0$.

An example of a weakly feasible set in semidefinite programming is the Gram spectrahedron of a sum-of-squares polynomial which has no rational Gram matrices, for instance one of the polynomials given by Scheiderer in [9]:

$$x^4 + xy^3 + y^4 - 3x^2yz - 4xy^2z + 2x^2z^2 + xz^3 + yz^3 + z^4.$$

An example of weakly infeasible semidefinite program (see [8]) is given by the linear matrix inequality

$$\begin{bmatrix} 0 & 1 \\ 1 & x \end{bmatrix} \succeq 0$$

since the previous inequality has no solutions, but the set $\left\{ \left[\begin{smallmatrix} 1/n & 1 \\ 1 & n \end{smallmatrix}\right] : n \in \mathbb{N} \right\}$ is included in the cone $\mathbf{K} = \mathbb{S}^2_+$ and has distance zero from the line $L = \{ \left[\begin{smallmatrix} 0 & 1 \\ 1 & x \end{smallmatrix}\right] : x \in \mathbb{R} \}$. Let us also mention that such degenerate programs might seem artificial, but actually can arise quite naturally in concrete situations, for instance in the case of Lasserre relaxations of the minimization of the Motzkin polynomial [10].

The existence of "weak types" makes the feasibility problem harder, since for these types, the feasibility is not robust under small perturbations of the input data: a weakly feasible program will become either strongly feasible or strongly infeasible after small perturbation, and the same for a weakly infeasible program. It is a typical behavior of numerical solvers that of giving a wrong answer in these degenerate situations.

The main contribution of our work [7] is the construction of a general framework for homogenizing a conic program. Indeed, different algorithms for solving conic (feasibility) problems are based on homogenization of the constraints, see for instance [11, 4] or [3, Ch. 4] and references therein. The idea at the core of our method is as follows: the vector space U is seen as an affine hyperplane in a vector space V, and $\mathbf{K}$ is *lifted* to a cone $\widehat{\mathbf{K}} \subset V$, such that $\widehat{\mathbf{K}} \cap U = \mathbf{K}$. The idea of this projective point of view is that the original cone $\mathbf{K}$ is now the part of the homogeneous cone $\widehat{\mathbf{K}}$ that *can be observed* if one sits in the affine chart U of $\mathbb{P}(V)$.

In this framework, it is possible to get information about the feasibility type of the original program $\mathbf{K} \cap L = \widehat{\mathbf{K}} \cap L$ and the *homogenized conic program* $\widehat{\mathbf{K}} \cap \mathrm{span}(L)$. One example is the following characterization of infeasibility:

$$\mathbf{K} \cap L = \emptyset \text{ and } (-\mathbf{K}) \cap L = \emptyset \;\Leftrightarrow\; \widehat{\mathbf{K}} \cap \mathrm{span}(L) \subset \mathrm{lin}(L) \tag{1}$$

where $\mathrm{lin}(L)$ is the direction of L: this tells us that L does not meet neither $\mathbf{K}$ nor $(-\mathbf{K})$ exactly when the homogenized feasible set lies at infinity.

Another interesting caracterization concerns a special subclass of strongly infeasible programs. A conic program $\mathbf{K} \cap L$ is called *stably infeasible* if there exists a neighbourhood $\mathcal{L}$ of L in the Grasmanniann of affine subspaces of U of the same dimension as L, such that if $L' \in \mathcal{L}$, then $\mathbf{K} \cap L' = \emptyset$. In other words, stably infeasible conic programs are strongly infeasible and keep this property after small perturbations. However it is easy to construct strongly infeasible conic programs that are not stable, already in the class of linear programming.

A second aspect that has been highlighted in [7] and in the talk at MFO, is to which extent the infeasibility can be certified over the field of definition of the conic program. Indeed, the expected output of a decision algorithm is a yes-no answer and in the case of conic infeasibility, one could also ask for a *rational infeasibility certificate*, that is, an element $\ell \in \mathbf{K}^*$ (in the dual cone $\mathbf{K}^* \subset V^\vee$), which is (nonnegative on $\mathbf{K}$ and) strictly negative on L, and that can be defined over the smallest field containing the input. The element ℓ is a separating hyperplane for $\mathbf{K}$ and L and the question is wheter such certificate can be made rational. Let us remarl that this is a natural question for infeasibility, since the same question has in general a "no" answer for the feasibility of non-linear conic programs, as already mentioned for the examples in [9].

In [7], we prove two main results concerning rationality of infeasibility certificates: first, every stably infeasible conic program admits a rational infeasibility certificate; second, we construct explicit strongly infeasible semidefinite programs that do not admit rational infeasibility certificates.

Finally, let us mention two last aspects of our contribution. The first concerns the property mentioned in (1), which is used in [7] to design a variant of the classical facial reduction algorithm [2], for the case of infeasible conic programs: if $\mathbf{K}$ is a nice cone, and if $\mathbf{K} \cap L = (-\mathbf{K}) \cap L = \emptyset$, then there exists a sequence of functionals $\ell_1, \ldots, \ell_k \in \mathbf{K}^*$, $k \leq 1 + \dim L$, that facially reduce the homogenized program by sending the feasible set $\widehat{\mathbf{K}} \cap \mathrm{span}(L)$ to infinity (see [7, Th. 3.4]).

The second concerns the complexity theory of semidefinite programming. The homogenization we have described allows us to give an alternative proof of the result by Ramana [6] that feasibility of semidefinite programming is in $\mathrm{coNP}_{\mathbb{R}}$ (Blum-Shub-Smale complexity model [1]), see [7, Sec. 4].

References

[1] L. Blum, F. Cucker, M. Shub, S. Smale. *Complexity and Real Computation*, Springer-Verlag New York Inc. (1998).

[2] J.M. Borwein, H. Wolkowicz. *Facial reduction for a cone-convex programming problem*, J. Aust. Math. Soc. **30:3** (1981).

[3] E. de Klerk, *Aspects of semidefinite programming*, Springer (2010).
[4] E. de Klerk, C. Roos, T. Terlaky *Initialization in semidefinite programming via a self-dual, skew-symmetric embedding*, Op. Res. Lett. **20** (1997), 213–221.
[5] I. Klep, M. Schweighofer. *An exact duality theory for semidefinite programming based on sums of squares*, Math. Oper. Res. **38:3** (2013), 569–590.
[6] M.V. Ramana. *An exact duality theory for semidefinite programming and its complexity implications*, Math. Prog. Serie B **77:2** (1997), 129–162.
[7] S. Naldi, R. Sinn. *Conic programming: Infeasibility certificates and projective geometry*, Journal of Pure and Applied Algebra **225:7** (2021), 106606.
[8] G. Pataki, *Bad semidefinite programs: they all look the same*, SIAM J. Optim. **27:1** (2017), 146–172.
[9] C. Scheiderer, *Sums of squares of polynomials with rational coefficients*, J. Eur. Math. Soc. (2016), 1495–1513.
[10] H. Waki, *How to generate weakly infeasible semidefinite programs via Lasserre relaxations for polynomial optimization*, Optim. Lett. **6:8** (2012), 1883-1896.
[11] Y. Ye, M.J. Todd, S. Mizuno *An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm*, Math. Oper. Res. **19:1** (1994), 53–67.

Obtaining Rational $\mathcal{CP}$-Certificates via Perfect (Generalized) Copositive Matrices

Valentin Dannenberg

An interesting subfield of study in recent years is the cone of *completely positive matrices*

$$\mathcal{CP}^n = \left\{Q \in \mathcal{S}^n : \ Q = \sum_{i=1}^{k} x_i x_i^T, \ k \in \mathbb{N}, \ x_i \in \mathbb{R}^n_{\geq 0}, i = 1, \ldots, k\right\},$$

which, together with its dual cone

$$\mathcal{COP}^n = \left\{Q \in \mathcal{S}^n : \ x^T Q x \geq 0 \text{ for all } x \geq 0\right\}$$

of *completely positive matrices*, can be used to reformulate various difficult optimization problems as convex conic ones, see e.g. [1].

One fundamental computational problem in this area is obtaining a rational factorization of a rational completely positive matrix, if it exists. In other words for a given rational completely positive matrix Q, we want to compute nonnegative rational vectors $x_1, \ldots, x_k$, such that $Q = \sum_{i=1}^{k} x_i x_i^T$. Since such a factorization gives a proof that Q is indeed completely positive, it is also called a (rational) $\mathcal{CP}$-certificate.

Recently a simplex-like algorithm has been proposed by Dutour Sikirić, Schürmann and Vallentin in [3], that can compute such a factorization, whenever it exists.

The talk is centered around explaining and visualizing the key concepts of the algorithm. The main one is that of *perfect copositive matrices*. Here we call a matrix $P \in \text{int}\mathcal{COP}^n$ perfect copositive, if it is the unique solutions to the linear equation system

$$x^T P x = \min_{\mathcal{COP}}(P), \ x \in \text{Min}_{\mathcal{COP}}(P),$$

where $\min_{\mathcal{COP}}(P) = \min_{x \in \mathbb{Z}^n_{\geq 0} \setminus \{0\}} x^T P x$ is the *copositive minimum* of P and $\mathrm{Min}_{\mathcal{COP}}(P) = \{x \in \mathbb{Z}^n_{\geq 0} \setminus \{0\} : \ x^T P x = \min_{\mathcal{COP}}(P)\}$ is the set of *minimal vectors* of P. These definitions essentially mirror the classical case of positive definite matrices, see also [4]. The copositive perfect matrices (with copositive minimum 1) form the vertex set of

$$\mathcal{R}_{\mathcal{COP}} = \{Q \in \mathcal{S}^n : \ \min_{\mathcal{COP}}(P) \geq 1\}.$$

It can further be shown that every rational completely positive matrix, which possesses a rational $\mathcal{CP}$-factorization, lies in the set

$$\bigcup_{P \text{ perfect copositive, } \min_{\mathcal{COP}}(P) = 1} \mathrm{cone}\{xx^T : \ x \in \mathrm{Min}_{\mathcal{COP}}(P)\}.$$

Therefore, the simplex-like nature of the algorithm consists in visiting the vertices of $\mathcal{R}_{\mathcal{COP}}$, and checking whether a given matrix Q factorizes over the minimal vectors of that vertex.

Using these ideas, we also present an adaption of this algorithm in [3], to solve the *$\mathcal{CP}$-membership problem*: Given a rational matrix Q, decide if Q is completely positive. A nice property of the resulting algorithm is, that it produces rational certificates for both the cases $Q \in \mathcal{CP}^n$ and $Q \notin \mathcal{CP}^n$.

We conclude by presenting recent new results as well as questions (and partial answers):

- Does every rational completely positive matrix possess a rational $\mathcal{CP}$-certificate?
  - Yes, for matrices in the interior of $\mathcal{CP}^n$, see [2].
  - Open, for matrices on the boundary, see [6] for a recent survey on this topic
- Does the $\mathcal{CP}$-membership algorithm always terminate in finitely many steps?
  - Yes, for $n = 2$
  - Open, for $n \geq 3$
- Recently, the speaker was able to obtain a new discretization of $\mathcal{CP}^n$,

$$\mathcal{CP}^n = \{Q \in \mathcal{N}^n : \ \mathrm{Trace}(PQ) \geq 0 \text{ for all perfect copositive matrices } P\},$$

  which refines a similar one given in [3]. Using it we were able to remove certain costly copositivity tests in the algorithm
- New rough approaches and ideas for (better) computing the copositive minimum as well as optimizing other steps in the algorithm
- New results on the topic of perfect copositive matrices, such as infinite non-equivalent series of perfect copositive matrices, that will also appear in [5]

References

[1] I.M. Bomze, M. Dür, E. de Klerk, C. Roos, A.J. Quist and T. Terlaky, *On Copositive Programming and Standard Quadratic Optimization Problems*, J. Global Optim. **18** (2000), 301–320.

[2] M. Dutour Sikirić, A. Schürmann and F. Vallentin, *Rational factorizations of completely positive matrices*, Linear Algebra Appl. **523** (2017), 46-51.

[3] M. Dutour Sikirić, A. Schürmann and F. Vallentin, *A simplex algorithm for rational cp-factorization*, Math. Prog. **187** (2021), 25–45.

[4] A. Schürmann, *Computational geometry of positive definite quadratic forms*, Amer. Math. Soc. (2009).

[5] A. Schürmann and V. Dannenberg, *Perfect Copositive Matrices*, in preparation.

[6] N. Shaked-Monderer, M. Dür and A. Berman, *Complete positivity over the rationals*, Pure and Appl. Func. Anal **3** (2018), 681 - 691

Exploiting symmetry in conic optimization

David de Laat

Conic optimization problems can often be simplified significantly by exploiting symmetries in the formulation of the problems. In this tutorial two examples are considered, and connections are made to recent developments.

In the most general situation we consider conic programs of the form

$$\begin{aligned} \text{minimize} \quad & L(x) \\ \text{subject to} \quad & x \in K, \\ & Ax = b \end{aligned}$$

where V is a locally convex topological vector space, K is a closed, convex cone in V, and $L\colon V \to \mathbb{R}$ is a continuous linear functional. Let Γ be a compact group with a continuous, linear action on V. The conic program is said to be Γ-invariant if $L(\gamma x) = L(x)$, $\gamma x \in K$, and $A\gamma x = b$ for all $x \in K$ and $\gamma \in \Gamma$. Then it follows that for feasible x, the vector $\overline{x}$ given by the weak vector-valued integral $\int_\Gamma \gamma x \, d\gamma$ over the Haar measure of Γ, is also feasible and has the same objective value. This shows we may restrict to invariant vectors without changing the optimal objective value.

In the first part of the tutorial we consider positive kernels invariant under the action of a compact group. Here we start by considering positive, continuous kernels $K\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ on the unit sphere that are invariant under the action of the orthogonal group $O(n)$. By Schoenberg's theorem [11] such kernels $K(x, y)$ can be written as uniformly absolutely converging series with nonnegative coefficients of Gegenbauer polynomials evaluated in $x \cdot y$. This can be used to show the Lovász theta number for the spherical code problem reduces to the Delsarte-Goethals-Seidel bound after symmetry reduction [1].

Under the action of $O(n)$ we can decompose the space $\mathcal{C}(S^{n-1})$ of continuous functions into irreducible subspaces H_k consisting of the spherical harmonics of degree k. With bases $e^n_{k,1}(x), \ldots, e^n_{k,d_k}(x)$ of H_k that are orthonormal with respect

to an $O(n)$-invariant inner product, the kernels

$$(x, y) \mapsto \sum_{j=1}^{d_k} \overline{e_{k,j}^n(x)} e_{k,j}^n(y)$$

are positive and $O(n)$-invariant. Using Bochner's characterization it follows each invariant, positive, continuous kernel is an absolutely uniformly converging series of such kernels with nonnegative coefficients. Schoenberg's result is then recovered using the addition formula, which shows the above kernels are exactly the Gegenbauer polynomials evaluated in $x \cdot y$.

In general Bochner's characterization applies for a compact group Γ acting transitively on X [4]. Then we get a decomposition

$$\mathcal{C}(X) = \bigoplus_{\pi} \bigoplus_{i=1}^{m_\pi} H_{\pi,i},$$

where π ranges over the irreducible representations of Γ, and where $H_{\pi,i}$ and $H_{\pi',i'}$ are equivalent as representations if $\pi = \pi'$. Let $e_{\pi,i,1}, \ldots, e_{\pi,i,d_\pi}$ be bases of $H_{\pi,i}$ which are orthonormal with respect to some Γ-invariant inner product and symmetry adapted in the sense that any two bases $\{e_{\pi,i,j}\}_j$ and $\{e_{\pi,i',j}\}_j$ transform in the same way under the action of Γ. Then we can construct positive, Γ-invariant kernels

$$(x, y) \mapsto \sum_{i,i'} A_{i,i'} \sum_{j=1}^{d_\pi} \overline{e_{\pi,i,j}(x)} e_{\pi,i',j}(y), \quad A \succeq 0.$$

If the action of Γ on X is transitive or has finitely many orbits, then each positive, invariant kernel is an absolutely uniformly converging series of such kernels ([4] and [5, Theorem 3.4.4]). Otherwise the finite sums of such kernels are uniformly dense in the cone of positive, invariant kernels [5, Theorem 3.4.5].

For example, under the action of $O(n-1)$ we get the decomposition

$$\mathcal{C}(S^{n-1}) = \bigoplus_{k \geq 0} \bigoplus_{i \geq 0} \langle (x \cdot e)^i \rangle \otimes \langle e_{k,j}^{n-1}\,(x - (x \cdot e)e) : j = 1, \ldots, d_\pi \rangle.$$

Here the irreducible subspaces for fixed k are all equivalent. For positive semidefinite A this then leads to the Bachoc-Vallentin kernels

$$(x, y) \mapsto \sum_{i,i'} A_{i,i'} u^i v^{i'} (1-u^2)^{j/2} (1-v^2)^{j/2} P_k^{n-1} \left(\frac{t - uv}{\sqrt{1-u^2}\sqrt{1-v^2}} \right),$$

which were used in the three-point semidefinite programming bound for the spherical code problem [2].

In [3] Bachoc and Vallentin show this three-point bound, which can be solved using semidefinite programming, is sharp for the problem of computing the largest size of a subset of the sphere for which the inner product between each distinct pair of points is at most $(2\sqrt{2} - 1)/7$. In [7] a method is given for rounding the numerical output of a semidefinite programming solver to exact numbers, which is nontrivial in the common case where the dimension of the optimal face is smaller

than the dimension of the affine space defining the semidefinite program. For the above example this can be used to round the output to the quadratic number field $\mathbb{Q}[\sqrt{2}]$, showing the three-point bound is sharp in this case.

In the second part of this tutorial we consider invariant polynomial inequality constraints. Suppose we have a constraint of the form

$$p(x) \geq 0 \quad \text{for all} \quad x \in \mathbb{R}^n \quad \text{with} \quad g_1(x) \geq 0, \ldots, g_m(x) \geq 0,$$

where p is a polynomial whose coefficients depend linearly on other variables in the optimization problem. Then Putinar's representation [10] can be used to relax the problem by instead requiring

$$p(x) = s_0(x) + \sum_{i=1}^{m} s_i(x) g_i(x), \tag{1}$$

where $s_i(x)$ is a sum-of-squares polynomial of degree $d - \lfloor \deg(g_i) \rfloor$. By writing $s_i(x) = [x]^\mathsf{T} A_i [x]$, where $[x]$ is a vector whose entries form a basis for the polynomials of degree at most d, and equating coefficients in a common basis these become semidefinite constraints.

If the polynomials $p, g_1, \ldots, g_m$ are invariant under the action of a group, then we may assume the sum-of-squares polynomials $s_1, \ldots, s_m$ are invariant as well, and the techniques by Gatermann and Parrilo [8] can be used to find block-diagonalized representations of the form

$$s_i(x) = \sum_{\pi} \langle A_\pi, E_\pi(x) \rangle,$$

where the matrices A_π are positive semidefinite. In [9], for example, the invariance of the polynomials in the three-point bound for the action of the symmetric group S_3 is used to simplify the semidefinite programming formulation.

Löfberg and Parrilo suggest to model the identity (1) by sampling at a unisolvent set of points, as this leads to rank one constraint matrices in the resulting semidefinite program, which can be exploited to speed up interior point solvers. In [6] a new solver is developed that takes advantage of these low rank constraints. This solver works particularly well for problems such as the three-point bound discussed above even when symmetry reduction is also used, which increases the ranks of the constraint matrices.

References

[1] C. Bachoc, G. Nebe, F. M. de Oliveira Filho, and F. Vallentin, *Lower bounds for measurable chromatic numbers*, Geom. Funct. Anal. **19** (2009), no. 3, 645–661. MR 2563765

[2] C. Bachoc and F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. **21** (2008), no. 3, 909–924. MR 2393433

[3] C. Bachoc and F. Vallentin, *Optimality and uniqueness of the* $(4, 10, 1/6)$ *spherical code*, J. Combin. Theory Ser. A **116** (2009), no. 1, 195–204. MR 2469257

[4] S. Bochner, *Hilbert distances and positive definite functions*, Ann. of Math. (2) **42** (1941), 647–656. MR 0005782

[5] D. de Laat, *Moment methods in extremal geometry*, Ph.D. thesis, Delft University of Technology, 2017.

[6] D. de Laat and N. Leijenhorst, *Solving clustered low-rank semidefinite programs arising from polynomial optimization*, 2022.
[7] M. Dostert, D. de Laat, and P. Moustrou, *Exact Semidefinite Programming Bounds for Packing Problems*, SIAM Journal on Optimization **31** (2021), no. 2, 1433–1458, Publisher: Society for Industrial and Applied Mathematics.
[8] K. Gatermann and P. A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Algebra **192** (2004), no. 1-3, 95–128. MR 2067190
[9] F. C. Machado and F. M. de Oliveira Filho, *Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry*, *Aceito em* Experiment. Math. (2017), arXiv:1609.05167.
[10] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J. **42** (1993), no. 3, 969–984. MR 1254128
[11] I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108. MR 0005922

Exact algorithm for symmetric polynomial optimization problems and applications

Mohab Safey El Din
(joint work with Jean-Charles Faugère, George Labahn, Éric Schost, Thi Xuan Vu)

Computer algebra algorithms allow one to solve polynomial systems exactly by parametrizing the solution set with the set of roots of some univariate polynomial. Precisely, given a system of polynomial equations $f_1 = \cdots = f_s = 0$ in $\mathbb{Q}[x_1, \ldots, x_n]$ with finitely many complex solutions, one can compute a sequence of polynomials $w, v_1, \ldots, v_n$ in $\mathbb{Q}[t]$ where t is a new variable and w is monic and square-free such that the solution set is defined by

$$\{((v_1(\vartheta), \ldots, v_n(\vartheta))) \mid \vartheta \in \mathbb{C} \quad w(\vartheta) = 0\} \subset \mathbb{C}^n$$

This appealing output data structure may suffer from the curse of dimensionality as it has size which is linear in the number n of variables and in the number δ of points (this latter quantity may be exponential in n).

Applying this to polynomial optimization problems by computing critical points of the polynomial map to optimize still allows one to solve exactly polynomial optimization problems but yields similar issues of scalability. This can be overcome by improving the computational efficiency of fundamental algebraic algorithms and by exploiting any structural property of optimization problems.

In this talk, we study how one can bypass these scalability issues for the case of optimization problems which are invariant by the action of the symmetric group S_n. We consider the case where input polynomials $f_1, \ldots, f_s$ are invariant by the action of S_n and a polynomial map φ as well. In this setting, the set of critical points of the restriction of the map $z \mapsto \varphi(x)$ to the complex zero set defined by $f_1 = \cdots = f_s = 0$ is also invariant by the action of S_n. It is then relevant to compute one point per orbit to save the computational complexity of the symmetric group S_n. A classical strategy to do this is to rewrite a system

defining this set w.r.t. the symmetric elementary functions which are invariant by the action of S_n.

However, under some classical regularity assumptions, defining critical sets in the above context leads to the consider the maximal minors of the Jacobian matrix associated to $f_1, \ldots, f_s, \varphi$ to express that this matrix is rank deficient. Even if each of these polynomials is invariant by the action of S_n, these minors are not and then extra ingredients are needed.

A first observation is that since the polynomials $f_1, \ldots, f_s, \varphi$ are invariant by the action of S_n, they can be expressed as a composition of some polynomials with the elementary symmetric polynomials $e_1, \ldots, e_n$ as follows:

$$f_i = g_i(e_1, \ldots, e_n), \quad \text{for } 1 \leq i \leq s \quad \text{and } \varphi = \psi(e_1, \ldots, e_n)$$

Using the chain rule formula, this yields a natural factorisation of the Jacobian matrix associated to $f_1, \ldots, f_s, \varphi$ as the product of the Jacobian matrix associated to $g_1, \ldots, g_s, \psi$ with the one of $e_1, \ldots, e_n$. The latter one is a Vandermonde matrix since it has defective rank if and only if $x_i = x_j$ for $i \neq j$.

Hence, computing one point per orbot of the critical points set with distinct coordinates boil down to solving a system written w.r.t. $e_1, \ldots, e_n$ which enjoys a weighted degree (and then some sparsity) structure since the e_i's model a relation of degree i in the variables $x_1, \ldots, x_n$.

Computing one point per orbit in the subset of critical points which have equalled coordinates is done similarly, considering partitions of integers and more involved group actions by cartesian products of symmetric groups. This also yields determinantal polynomial systems expressed by elementary symmetric functions and then enjoying a special sparsity structure.

We then design a dedicated symbolic homotopy algorithm for solving such polynomial systems that exploits both the determinantal and sparsity structure. All in all, this leads to an algorithm whose complexity is polynomial in $E, D^s, \binom{n}{s+1}$ and $\binom{n+D}{n}$ where E is the evaluation complexity of $f_1, \ldots, f_s, \varphi$ and D the maximum degree of these polynomials. Note that when s and D are fixed, the algorithm runs in time which is polynomial in n. This constitutes an algebraic proof of the so-called degree principle which states that S_n invariant polynomial optimization problems can be solved in polynomial time for D fixed.

Note also that when s is fixed and $D \leq n^\alpha$ with $\alpha < 1$, the algorithm runs in time which is subexponential in n, hence extending the degree principle. Practical experiments show the interest of such approaches and algorithms.

Symmetry reduction in AM/GM-based optimization

Philippe Moustrou

(joint work with H. Naumann, C. Riener, T. Theobald, H. Verdure)

Deciding whether a real function only takes non-negative values is a fundamental question in real algebraic geometry. Non-negativity certificates and optimization approaches are tightly related to each other by observing that the infimum f^* of

a function $f : \mathbb{R}^n \to \mathbb{R}$ can be expressed as the largest $\lambda \in \mathbb{R}$ for which $f - \lambda$ is non-negative on $\mathbb{R}^n$:

$$f^* = \inf\{f(x) \ : \ x \in \mathbb{R}^n\} \ = \ \sup\{\lambda \in \mathbb{R} \ : \ f - \lambda \text{ is non-negative on } \mathbb{R}^n\}.$$

Both in the context of polynomials and in the broader context of exponential sums, the last years have seen strong interest in non-negativity certificates and optimization techniques based on the arithmetic mean/geometric mean-inequality (AM/GM inequality). More precisely, an exponential sum (or *signomial*) supported on a finite subset $\mathcal{T} \subset \mathbb{R}^n$ is a linear combination $\sum_{\alpha\in\mathcal{T}} c_\alpha \exp(\langle\alpha, x\rangle)$ with real coefficients c_α. In particular cases, the non-negativity of the real function defined by an exponential sum can be decided via the arithmetic-geometric mean inequality. For example, for support points $\alpha_0, \ldots, \alpha_m \in \mathbb{R}^n$ and coefficients $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^n_+$ satisfying $\sum_{i=1}^m \lambda_i = 1$ and $\sum_{i=1}^m \lambda_i\alpha_i = \alpha_0$, the exponential sum

$$\sum_{i=1}^m \lambda_i \exp(\langle\alpha_i, x\rangle) - \exp(\langle\alpha_0, x\rangle)$$

is non-negative on $\mathbb{R}^n$ as a consequence of the weighted arithmetic-geometric mean inequality, namely $\sum_{i=1}^m \lambda_i \exp(\langle\alpha_i, x\rangle) \geq \prod_{i=1}^m (\exp(\langle\alpha_i, x\rangle))^{\lambda_i}$. Clearly, sums of such exponential sums are non-negative as well. Note that exponential sums can be seen as a generalization of polynomials: when $\mathcal{T} \subset \mathbb{N}^n$, the transformation $x_i = \ln y_i$ gives polynomial functions $y \mapsto \sum_{\alpha\in\mathcal{T}} c_\alpha y^\alpha$ on $\mathbb{R}^n_{>0}$.

These AM/GM-based certificates appear to be particularly useful in sparse settings. In the specialized situation of polynomials, they can be seen as an alternative to non-negativity certificates based on sums of squares. The ideas of these approaches go back to Reznick [11] and have been recently brought back into the focus of the developments by Pantea, Koeppl, and Craciun [10], Chandrasekaran and Shah [1] ("*SAGE*" cone: *sums of arithmetic-geometric exponentials*) and Iliman and de Wolff [4] ("*SONC*" cone: *sums of non-negative circuit polynomials*), see also [6] for a generalized, uniform framework. The AM/GM certificates can be effectively obtained by relative entropy programming (see [1, 2]), and in restricted settings these relative entropy programs become geometric programs [5]. Other recent approaches to sparse polynomials besides the ones based on the AM/GM inequality can be found in the sparse moment hierarchies [14, 15] and in the works exploiting correlative sparsity [7, 13].

From an algebraic point of view, a problem is *symmetric* when it is invariant under some group action. Symmetries are ubiquitous in the context of polynomials and optimization, since they manifest both in the problem formulation and the solution set. This often allows to reduce the complexity of the corresponding algorithmic questions: Symmetry reduction has provided essential advances in many situations, especially in the context of sums of squares (see [3, 12]).

This work deals with the question to which extent symmetries can be exploited in AM/GM-based optimization assuming that the problem affords symmetries. We provide a first systematic study of the AM/GM-based approaches in G-invariant

situations under the action of a group G. Our focus is on symmetry-adapted representation theorems, and algorithmic symmetry reduction techniques.

First, we prove a symmetry-adapted decomposition theorem and develop a symmetry-adapted relative entropy formulation of the cone of SAGE exponentials in a general G-invariant setting. This adaption reduces the size of the resulting relative entropy programs or geometric programs, and the gain depends on the orbit structure of the group action.

In the case of the symmetric group, we use combinatorial aspects of the representation theory of the symmetric group in order to measure the size of the resulting relative entropy program. In particular, we identify situations in which the size of the symmetry adapted relative entropy program stabilizes with respect to the number of variables.

These structural results can be evaluated in terms of computations. In situations with strong symmetry structure, the number of variables and the number of equations and inequalities becomes substantially smaller. Accordingly, the interior-point solvers underlying the computation of SAGE bounds then show strong reductions of computation time. In various cases, the symmetry-adapted computation succeeds when the conventional SAGE computation fails.

References

[1] V. Chandrasekaran and P. Shah. Relative entropy relaxations for signomial optimization. *SIAM J. Optim.*, 26(2):1147–1173, 2016.

[2] V. Chandrasekaran and P. Shah. Relative entropy optimization and its applications. *Math. Program., Ser. A*, 161(1-2):1–32, 2017.

[3] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *J. Pure & Applied Algebra*, 192(1-3):95–128, 2004.

[4] S. Iliman and T. de Wolff. Amoebas, nonnegative polynomials and sums of squares supported on circuits. *Res. Math. Sci.*, 3(paper no. 9), 2016.

[5] S. Iliman and T. de Wolff. Lower bounds for polynomials with simplex Newton polytopes based on geometric programming. *SIAM J. Optim.*, 26(2):1128–1146, 2016.

[6] L. Katthän, H. Naumann, and T. Theobald. A unified framework of SAGE and SONC polynomials and its duality theory. *Math. Comput.*, 90:1297–1322, 2021.

[7] J. B. Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity. *SIAM J. Optim.*, 17:822–843, 2006

[8] R. Murray, V. Chandrasekaran, and A. Wierman. Newton polytopes and relative entropy optimization. *Found. Comput. Math.*, 21:1703–1737, 2021.

[9] R. Murray, V. Chandrasekaran, and A. Wierman. Signomial and polynomial optimization via relative entropy and partial dualization. *Math. Program. Comput.*, 13:257–295, 2021.

[10] C. Pantea, H. Koeppl, and G. Craciun. Global injectivity and multiple equilibria in uni- and bi-molecular reaction networks. *Discrete and Continuous Dynamical Systems - Series B*, 17(6):2153–2170, May 2012.

[11] B. Reznick. Forms derived from the arithmetic-geometric inequality. *Math. Annalen*, 283(3):431–464, 1989.

[12] C. Riener, T. Theobald, L. Jansson-Andrén, and J. B. Lasserre. Exploiting symmetries in SDP-relaxations for polynomial optimization. *Math. Oper. Res.*, 38(1):122–141, 2013.

[13] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM J. Optim.*, 17:218–242, 2006

[14] J. Wang, V. Magron, and J. B. Lasserre. Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension. *SIAM J. Optim.*, 31:114–141, 2021.
[15] J. Wang, V. Magron, and J. B. Lasserre. TSSOS: a moment-SOS hierarchy that exploits term sparsity. *SIAM J. Optim.*, 31:1–29, 2021.

New results in exact computations with group representations

Dmitrii Pasechnik

Real representations of finite groups play an important role in applications of convex optimisation to extremal combinatorics, quantum physics, etc. Often they arise while exploiting symmetry in dimension reduction of data for convex optimisation, see e.g. [6, 1]. A typical task is decomposing a d-dimensional linear represenation $\tau : G \to GL_d(\mathbb{R})$ of a finite group G into real irreducible representations. Such computations are very sensitive to rounding errors, and therefore are best done exactly. For such a task, explicit knowledge of real irreducible representations is crucial.

It is classically known that every complex representation $G \to GL_d(\mathbb{C})$ may be realised over the field F_n of n-th order cyclotomic numbers, $F_n := \mathbb{Q}(e^{\frac{2\pi i}{n}})$, where $n = \exp(G)$, the exponent of G. For real representations, there was empirical evidence that it suffices to take real cyclotomics of n-th order, i.e. elements of $E_n := F_n \cap \mathbb{R}$. The speaker succeeded to prove this in [4]. Specifically, the following holds.

Theorem. Let G be a finite group of exponent n, and ρ a G-representation $\rho : G \to GL_d(\mathbb{R})$. Then ρ may be realised over E_n.

To sketch the proof, consider the decomposition of the character χ_ρ into real irreducible characters χ_μ. They are of three types:

- *complex*, i.e. of the form $\chi_\tau + \chi_{\overline{\tau}}$ with τ a complex representation
- *quaternionic*, i.e. of the form $2\chi_\tau$, with τ a quaternionic representation;
- *real*, i.e. not of the above two types.

In the first two cases it is easy to show that μ may be written over E_n, by looking at the underlying half-dimensional representation τ as $G \ni g \mapsto A_g + iB_g$, with A_g and B_g real, and then forming μ by setting $g \mapsto \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$.

In the latter case, we start with $\tau : G \to GL_d(F_n)$ with $\chi_\tau = \chi_\mu$, i.e. τ is not necessary real, although given by cyclotomics; as τ and μ have the same character, they are equivalent, i.e. there exists Q such that $\mu = Q^{-1}\tau Q = \overline{Q^{-1}\tau Q}$, thus $\overline{Q}Q^{-1}\tau = \overline{\tau}\overline{Q}Q^{-1}$, and $P := \overline{Q}Q^{-1}$ transforms τ to $\overline{\tau}$, i.e. $P^{-1}\overline{\tau}P = \tau$. Then we show that $P = \Sigma^{-1}M$, with Σ a Hermitean G-invariant form, and M a symmetric bilinear G-invariant form; in particular, $P \in GL_d(F_n)$. Now we have

$$PQ = \overline{Q}, \qquad \det Q \neq 0 \tag{1}$$

implying $\overline{P}PQ = \overline{PQ} = Q$, i.e. $\overline{P}P = I$. The latter is an extra restriction: our construction only guarantees that $\overline{P}P = aI$ for some $a = \sum_{k=1}^{d} b_k\overline{b_k}$, $b_k \in F_n$. Using a result by Serre on induced characters [5], we show that in fact there exists

P so that $a = 1$; this means that in practice always $a = b\overline{b}$ for $b \in F_n$, and so we can replace P with P/b, if needed. Finally, for a randomly chosen $Y \in M_d(F)$ setting $Q = \overline{Y} + \overline{P}Y$ produces a solution to (1) with high probability.

Note that the proof is constructive; the algorithm fits well with the usual computational group theory systems such as GAP [2], which work with cyclotomic numbers very efficiently. We plan to extend our GAP package [3] to incorporate it.

Another interesting research direction is to produce a number-theoretic proof of this result, avoiding [5], and directly showing that $a = \sum_{k=1}^{d} b_k\overline{b_k}$ can be written as $a = b\overline{b}$, $b \in F_n$.

References

[1] Christine Bachoc, Dion C. Gijswijt, Alexander Schrijver, and Frank Vallentin, *Invariant semidefinite programs*,Handbook on semidefinite, conic and polynomial optimization, Internat. Ser. Oper. Res. Management Sci., vol. 166, Springer, New York, 2012, pp. 219–269.

[2] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.11.1*, 2021

[3] Kaashif Hymabaccus and Dmitrii V. Pasechnik, *RepnDecomp: A GAP package for decomposing linear representations of finite groups*, Journal of Open Source Software **5** (2020), no. 50, 1835–1836.

[4] Dmitrii Pasechnik, *Splitting fields of real irreducible representations of finite groups*, Represent. Theory **25** (2021), 897–902. MR 4324950

[5] Jean-Pierre Serre, *Conducteurs d'Artin des caractères réels*, Invent. Math. **14** (1971), 173–183. MR 321908

[6] Frank Vallentin, *Symmetry in semidefinite programs*, Linear Algebra Appl. **430** (2009), no. 1, 360–369. MR 2460523

Towards a computer-assisted proof for a conjecture from graph theory

Elisabeth Gaar

(joint work with Daniel Krenn, Susan Margulies, Melanie Siebenhofer, Angelika Wiegele)

We consider an open conjecture from Vizing in 1968 [5], namely that the product of the domination numbers of two graphs G and H is always smaller or equal to the domination number of the product graph of G and H.

In this talk we present initial steps on the way to a computer-assisted proof (or counter example) for Vizing's conjecture and discuss recent results. The foundation of this work is done in Gaar, Krenn, Margulies, and Wiegele [1, 2]. In particular, we start by considering the graph class $\mathcal{G}$ of all graphs with $n_{\mathcal{G}}$ vertices and domination number $k_{\mathcal{G}}$, and the graph class $\mathcal{H}$ of all graphs with $n_{\mathcal{H}}$ vertices and domination number $k_{\mathcal{H}}$. Then we continue by building an algebraic model of the conjecture for each possible tuple $(n_{\mathcal{G}}, k_{\mathcal{G}}, n_{\mathcal{H}}, k_{\mathcal{H}})$ by creating the Vizing ideal. Then we translate Vizing's conjecture for each tuple into the question of whether a specific polynomial f_{viz} is nonnegative over the Vizing ideal. Then we do another reformulation to the question of whether this specific polynomial f_{viz} is a sum-of-squares polynomial on a certain level of the sum-of-squares-hierarchy.

Finally we use semidefinite programming (SDP) to answer these kind of questions. Unfortunately any (rational) solution of an SDP solver still has to be transformed into an exact (algebraic) sum-of-squares certificates, which is a challenging task and can for example be done by cleverly guessing the exact SDP solution.

We give insight in the recent methods that have been used for developing new sum-of-squares certificates for particular parameters, i.e. tuples. In particular, we present certificates for the graph classes $\mathcal{G}$ and $\mathcal{H}$ with the property that $n_{\mathcal{G}}$, $k_{\mathcal{G}}$, $n_{\mathcal{H}}$ and $k_{\mathcal{H}}$ satisfy $k_{\mathcal{G}} = \mathcal{G} - 1 \geq 1$ and $k_{\mathcal{H}} = n_{\mathcal{H}} - 1$ for $n_{\mathcal{H}} \in \{2, 3\}$ and the graph classes $\mathcal{G}$ and $\mathcal{H}$ with $k_{\mathcal{G}} = n_{\mathcal{G}}$ and $k_{\mathcal{H}} = n_{\mathcal{H}} - d$ for $d \leq 4$ which were obtained by Gaar, Krenn, Margulies, and Wiegele [1, 2].

Furthermore, we present the findings of the master thesis of Siebenhofer [4], which can also be found in the preprint Gaar and Siebenhofer [3]. The authors are able to derive the unique reduced Gröbner basis of the Vizing ideal in the case that $k_{\mathcal{G}} = k_{\mathcal{H}} = 1$. This makes it possible for them to obtain the minimum degree $(n_{\mathcal{G}} + n_{\mathcal{H}} - 1)/2$ of a sum-of-squares certificate for Vizing's conjecture. Moreover, we present a method to find certificates for $n_{\mathcal{G}} + n_{\mathcal{H}} - 1 = d$ for general d, where only a much smaller SDP has to be solved compared to the method of Gaar, Krenn, Margulies, and Wiegele [1, 2]. Finally, certificates for all graph classes $\mathcal{G}$ and $\mathcal{H}$ with $k_{\mathcal{G}} = k_{\mathcal{H}} = 1$ and $n_{\mathcal{G}} + n_{\mathcal{H}} \leq 15$ are presented.

Finally, it should be mentioned that for all specifically considered graph classes it is already known in the literature that the inequality in Vizing's conjecture is true. As a result, we do not gain new knowledge on whether Vizing's conjecture is true or not for any graph classes. However, obtaining sum-of-squares certificates via an algebraic method is an important step in the area of using conic linear optimization for computer-assisted proofs, because it shows that deriving such proofs is possible for a set of graph classes.

References

[1] E. Gaar, D. Krenn, S. Margulies, and A. Wiegele, *An Optimization-Based Sum-of-Squares Approach to Vizing's Conjecture*, Proceedings of the International Symposium on Symbolic and Algebraic Computation (ISSAC '19), July 15–18, 2019, Beijing, China.

[2] E. Gaar, D. Krenn, S. Margulies, and A. Wiegele. *Towards a computational proof of vizing's conjecture using semidefinite programming and sums-of-squares*, J. Symb. Comput., 107:67–105, 2021.

[3] E. Gaar and M. Siebenhofer. *Sum-of-squares certificates for Vizing's conjecture via determining Gröbner bases*, https://arxiv.org/abs/2112.04007, December 2021.

[4] M. Siebenhofer. *Establishing new sum-of-squares certificates for vizing's conjecture*, Master's thesis, University of Klagenfurt, 2021.

[5] V. G. Vizing. *Some unsolved problems in graph theory*, Uspehi Mat. Nauk, 23(6 (144)):117–134, 1968.

The Chvátal-Gomory Procedure for Integer Semidefinite Programs

Renata Sotirov
(joint work with Frank de Meijer)

The Chvátal-Gomory (CG) cutting-plane procedure is introduced by Chvátal [1] and Gomory [2] and it is considered to be among the most celebrated results in integer programming. We consider the Chvátal-Gomory (CG) procedure for integer semidefinite programs (ISDPs). We present a formulation of the elementary closure of spectrahedra that relies on the data matrices of the ISDP and positive semidefinite matrices. This formulation provides a constructive description of the elementary closure of spectrahedra rather than the implicit description that is known for general convex sets. Equivalent to the case of polyhedra, the elementary closure operation can be repeated, leading to a hierarchy of stronger approximations of the integer hull of the spectrahedron. Our explicit formulation of the elementary closure enables us to introduce Chvátal-Gomory cuts and strengthened Chvátal-Gomory cuts for ISDPs. We also show how to derive a polyhedral description of the CG closure for a specific class of spectrahedra based on the SDP equivalent of total dual integrality.

In the second part of the talk, we show how to exploit (strengthened) CG cuts in a branch-and-cut framework for ISDPs. Different from existing algorithms for solving ISDPs, the separation routine in our approach exploits both the semidefinite and the integrality constraints. We present separation routines for common classes of binary SDPs resulting from combinatorial optimization problems.

In the third part of the talk, we apply our approach to the quadratic traveling salesman problem (QTSP). The QTSP is the problem of finding a Hamiltonian cycle in a graph that minimizes the total interaction costs among consecutive arcs. By exploiting the algebraic connectivity of the directed Hamiltonian cycle, we introduce an ISDP model for the QTSP. We show that the CG cuts resulting from these formulations contain several well-known families of cutting planes. Numerical results illustrate the practical strength of the CG cuts in our branch-and-cut algorithm, which outperforms alternative ISDP solvers and is able to solve large QTSP instances to optimality. For details on the CG-cuts for ISDPs see [3].

References

[1] V. Chvátal, *Edmonds polytopes and a hierarchy of combinatorial problems*, Discrete Math. **4** (1973), 305–337.

[2] R.E. Gomory, *Outline of an algorithm for integer solutions to linear programs*, Bull. Am. Math. Soc. **64** (1958), 275–278.

[3] Frank de Meijer, Renata Sotirov, The Chvátal-Gomory Procedure for Integer SDPs with Applications in Combinatorial Optimization, Preprint, 2022, https://arxiv.org/abs/2201.10224

High-Dimensional Hoffman Bound and its applications to extremal combinatorics

Konstantin Golubev
(joint work with Yuval Filmus and Noam Lifshitz)

The celebrated Hoffman bound [10] connects spectral graph theory with extremal combinatorics, by upper-bounding the independence number of a graph in terms of the minimal eigenvalue of its adjacency matrix. Hoffman's original paper only states a bound on the chromatic number. The corresponding bound on the independence number appears in works of Haemers [9] and Lovász [11], where it is attributed to Hoffman. The weighted version of the Hoffman bound, used by Lovász [11] to compute the Shannon capacity of the pentagon, had appeared earlier in foundational work of Delsarte [2] in coding theory. The Hoffman bound, in a generalized version due to Lovász [11], has seen many applications in extremal set theory and theoretical computer science.

Let $G = (V, \mu)$ be a graph, that is, V is the set of vertices and μ is a probability measure on the set of unordered pairs $V^{[2]}$ with repetitions allowed. By edges of G we mean the support of μ in $V^{[2]}$. A subset of V is called *independent* if it does not contain any edges. By μ_1 we denote the induced probability measure on V, that is, $\mu_1(v) = \mu(\{v, v\}) + \frac{1}{2}\sum_{u \neq v} \mu(\{u, v\})$. The *normalized adjacency operator* of G can defined by the matrix T_G given by

$$T_X(u, v) = \begin{cases} \frac{\mu([u,u])}{\mu_1(u)} & \text{if } u = v; \\ \frac{\mu([u,v])}{2\mu_1(u)} & \text{if } u \neq v. \end{cases}$$

The largest eigenvalue of T_G is 1, by $\lambda(G)$ we denote its minimal eigenvalue. In this terms, the Hoffman bound reads as follows

$$\alpha(G) \leq 1 - \frac{1}{1 - \lambda(G)}.$$

The Hoffman bound can be used to solve problems in extremal set theory in which the constraints can be modeled as a graph. As an example, the Hoffman bound can be used to prove the fundamental Erdős–Ko–Rado theorem on the size of intersecting families, in which the constraint is that every two sets in the family have nonempty intersection, as well as many other Erdős–Ko–Rado theorems on various domains [3, 4, 7]. Other problems involve more complex constraints, and so are not amenable to this method. A simple example is the s-wise intersecting Erdős–Ko–Rado theorem, due to Frankl [5], which concerns families in which every s sets have nonempty intersection. In this case the constraints can be modeled as a *hypergraph* rather than as a graph.

Recently, the Hoffman bound has been generalized to hypergraphs [1, 8]. Our new bound is particularly attractive for upper-bounding independent sets in tensor powers of hypergraphs. We demonstrate the power of this method by solving a problem of Frankl on triangle-free families and by giving a spectral proof of Mantel's theorem.

In order to formulate the bound, we set the notations first. This will take about a page, more details can be found in our paper, [6]. A *multiset* is an unordered collection of elements that is allowed to have repetitions, its size is the number of its elements counting the multiplicity. An *i-multiset* is a multiset of size i. Let V be a set. We denote by $V^{[i]}$ the collection of all i-multisets of elements of V, and elements of $V^{[i]}$ will be denoted by $[v_1, \ldots, v_i]$. The collection $V^{[0]}$ consists of the empty set. A *weighted k-uniform hypergraph* is a pair $X = (V, \mu)$ where V is the vertex set and μ is a probability distribution on $V^{[k]}$. For $0 \leq i \leq k-1$, define a probability measure μ_i on $V^{[i]}$ by the following process. First, choose a multiset $[v_1, \ldots, v_k]$ according to μ, and then choose an i-submultiset of it uniformly at random. We write $X^{(i)}$ for the set of elements of $V^{[i]}$ whose μ_i measure is positive. The elements of $X^{(i)}$ are called the *i-faces* of X, and the elements of $X^{(0)} \cup \cdots \cup X^{(k)}$ are called the *faces* of X. Note that if σ_2 is a face of X, then σ_1 is a face of X for any $\sigma_1 \subseteq \sigma_2$. Note that $X^{(0)} = \{\emptyset\}$, i.e., the empty set is the one and only 0-face of X. We assume without loss of generality that $X^{(1)} = V$, that is, $\mu_1(v) > 0$ for all $v \in V$ (otherwise, we can replace V with $X^{(1)}$).

A set $I \subseteq V$ is said to be *independent in a k-uniform hypergraph* X if no k-face of X is contained in I. The largest possible value of $\mu_1(I)$, where $I \subseteq V$ is an independent set in X, is called the *independence number* of X and denoted $\alpha(X)$. A subset $I \subseteq V$ is said to be an *extremal independent set* of X if $\mu_1(I) = \alpha(X)$.

For an i-face $\sigma \in X^{(i)}$, its *link* in X is the $(k-i)$-uniform hypergraph $X_\sigma = (V, \mu_\sigma)$, where μ_σ is the probability distribution that corresponds to the following process: sample a random flag according to $\boldsymbol{\mu}$ subject to $\boldsymbol{\mu}_i = \sigma$, and output $\boldsymbol{\mu}_k \setminus \sigma$. Note that the link of the empty set is the whole hypergraph X itself. For an element $v \in X$, the link of v is $X_v = X_{\{v\}}$.

The *skeleton* of X is the graph $(X^{(1)}, \mu_2)$, in other words, its vertex set is $X^{(1)}$, the edges are $X^{(2)}$, with the weights given by μ_2.

By $\lambda(X)$ we denote the smallest eigenvalue of T_X. For all $0 \leq i \leq k-2$, we write

$$\lambda_i(X) = \min\left\{\lambda(S(X_\sigma)) \,\middle|\, \sigma \in X^{(i)}\right\}.$$

In other words, $\lambda_i(X)$ is the minimal possible value of an eigenvalue of the normalized adjacency matrix of a skeleton of the link of an i-face of X. Note that $\lambda_0(X)$ is just the smallest eigenvalue of the normalized adjacency operator on the skeleton of X.

We conclude the notations with the notion of the tensor product $X \otimes X'$ of two k-uniform hypergraphs $X = (V, \mu)$ and $X' = (V', \mu')$. It is a k-uniform hypergraph $(V \times V', \mu \times \mu')$, where $\mu \times \mu'$ stands for the following measure on $(V \times V')^{[k]} \simeq V^{[k]} \times V'^{[k]}$: an ordering of an edge $(\sigma, \sigma') \in V^{[k]} \times V'^{[k]}$ implies an ordering on σ and on σ', we define $(\mu \times \mu')_k(\sigma, \sigma')$ as the sum of $\widetilde{\mu_k}(\sigma)\widetilde{\mu'_k}(\sigma')$ over all orderings of (σ, σ'). The measure $\widetilde{\mu_k}$ is constructed from μ by choosing an ordering on an edge uniformly at random. For a k-uniform hypergraph X, we denote by $X^{\otimes n} = \underbrace{X \otimes \cdots \otimes X}_{n}$ its n-th tensor power.

We prove a new upper bound for the independence number of a hypergraph, and its invariance under the tensor power operation.

$$\alpha(X) \leq 1 - \frac{1}{(1-\lambda_0)(1-\lambda_1)\cdots(1-\lambda_{k-2})}.$$

If in addition $\lambda_i \leq 0$ for all $0 \leq i < k-2$, then for any positive integer n the following inequality holds for $X^{\otimes n}$:

$$\alpha(X^{\otimes n}) \leq 1 - \frac{1}{(1-\lambda_0)(1-\lambda_1)\cdots(1-\lambda_{k-2})}.$$

In particular, if the bound is sharp for X, it remains sharp for its tensor powers, as $\alpha(X^{\otimes n}) = \alpha(X)$.

Furthermore, if $\lambda_0 > -1$ and I is an independent set attaining the bound, then I is a dictator (viewed as a subset of V^n, membership in I depends on a single coordinate).

Refer to our paper,[6], for more details.

References

[1] Christine Bachoc, Anna Gundert, and Alberto Passuello, *The theta number of simplicial complexes*, Israel Journal of Mathematics **232** (2019), no. 1, 443–481.

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

[3] David Ellis, Yuval Filmus, and Ehud Friedgut, *Triangle-intersecting families of graphs*, Journal of the European Mathematical Society **14** (2012), 841–885.

[4] David Ellis, Ehud Friedgut, and Haran Pilpel, *Intersecting families of permutations*, Journal of the American Mathematical Society **24** (2011), no. 3, 649–682.

[5] Péter Frankl, *On Sperner families satisfying an additional condition*, Journal of Combinatorial Theory, Series A **20** (1976), no. 1, 1–11.

[6] Yuval Filmus, Konstantin Golubev, Noam Lifshitz, *High dimensional Hoffman bound and applications in extremal combinatorics*, Algebraic Combinatorics **4, no. 6** (2021): 1005–1026.

[7] Chris Godsil and Karen Meagher, *Erdős–Ko–Rado theorems: algebraic approaches*, Cambridge Studies in Advanced Mathematics, vol. 149, Cambridge University Press, Cambridge, 2016.

[8] Konstantin Golubev, *On the chromatic number of a simplicial complex*, Combinatorica (2016), 1–12.

[9] Willem Haemers, *A generalization of the Higman–Sims technique*, Nederl. Akad. Wetensch. Indag. Math. **40** (1978), no. 4, 445–447.

[10] Alan J. Hoffman, *On eigenvalues and colorings of graphs*, in Selected Papers of Alan J. Hoffman — with Commentary, World Scientific, 2003, pp. 407–419.

[11] László Lovász, *On the Shannon capacity of a graph*, IEEE Transactions on Information theory **25** (1979), no. 1, 1–7.

Razborov's Flag Algebras

Fernando Mário de Oliveira Filho

The theory of flag algebras, introduced by Razborov in 2007 [2], has opened the way to a systematic approach to the development of computer-assisted proofs in extremal combinatorics. It makes it possible to derive bounds for parameters in extremal combinatorics with the help of a computer, in a semi-automated manner. In this talk, based on an expository paper [1] written jointly with M.K. de Carli Silva and C.M. Sato, I describe the main points of the theory in a complete way, using Mantel's theorem as a guiding example.

Mantel's theorem is perhaps the first result in extremal graph theory. It states that the maximum number of edges in an n-vertex triangle-free graph is $\lfloor n^2/4 \rfloor$, the maximum being achieved by a complete bipartite graph with parts of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$. This is an example of the prototypical question in extremal graph theory: given a collection $\mathcal{H}$ of forbidden graphs, how many copies of a graph C can an $\mathcal{H}$-free graph G have?

Let us be more precise. Denote by $|G|$ the number of vertices of a graph G. Given graphs F and G, let $p(F;G)$ be the (induced) *density* of F in G, that is, if $c(F;G)$ is the number of times F occurs as an induced subgraph of G, then

$$p(F;G) = c(F;G)\binom{|G|}{|F|}^{-1}.$$

Let $\mathcal{H}$ be a collection of graphs. A graph is *$\mathcal{H}$-free* if no induced subgraph of G is isomorphic to a graph in $\mathcal{H}$. A fundamental problem in extremal graph theory is to determine, for a given graph C, the maximum asymptotic density of C in $\mathcal{H}$-free graphs, namely

$$\operatorname{ex}(C,\mathcal{H}) = \sup_{(G_k)_{k\geq 0}} \limsup_{k\to\infty} p(C;G_k), \tag{1}$$

where the supremum is taken over all sequences $(G_k)_{k\geq 0}$ of $\mathcal{H}$-free graphs that are *increasing*, that is, $(|G_k|)_{k\geq 0}$ is strictly increasing. Mantel's theorem asserts that $\operatorname{ex}(e,\{T\}) = 1/2$, where e is the graph consisting of a single edge and T is a triangle.

Let $\mathcal{G}$ be the set of all finite $\mathcal{H}$-free graphs taken up to isomorphism. An increasing sequence $(G_k)_{k\geq 0}$ is *convergent* if $\lim_{k\to\infty} p(F;G_k)$ exists for every $F \in \mathcal{G}$. Every increasing sequence of $\mathcal{H}$-free graphs has a convergent subsequence, since densities are numbers in $[0,1]$ and so for $k \geq 0$ the function $F \mapsto p(F;G_k)$ can be identified with a point in $[0,1]^{\mathcal{G}}$, which is a compact space by Tychonoff's theorem.

In (1) we can therefore restrict ourselves to convergent sequences and then work with their limits. Call $\phi\colon \mathcal{G} \to \mathbb{R}$ a *limit functional* if there is a convergent sequence $(G_k)_{k\geq 0}$ of $\mathcal{H}$-free graphs such that

$$\phi(F) = \lim_{k\to\infty} p(F;G_k)$$

for all $F \in \mathcal{G}$. Let Φ be the set of all limit functionals. Computing $\mathrm{ex}(C, \mathcal{H})$ is the same as solving an optimization problem over Φ, namely

$$\mathrm{ex}(C, \mathcal{H}) = \sup\{\, \phi(C) : \phi \in \Phi \,\}.$$

This rewording of the original problem makes it clear that the issue is understanding Φ; Razborov's theory of flag algebras is really a study of the set Φ of limit functionals.

The theory is built around the concept of density defined above, for which the *chain rule* holds: if F and G are graphs, if n is an integer such that $|F| \leq n \leq |G|$, and if $\mathcal{G}_n$ is the set of all graphs on n vertices up to isomorphism, then

$$p(F; G) = \sum_{F' \in \mathcal{G}_n} p(F; F')p(F'; G).$$

Now consider the free vector space $\mathbb{R}\mathcal{G}$. Given a limit functional ϕ, we may extend it to a linear functional of $\mathbb{R}\mathcal{G}$, and using the chain rule we may verify that

$$\phi(F) = \phi\Bigg(\sum_{F' \in \mathcal{G}_n} p(F; F')F' \Bigg)$$

for any integer $n \geq |F|$. (Note: this holds for limit functionals, not for every linear functional on $\mathbb{R}\mathcal{G}$.)

This means that, to limit functionals (which are our objects of interest),

$$F \quad \text{and} \quad \sum_{F' \in \mathcal{G}_n} p(F; F')F'$$

are the same. Let $\mathcal{K}$ be the space spanned by all expressions of the form

$$F - \sum_{F' \in \mathcal{G}_n} p(F; F')F'$$

and consider the quotient space $\mathcal{A} = \mathbb{R}\mathcal{G}/\mathcal{K}$. Since $\mathcal{K}$ is a subset of the kernel of every limit functional, every limit functional is a linear functional of $\mathcal{A}$.

On $\mathcal{A}$ it is possible to define, again with the help of the chain rule, an associative and commutative product which turns $\mathcal{A}$ into an algebra with identity (given by the empty graph $\varnothing$), called the *flag algebra*. With respect to this product, limit functionals are multiplicative: if $f, g \in \mathcal{A}$, then $\phi(f \cdot g) = \phi(f)\phi(g)$. Moreover, $\phi(\varnothing) = 1$ for every limit functional ϕ, and so limit functionals are algebra homomorphisms from $\mathcal{A}$ to $\mathbb{R}$. Since they are defined in terms of densities, limit functionals are also nonnegative: for every graph F we have $\phi(F) \geq 0$. Razborov shows that Φ is exactly the set of nonnegative algebra homomorphisms from $\mathcal{A}$ to $\mathbb{R}$.

This gives an exact characterization of the set Φ. Building on this, the theory provides, among other things, a systematic way of computing hierarchies of convex relaxations of Φ, that is, sequences

$$[0, 1]^{\mathcal{G}} \supseteq \Phi_1 \supseteq \Phi_2 \supseteq \cdots$$

of convex sets Φ_k, each containing Φ. Moreover, with appropriate control, optimization over the relaxations Φ_k can be carried out efficiently, for instance via semidefinite programming. This allows us to compute upper bounds for $\mathrm{ex}(C, \mathcal{H})$.

Many significant results in extremal combinatorics have been obtained via the theory of flag algebras, for instance: computing the minimal number of triangles in graphs with given density, computing the maximum number of pentagons in triangle-free graphs, and obtaining new results regarding the Cacceta-Häggkvist conjecture (see Carli Silva, Oliveira, and Sato [1] for references). Though presented here in the context of graphs, the theory is more general, and can be applied also to digraphs, permutations, and other objects.

References

[1] M.K. de Carli Silva, F.M. de Oliveira Filho, and C.M. Sato, Flag algebras: a first glance, *Nieuw Archief voor Wiskunde* 5/17 (2016) 193–199.

[2] A. Razborov, Flag algebras, *Journal of Symbolic Logic* 72 (2007) 1239–1282.

A Recursive Theta Body for Hypergraphs

Fernando Mário de Oliveira Filho

(joint work with D. Castro-Silva, L. Slot, and F. Vallentin)

Let $H = (V, E)$ be an r-uniform hypergraph. A set $I \subseteq V$ is *independent* if no edge of H is contained in I. Given a weight function $w \in \mathbb{R}^V$, the *weighted independence number* of H is

$$\alpha(H, w) = \max\{\, w(I) : I \subseteq V \text{ is independent} \,\},$$

where $w(I) = \sum_{x \in I} w(x)$.

The *independent-set polytope* of H is

$$\mathrm{IND}(H) = \mathrm{conv}\{\, \chi_I \in \mathbb{R}^V : I \subseteq V \text{ is independent} \,\},$$

where $\chi_S \in \mathbb{R}^V$ is the characteristic function of $S \subseteq V$. The weighted independence number of H can be computed by maximizing $w^{\mathsf{T}} f$ over $f \in \mathrm{IND}(H)$, and so optimizing over the independent-set polytope is an NP-hard problem.

A *clique* of H is a set $C \subseteq V$ such that every r-subset of C is an edge. If C is a clique of H and $f \in \mathrm{IND}(H)$, then $f(C) \leq r - 1$. These valid inequalities for $\mathrm{IND}(H)$ are called *clique inequalities*; they give a relaxation of the independent set polytope, namely

$$\mathrm{QIND}(H) = \{\, f \in [0,1]^V : f(C) \leq r - 1 \text{ for every } C \subseteq V \,\}.$$

The integer hull of this relaxation is $\mathrm{IND}(H)$. The separation problem over $\mathrm{QIND}(H)$ is NP-hard, and hence optimizing over $\mathrm{QIND}(H)$ is also an NP-hard problem.

Grötschel, Lovász, and Schrijver [1] defined the theta body of a graph: a convex relaxation of $\mathrm{IND}(H)$ stronger than $\mathrm{QIND}(H)$ over which it is possible to optimize a linear function in polynomial time. We extend this definition recursively to hypergraphs with uniformity $r \geq 3$.

We take as our base case $r = 1$; though 1-uniform hypergraphs are a rather degenerate case, this allows us to give uniform proofs that recover results for graphs proved by Grötschel, Lovász, and Schrijver [1]. So we define the *theta body* of a 1-uniform hypergraph as $\mathrm{TH}(H) = \mathrm{IND}(H)$.

Let $H = (V, E)$ be an r-uniform hypergraph for some $r \geq 2$. Given $x \in V$, the *link* of x in H is the $(r-1)$-uniform hypergraph H_x with vertex set

$$V_x = \{\, y \in V : \text{there is } e \in E \text{ containing } x \text{ and } y \,\},$$

in which an $(r-1)$-subset e of V_x is an edge if $e \cup \{x\}$ is an edge of H.

Given a matrix $A \in \mathbb{R}^{V \times V}$ and $x \in V$, let $A_x \in \mathbb{R}^V$ denote the row of A indexed by x, that is, $A_x(y) = A(x, y)$. If $f \colon V \to \mathbb{R}$ is any function and $U \subseteq V$ is any set, denote by $f[U]$ the restriction of f to U. The theta body of the r-uniform hypergraph $H = (V, E)$ for $r \geq 2$ is the convex set

$$\begin{aligned}\mathrm{TH}(H) = \{\, f \in \mathbb{R}^V : \text{there is } F \in \mathbb{R}^{V \times V} \text{ such that } & f = \operatorname{diag} F, \\ & F_x[V_x] \in F(x,x)\,\mathrm{TH}(H_x) \quad \text{for every } x \in V, \text{ and} \\ & \begin{pmatrix} 1 & f^{\mathsf{T}} \\ f & F \end{pmatrix} \text{ is positive semidefinite} \,\}.\end{aligned}$$

It is easy to see that $\mathrm{IND}(H) \subseteq \mathrm{TH}(H)$. Indeed, if $I \subseteq V$ is any independent set, then for every $x \in I$ we have that $I \cap V_x$ is an independent set of the link H_x of x. This implies that $f = \chi_I$ and $F = \chi_I \chi_I^{\mathsf{T}}$ satisfy the constraints above, so $\chi_I \in \mathrm{TH}(H)$. It is also possible to prove that $\mathrm{TH}(H) \subseteq \mathrm{QIND}(H)$, and so the theta body is a convex relaxation of the independent-set polytope stronger than $\mathrm{QIND}(H)$.

For every fixed r it is possible to optimize over $\mathrm{TH}(H)$ in polynomial time. Many other properties of the theta body of a graph, such as antiblocking relations, can also be shown for this recursive extension.

By optimizing over the theta body we obtain upper bounds for the weighted independence number. These upper bounds can be used to prove results in extremal combinatorics. For instance, it is possible to reprove Mantel's theorem by considering the hypergraph whose vertex set is $V = \binom{[n]}{2}$, where $[n] = \{1, \ldots, n\}$, and in which three elements of V are adjacent if they form a triangle of the complete graph K_n. Another application is computing upper bounds for the maximum number of binary words of length n no three of which are mutually at Hamming distance d from each other for some fixed d.

Finally, Castro-Silva, Oliveira, Slot, and Vallentin [2] consider an extension of the theta body to infinite geometrical graphs which provides upper bounds for the maximum measure of sets of points on the sphere or in Euclidean space that avoid regular simplices of a given side length. These bounds are strong enough to show that the maximum measure of such sets decays exponentially fast in the dimension of the space.

References

[1] M. Grötschel, L. Lovász, and A. Schrijver, Relaxations of vertex packing, *Journal of Combinatorial Theory, Series B* 40 (1986) 330–343.

[2] D. Castro-Silva, F.M. de Oliveira Filho, L. Slot, and F. Vallentin, A recursive Lovász theta number for simplex-avoiding sets, to appear in *Proceedings of the American Mathematical Society*, 2021, 13pp, arXiv:2106.09360.

The Symmetries of Flag Algebras

Daniel Brosch

Flag algebras, first introduced by Razborov [1] in 2007, remain one of the most powerful tools in extremal combinatorics today. Recently, Raymond, Saunderson, Singh and Thomas [2] discovered a connection to polynomial optimization: we can recover flag sums-of-squares hierarchies by partially exploiting the symmetries of a sequence of symmetric polynomial optimization hierarchies, and taking the limit as the number of variables approaches infinity. We continue from there, fully exploiting the symmetries of two different hierarchies, one focusing on a low number of edges, and another focusing on a low number of vertices.

We here optimize over polynomials in binary variables x_{ij}, with indices $1 \leq i < j \leq n$, which correspond to edges in a graph on n vertices. Each of the polynomials in the problem is fully symmetric for the simultaneous action of the symmetric group on the indices:

$$\sigma(x_{ij}) = x_{\sigma(i)\sigma(j)}.$$

The edge-truncated hierarchy is exactly the Lasserre hierarchy for this problem, which relaxes the polynomial optimization problem to a convex problem in positive semidefinite variables. The convexity of the relaxation allows us to exploit the symmetries, and block-diagonalize the problem using Artin-Wedderburn theory. We can see the ring of polynomials, truncated to a given degree, as an S_n-module. To compute the block-diagonalization, we need to decompose this module into Specht modules, the irreducible modules of S_n. Due to the high dimension of the spans of orbits of monomials, we needed to come up with an efficient algorithm to decompose quotients of permutation modules M^λ/G, where G is a group of permutations acting on the rows of the tabloids in M^λ.

The resulting symmetry reduced hierarchy is equivalent to optimization over flag sums-of-squares, but our flags are not obtained by labeling certain vertices, but by grouping vertices together, and is computationally more efficient.

The second, vertex-truncated hierarchy is closer to what Razborov [1] defines. We generalize the hierarchy, reduce it further, and show how it can be obtained as a truncation of a high level of the Lasserre hierarchy. This allows us to compare the Razborov hierarchy to the Lasserre hierarchy in both directions.

We apply the reduced hierarchies to compute outer approximations of graph profiles, which are the sets of simultaneously achievable subgraph densities of a given set of graphs.

References

[1] A. Razborov, *Flag algebras*, Journal Of Symbolic Logic **72** (2007), 1239-1282.
[2] A. Raymond, J. Saunderson, M. Singh and R. Thomas, *Symmetric sums-of-squares over k-subset hypercubes*, Mathematical Programming **2** (2018), 315–354.

Worst-case analyses of simple iterative methods via semidefinite programming

Adrien Taylor

In this talk, we provide an introductory overview on how simple iterative methods can be analyzed via semidefinite programming. In short, the idea is that finding worst-case scenarios (a.k.a. "worst-case analyses") can be cast as solutions to certain optimization problems, referred to as "performance estimation problems" (PEPs). In the cases of interest for this talk, we show that those problems are actually instances of linear semidefinite programs (SDPs). As a result, by strong duality, any worst-case guarantee on the performance of the method at hand can be obtained as a feasible point to the dual SDP (to that of finding a worst-case instance). The PEP framework was originally proposed by Yoel Drori and Marc Teboulle in [1] for analyzing first-order optimization methods (namely gradient descent, its accelerated variants by Nesterov [2], and the heavy-ball method by Polyak [3]). In this talk, we follow a principled approach to PEPs by [4, 5], which shows that PEPs construct non-improvable worst-case guarantees for a large class of first-order optimization methods. We also mention the possibility of using this framework for searching for Lyapunov functions and its link with the integral quadratic constraints [6] for analyzing iterative optimization methods.

In the first part of the presentation, we show how to analyze an instance of gradient descent for smooth strongly convex minimization, using a few principled steps. Those stages finally lead us to an equivalence between performing this worst-case analysis and solving a linear semidefinite program. We further provide the key elements allowing to generalize the approach to the analysis of oracle-based first-order methods for composite convex optimization, including those performing explicit, projected, proximal, conditional and inexact (sub)gradient steps. For all those methods (and the corresponding problem classes), the PEP framework allows to simultaneously obtain tight worst-case guarantees and explicit instances of optimization problems on which the worst-cases are attained, by solving SDPs.

In the second part of the talk, we discuss two recent pieces of software [7, 8], allowing to perform such analyses (numerically) without going into the SDP modelling details (which are very error-prone). We also discuss a few recent (tight) analyses that were obtained using those tools: the Halpern-iteration for solving fixed-point problems [9], the proximal-point algorithm for solving monotone inclusions [10], gradient descent for nonconvex minimization [11, 12], as well as a few methods that were entirely obtained using PEPs [13, 14, 15, 16]. We conclude this part by discussing gradient descent with exact line-searches [17].

Finally, we illustrate that PEPs can be used for designing methods with optimized (and often optimal) worst-case guarantees, using results from [18, 19]. We finally draw a few vague links with Chebyshev methods, which are optimal for solving smooth strongly convex quadratic minimization problems [20].

References

[1] Yoel Drori and Marc Teboulle. Performance of first-order methods for smooth convex minimization: a novel approach. *Mathematical Programming*, 145(1):451–482, 2014.

[2] Yurii Nesterov. A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Soviet Mathematics Doklady*, 27(2):372–376, 1983.

[3] Boris T. Polyak. *Introduction to optimization*. Optimization Software New York, 1987.

[4] Adrien B. Taylor, Julien M. Hendrickx, and François Glineur. Smooth strongly convex interpolation and exact worst-case performance of first-order methods. *Mathematical Programming*, 161(1-2):307–345, 2017.

[5] Adrien B. Taylor, Julien M. Hendrickx, and François Glineur. Exact worst-case performance of first-order methods for composite convex optimization. *SIAM Journal on Optimization*, 27(3):1283–1313, 2017.

[6] Laurent Lessard, Benjamin Recht, and Andrew Packard. Analysis and design of optimization algorithms via integral quadratic constraints. *SIAM Journal on Optimization*, 26(1):57–95, 2016.

[7] Adrien B. Taylor, Julien M. Hendrickx, and François Glineur. Performance estimation toolbox (PESTO): automated worst-case analysis of first-order optimization methods. In *Proceedings of the 56th Conference on Decision and Control (CDC)*, 2017.

[8] Baptiste Goujaud, Céline Moucer, François Glineur, Julien Hendrickx, Adrien Taylor, and Aymeric Dieuleveut. PEPit: computer-assisted worst-case analyses of first-order optimization methods in python. *preprint arXiv:2201.04040*, 2022.

[9] Felix Lieder. On the convergence rate of the Halpern-iteration. *Optimization Letters*, 15(2):405–418, 2021.

[10] Guoyong Gu and Junfeng Yang. Tight sublinear convergence rate of the proximal point algorithm for maximal monotone inclusion problems. *SIAM Journal on Optimization*, 30(3):1905–1921, 2020.

[11] Hadi Abbaszadehpeivasti, Etienne de Klerk, and Moslem Zamani. The exact worst-case convergence rate of the gradient method with fixed step lengths for L-smooth functions. *Optimization Letters*, 2021.

[12] Teodor Rotaru, François Glineur, and Panagiotis Patrinos. Tight convergence rates of the gradient method on hypoconvex functions. *preprint arXiv:2203.00775*, 2022.

[13] Donghwan Kim and Jeffrey A. Fessler. Optimized first-order methods for smooth convex minimization. *Mathematical Programming*, 159(1-2):81–107, 2016.

[14] Bryan Van Scoy, Randy A. Freeman, and Kevin M. Lynch. The fastest known globally convergent first-order method for minimizing strongly convex functions. *IEEE Control Systems Letters*, 2(1):49–54, 2017.

[15] Donghwan Kim. Accelerated proximal point method for maximally monotone operators. *Mathematical Programming*, pages 1–31, 2021.

[16] Donghwan Kim and Jeffrey A. Fessler. Optimizing the efficiency of first-order methods for decreasing the gradient of smooth convex functions. *Journal of Optimization Theory and Applications*, 188(1):192–219, 2021.

[17] Etienne De Klerk, François Glineur, and Adrien B Taylor. On the worst-case complexity of the gradient method with exact line search for smooth strongly convex functions. *Optimization Letters*, 11(7):1185–1199, 2017.

[18] Yoel Drori and Adrien B Taylor. Efficient first-order methods for convex minimization: a constructive approach. *Mathematical Programming*, 184(1):183–220, 2020.

[19] Adrien Taylor and Yoel Drori. An optimal gradient method for smooth strongly convex minimization. *arXiv:2101.09741*, 2021.
[20] Arkadi Nemirovski. Information-based complexity of convex programming. *Lecture notes, http://www2.isye.gatech.edu/ñemirovs/Lec_EMCO.pdf*, 1994.

Construction of multivariate polynomial approximation kernels via semidefinite programming

Felix Kirschner
(joint work with Etienne de Klerk)

The talk is based on [2]. A classical problem in approximation theory is uniform approximation of a given function by linear combinations of orthogonal polynomials. We say two functions $f, g \in \mathcal{C}(\mathbf{K})$ are orthogonal (w.r.t. a positive finite Borel measure μ on $\mathbf{K}$), if

$$\langle f, g\rangle_\mu := \int_{\mathbf{K}} f(\mathbf{x})g(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) = 0.$$

Let $\{p_\alpha\}_{\alpha\in\mathbb{N}^n}$ be a system of orthogonal polynomials with respect to a measure μ, where p_α is of degree $|\alpha| = \alpha_1+\cdots+\alpha_n$. Consider a kernel $K_r(\mathbf{x}, \mathbf{y}) : \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ given by

$$K_r(\mathbf{x}, \mathbf{y}) := \sum_{\alpha\in\mathbb{N}^n_r} g_\alpha p_\alpha(\mathbf{x})p_\alpha(\mathbf{y}), \tag{1}$$

for given constants g_α for $\alpha \in \mathbb{N}^n_r$, where $\mathbb{N}^n_r = \{\alpha \in \mathbb{N}^n : \alpha_1 + \cdots + \alpha_n \le r\}$. Then the convolution operator, defined as

$$\mathcal{K}^{(r)}(f)(\mathbf{x}) := \int_{\mathbf{K}} f(\mathbf{y})K_r(\mathbf{x}, \mathbf{y})\mathrm{d}\mu(\mathbf{y}),$$

maps any integrable function f to a polynomial of degree at most r, which will serve as an approximation of f. More precisely,

$$\mathcal{K}^{(r)}(f)(\mathbf{x}) = \sum_{\alpha\in\mathbb{N}^n_r} b_\alpha p_\alpha(\mathbf{x})\,, \text{ where } b_\alpha = \langle p_\alpha, f\rangle_\mu\, g_\alpha.$$

The coefficients g_α of the kernel K_r determine the approximation.

Let now $\mathbf{K} = [-1, 1]^n$ and let the system of orthogonal polynomials on $\mathbf{K}$ be given by the multivariate generalization of the well-known Chebyshev polynomials of the first kind given by $T_k(x) := \cos(k \arccos(x))$ for $k \in \mathbb{N}$. In other words, define for $\alpha \in \mathbb{N}^n_r$

$$T_\alpha(\mathbf{x}) = \prod_{i=1}^n T_{\alpha_i}(x_i).$$

The corresponding measure μ is given by

$$\mathrm{d}\mu(\mathbf{x}) = \prod_{i=1}^n \frac{1}{\pi\sqrt{1-x_i^2}}\mathrm{d}\mathbf{x}.$$

We associate the following quantity to a given kernel $K_r(\mathbf{x}, \mathbf{y})$

$$\sigma_r := \left(\int_{\mathbf{K}\times\mathbf{K}} \|\mathbf{x}-\mathbf{y}\|^2 K_r(\mathbf{x},\mathbf{y}) \mathrm{d}\mu(\mathbf{x}) \mathrm{d}\mu(\mathbf{y}) \right)^{1/2},$$

called the *resolution* of the kernel $K_r(\mathbf{x}, \mathbf{y})$ in [3]. We show how to construct kernels K_r with minimal resolution such that $\mathcal{K}^{(r)}(f)$ converges to f uniformly on $\mathbf{K} = [-1, 1]^n$. One can show that the rate of convergence may be bounded in terms of σ_r. For the statement of the main result recall that the modulus of continuity of $f \in \mathcal{C}(\mathbf{K})$ is defined as

$$\omega_f(\delta) := \max_{\substack{\mathbf{x},\mathbf{y}\in\mathbf{K} \\ \|\mathbf{x}-\mathbf{y}\|\leq\delta}} |f(\mathbf{x}) - f(\mathbf{y})|.$$

Now, any kernel satisfying the following four properties may be used to approximate a continuous function on $\mathbf{K}$:

P1. $K_r(\mathbf{x}, \mathbf{y}) = \sum_{\alpha\in\mathbb{N}^n_r} g_\alpha T_\alpha(\mathbf{x}) T_\alpha(\mathbf{y})$, for $g_\alpha \in \mathbb{R}$ for $\alpha \in \mathbb{N}^n_r$
P2. $K_r(\mathbf{x}, \mathbf{y}) \geq 0$ for all $(\mathbf{x}, \mathbf{y}) \in \mathbf{K} \times \mathbf{K}$ and all r;
P3. $\int_{\mathbf{K}} K_r(\mathbf{x}, \mathbf{y}) \mathrm{d}\mu(\mathbf{y}) = 1$ for all $\mathbf{x} \in \mathbf{K}$ for all r;
P4. $\lim_{r\to\infty} \sigma_r = 0$.

Moreover, we have the following result.

Proposition 1. Let $\mathbf{K} = [-1, 1]^n$ and $f : \mathbf{K} \to \mathbb{R}$ be continuous on $\mathbf{K}$ with modulus of continuity ω_f. Under the above conditions P1-P4 on $K_r(\mathbf{x}, \mathbf{y})$ one has $\mathcal{K}^{(r)}(f) \to f$ as $r \to \infty$, uniformly on $\mathbf{K}$. Moreover,

$$\|\mathcal{K}^{(r)}(f) - f\|_{\infty,\mathbf{K}} \leq 2\left(1 + \frac{\pi}{\sqrt{2}}\right) \omega_f(\sigma_r).$$

The difficult property is clearly P2. The strategy is to strengthen the non-negativity condition to make it more tractable. For this we note that a kernel $K_r(\mathbf{x}, \mathbf{y})$ satisfying property P1 is non non-negative on $[-1, 1]^n \times [-1, 1]^n$ if and only if the trigonometric polynomial $G_r(\boldsymbol{\phi}) = \sum_{\alpha\in\mathbb{N}^n_r} g_\alpha \prod_{i=1}^n \cos\alpha_i\phi_i$ is non-negative on $[-\pi, \pi]^n$. Therefore, we can make use of the following result.

Theorem 1. (e.g. Theorem 3.5 in [4]) If p is a positive trigonometric polynomial, then there exists an $r \in \mathbb{N}$ and a hermitian positive semidefinite (psd) matrix M of order $\binom{n+r}{r}$ such that

$$p(\boldsymbol{\phi}) = \left[\exp(\imath\alpha^T\boldsymbol{\phi})\right]^*_{\alpha\in\mathbb{N}^n_r} M \left[\exp(\imath\alpha^T\boldsymbol{\phi})\right]_{\alpha\in\mathbb{N}^n_r}.$$

One can show that if G_r has a hermitian psd representation as in Theorem 1 then it has a real symmetric psd representation as well. We can therefore solve a semidefinite program (SDP) to obtain coefficients g_α such that the corresponding kernel K_r will satisfy P1-P4. For a given $r \in \mathbb{N}$ the SDP is given by

$$\sigma_r^2 = \min \sum_{i=1}^n (1 - g_{e_i})$$

subject to

$$\begin{aligned}\sum_{\alpha\in\mathbb{N}^n_r} g_\alpha \prod_{i\in[n]} \cos\alpha_i\phi_i &= \left[\exp(\imath\alpha^T\boldsymbol{\phi})\right]^*_{\alpha\in\mathbb{N}^n_r} M \left[\exp(\imath\alpha^T\boldsymbol{\phi})\right]_{\alpha\in\mathbb{N}^n_r}\\ g_{(0,\ldots,0)} &= 1\\ M &\succeq 0.\end{aligned}$$

Property P3 is enforced by the constraint $g_{(0,\ldots,0)} = 1$ and the non-negativity, i.e., P2 follows from Theorem 1. The optimal solution M provide the coefficients g_α leading to a kernel satisfying P1 with minimal resolution σ_r ensuring P4. We formulate upper bounds on the optimal values σ_r^2 of the above SDP in the following proposition.

Proposition 2 Fix $n \in \mathbb{N}$. For $r \in \mathbb{N}$ we have

$$\sigma_r^2 \leq n\left(1-\cos\frac{n\pi}{r+n}\right) \sim \frac{n^3\pi^2}{2(r+n)^2} \text{ if } r \gg 0.$$

In the univariate case there exists a closed form solution for the coefficients of the kernel K_r with minimal resolution. It is a open question whether it is possible to find a closed form for the multivariate case as well.

References

[1] Sergei Bernstein. Sur l'ordre de la meilleure approximation des fonctions continues par des polynômes de degré donné. *émoire couronné, Brussels*, 1912.

[2] Felix Kirschner and Etienne de Klerk, "Construction of multivariate polynomial approximation kernels via semidefinite programming," 2022, arXiv:2203.05892.

[3] Alexander Weiße, Gerhard Wellein, Andreas Alvermann, and Holger Fehske. The kernel polynomial method. *Rev. Mod. Phys.*, 78:275–306, Mar 2006.

[4] Bogdan Dumitrescu. *Positive Trigonometric Polynomials and Signal Processing Applications*. Signals and Communication Technology. Springer, Dordrecht, 2007.

The kissing number problem for regular tetrahedra

Andreas Spomer

(joint work with Frank Vallentin, Fernando Mário de Oliveira Filho, Stefan Krupp and Fabrício C. Machado)

What is the maximum number τ_3 of pairwise nonoverlapping regular tetrahedra in $\mathbb{R}^3$ that share a vertex? The question has a long history which can be traced back to Plato. An easy geometric argument shows that the number can not be bigger than 22. By associating each faced of the Icosaedron with a tetrahedron, we obtain a configuration of 20 tetrahedra, that share the center of the Icosaedron as a vertex. This configuration is conjectured to be optimal.

With the help of modern optimization techniques, we are able to improve the upper bound. The problem can be approached by using semidefinite programming. Nonoverlapping configurations can be considered as independent sets of an infinite graph. To relax this problem, we can develop an infinite version of Lovàsz Theta number [1] for this graph.

Let $\mathcal{T}$ be a regular tetrahedron with a vertex at the origin. Assume there exists a Hermitian kernel $\mathcal{K} : \mathcal{C}(\mathrm{SO}(3) \times \mathrm{SO}(3), \mathcal{C})_{\succeq 0}$ such that

$$\begin{aligned} &\mathrm{Re}(\mathcal{K}(A, A)) = \lambda - 1 && \text{for all } A \in \mathrm{SO}(3) \text{ and} \\ &\mathrm{Re}(\mathcal{K}(A, B)) \leq -1 && \text{whenever } A\mathcal{T} \cap B\mathcal{T} = \emptyset, \end{aligned}$$

then $\tau_3 \leq \lambda$.

This optimization problems tend to have a lot of symmetries, which can be exploited to reduce the complexity of the problem. With the help of the representation theory of the special orthogonal group, we are able to reformulate the bound as a semidefinite optimization problem, which can be solved numerically

$$\begin{aligned} \inf \quad & 1 + f(0,0,0) \\ \text{s.t.} \quad & f : [0, 2\pi) \times [0, \pi) \times [0, 2\pi) \to \mathbb{C}, \\ & F_\ell \succeq 0 \text{ for all } \ell \geq 0, \\ & f(\alpha, \beta, \gamma) = \sum_{\ell=0}^{\infty} \langle A_\ell(\alpha, \beta, \gamma), F_\ell \rangle \\ & \mathrm{Re}(f(A)) \leq -1, \text{whenever } \mathcal{T} \cap A\mathcal{T} = \emptyset. \end{aligned}$$

In my talk, I would like to give an introduction to thee-point bounds and illustrate how they can be applied to geometric optimization problems. I will explain how symmetry reduction and representation theory help to compute these bounds efficiently. Moreover, I will present some of my numerical results concerning the kissing number for regular tetrahedra and discuss applications of this framework.

References

[1] L. Lovàsz, *On the Shannon capacity of a graph*, IEEE Transactions on Information Theory **25** (1979), 1–7.

Semidefinite programming bounds for the average kissing number

Maria Dostert

(joint work with Alexander Kolpakov, Fernando Mário de Oliveira Filho)

The average kissing number of $\mathbb{R}^n$ is the supremum of the average degree of contact graphs of packings of finitely many balls (of any radii) in $\mathbb{R}^n$. A packing of balls in $\mathbb{R}^n$ is a finite set of interior-disjoint closed balls. Furthermore, the contact graph of a packing $\mathcal{P}$ is the graph with vertex set $\mathcal{P}$ in which two balls are adjacent if they intersect, that is, if they are tangent to each other.

Contact graphs of packings of disks on the plane are characterized by the Koebe-Andreev-Thurston theorem [7]: they are precisely the (simple) planar graphs. In higher dimensions, no such simple characterization is known (see the paper by Glazyrin [4] for a nice discussion), and therefore research has been focused on understanding the behavior of some specific parameters of contact graphs.

In this talk, we consider the average degree of contact graphs. More precisely, we are interested in the *average kissing number* of $\mathbb{R}^n$, namely

$$\kappa_n = \sup\{\, \overline{\delta}(G) : G \text{ is the contact graph of a packing of balls in } \mathbb{R}^n \,\},$$

where $\overline{\delta}(G)$ denotes the average degree of G.

Lower bounds for κ_n can be obtained by constructions; a simple idea is to consider lattice packings. Given a lattice $\Lambda \subseteq \mathbb{R}^n$ with shortest vectors of length d, we consider the set of all balls of radius $d/2$ centered on the lattice points. These balls have disjoint interiors and so we have a packing of infinitely many balls. Each ball in this packing has the same number of tangent balls, called the *kissing number* of the lattice Λ. The *lattice kissing number* of $\mathbb{R}^n$, denoted by τ_n^*, is the largest kissing number of any lattice in $\mathbb{R}^n$; immediately we have $\kappa_n \geq \tau_n^*$. Conway and Sloane [1, Table 1.2] list lower bounds for τ_n^*, and hence for κ_n, for n up to 128. For $n = 3$, a construction of Eppstein, Kuperberg, and Ziegler [3] gives $\kappa_3 \geq 12.612$, while $\tau_3^* = 12$.

On the side of upper bounds, it is easy to see that $\kappa_n \leq 2\tau_n$, where τ_n is the *kissing number* of $\mathbb{R}^n$, that is, the maximum number of interior-disjoint unit balls that can simultaneously touch a central unit ball. Indeed, say $\mathcal{P}$ is a packing of balls and let $r(X)$ be the radius of the ball $X \in \mathcal{P}$; let $G = (\mathcal{P}, E)$ be the contact graph of $\mathcal{P}$. In G, the number of neighbors of a ball $X \in \mathcal{P}$ that have radius at least $r(X)$ is at most the kissing number τ_n. So

$$|E| \leq \sum_{X \in \mathcal{P}} |\{\, \{X, Y\} \in E : r(X) \leq r(Y) \,\}| \leq \tau_n |\mathcal{P}|,$$

whence the average degree of G is $2|E|/|\mathcal{P}| \leq 2\tau_n$. Though simple, this bound is still the best known for all $n \geq 10$.

Dimension	*Lower bound*	*Previous upper bound*	*New upper bound*
3	12.612	13.955	13.606
4	24	34.681	27.439
5	40	77.757	64.022
6	72	156	121.105
7	126	268	223.144
8	240	480	408.386
9	272	726	722.629

TABLE 1. Lower and upper bounds for the average kissing number. The lower bound in dimension 3 was given by Eppstein, Kuperberg, and Ziegler [3]; all other lower are in listed by Conway and Sloane [1, Table 1.2]. Upper bounds in dimensions 3, ..., 5 are due to Glazyrin [4]; all other upper bounds are twice the best known upper bound for the kissing number; see Table 1 in Machado and Oliveira [6].

Kuperberg and Schramm [5] gave the first nontrivial upper bound for the average kissing number in dimension 3, proving that $\kappa_3 \leq 8 + 4\sqrt{3} = 14.928\ldots$.

Glazyrin [4] refined their approach and showed that $\kappa_3 \leq 13.955$; he also extended their result to higher dimensions and managed to beat the upper bound of $2\tau_n$ for $n = 4$ and 5. In this talk, I will explain how, we use, in [2], semidefinite programming to refine Glazyrin's approach, obtaining better upper bounds for $n = 3, \ldots, 9$; see Table 1.

References

[1] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices, and Groups*, Grundlehren der mathematischen Wissenschaften 290, Springer-Verlag, New York, 1988.

[2] M. Dostert, A. Kolpakov and F.M. de Oliveira Filho, *Semidefinite programming bounds for the average kissing number*, Israel Journal of Mathematics TBD (2022), 1–25.

[3] D. Eppstein, G. Kuperberg, and G.M. Ziegler, Fat 4-polytopes and fatter 3-spheres, in: *Discrete geometry*, Monographs and Textbooks in Pure and Applied Mathematics 253, Dekker, New York, 2003, pp. 239–265.

[4] A. Glazyrin, Contact graphs of ball packings, arXiv:1707.02526, 2017, 15pp.

[5] G. Kuperberg and O. Schramm, Average kissing numbers for non-congruent sphere packings, *Mathematical Research Letters* 1 (1994) 339–344.

[6] F.C. Machado and F.M. de Oliveira Filho, Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry, *Experimental Mathematics* 27 (2018) 362–369.

[7] K. Stephenson, *Introduction to Circle Packing: The Theory of Discrete Analytic Functions*, Cambridge University Press, Cambridge, 2005.

Three-point bounds for sphere packing

David de Laat

The sphere packing problem asks for the densest packing of Euclidean space by congruent balls. With B_r^n the ball of radius r centered at the origin, the optimal sphere packing density in $\mathbb{R}^n$ can be defined as the maximum of

$$\limsup_{R\to\infty} \frac{\mathrm{vol}(B_{1/2}^n)|P \cap B_R^n|}{\mathrm{vol}(B_R^n)}$$

over all $P \subseteq \mathbb{R}^n$ with distance at least 1 between any two distinct points. This problem has been solved in dimensions 1, 2, 3, 8, and 24. The 3-dimensional version of this problem was known as the Kepler conjecture and has been solved by Hales in what is one of the most well-known computer assisted proofs [8]. Dimensions 8 and 24 have been solved more recently by Viazovka and others [10, 4] using the Cohn-Elkies linear programming bound [3].

An easy nonconstructive argument shows the sphere packing density in $\mathbb{R}^n$ is at least 2^{-n}, and in [9] Kabatyanskii and Levenshtein give the asymptotic upper bound

$$2^{-(0.59905576+o(1))n}.$$

In [6] it is shown that this upper bound can be recovered by a sequence of feasible solutions to the Cohn-Elkies linear programming bound. The question then arises whether the linear programming bound can be used to obtain a better asymptotic bound. In [1] we compute the linear programming bound for high dimensions and

based on this we conjecture it can be used to get the asymptotic upper bound $2^{-(\lambda+o(1))n}$ with $0.604 < \lambda < 0.605$.

The Cohn-Elkies linear programming bound is not expected to be sharp for any dimensions other than 1, 2 (for which this is still a conjecture), 8, and 24, and apart from the small improvements in [7] no improvements on the linear programming bound for sphere packing are known. We therefore would like to have better bounds. In [5] we give two three-point bounds to go beyond the linear programming bounds (which can be called two-point bounds).

For the first bound we show that the Lovász theta prime number, which is a semidefinite program to upper bound the independence number of a finite graph, is multiplicative for the disjunctive product of graphs. Since the Cohn-Elkies linear programming bound can be thought of as the Lovász theta prime number for an infinite noncompact graph, this gives a new formulation for exactly the same sphere packing bound. The advantage of this new formulation is that we can strengthen it when considering the easier problem of finding the optimal lattice packing density. Based on the computational results we conjecture this new bound is sharp for the lattice sphere packing problem in dimension 4. This shows three-point bounds can be very strong for sphere packing problems. Moreover, it might allow for the use of modular form techniques such as those developed by Viazovska to give an alternative proof of the classical result by Korkine and Zolotareff that the D_4 root lattice gives the optimal lattice sphere packing in $\mathbb{R}^4$.

To get better upper bounds for the sphere packing problem we adapt the Bachoc-Vallentin three-point bound for the spherical code problem [2], which is the compact analogue of the sphere packing problem, to a noncompact setting. The crucial new ingredient that makes the bound work is that we show the constraints on the three-point function must be truncated. We use this bound to give improved upper bounds on the optimal sphere packing density in dimensions 4 through 7, and 9 through 12.

References

[1] N. Afkhami-Jeddi, H. Cohn, T. Hartman, D. de Laat, and A. Tajdini, *High-dimensional sphere packing and the modular bootstrap*, Journal of High Energy Physics **2020** (2020), no. 12, 66.

[2] C. Bachoc and F. Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. **21** (2008), no. 3, 909–924. arXiv:math/0608426 MR 2393433

[3] H. Cohn and N. Elkies, *New upper bounds on sphere packings I*, Ann. of Math. (2) **157** (2003), no. 2, 689–714. arXiv:math.MG/0110009 MR 1973059

[4] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska, *The sphere packing problem in dimension 24*, Ann. of Math. (2) **185** (2017), no. 3, 1017–1033. arXiv:1603.06518 MR 3664817

[5] H. Cohn, D. de Laat, and A. Salmon, *Three-point bounds for sphere packing*, in preparation, 2022.

[6] H. Cohn and Y. Zhao, *Sphere packing bounds via spherical codes*, Duke Math. J. **163** (2014), no. 10, 1965–2002. arXiv:1212.5966 MR 3229046

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

[8] T. C. Hales, *A proof of the Kepler conjecture*, Ann. of Math. (2) **162** (2005), no. 3, 1065–1185. MR 2179728
[9] G. A. Kabatyanskii and V. I. Levenshtein, *Bounds for packings on a sphere and in space*, Problemy Peredachi Informatsii **14** (1978), no. 1, 3–25, English translation in Problems of Information Transmission **14**, 1–17, 1978.
[10] M. S. Viazovska, *The sphere packing problem in dimension 8*, Ann. of Math. (2) **185** (2017), no. 3, 991–1015. arXiv:1603.04246 MR 3664816

Bounding the Distance to Unsafe Sets

Jared Miller
(joint work with Mario Sznaier)

This talk presents a method to lower-bound the distance of closest approach between points on an unsafe set and points along system trajectories [1]. Such a minimal distance is a quantifiable and interpretable certificate of safety of trajectories, as compared to prior art in barrier which offers a binary indication of safety/unsafety [2]. As an example, receiving a report that a planned course of action is safe (car traveling at 60 km/hr does not crash into a tree) yields a different reaction as compared to a report that includes the minimal distance (car traveling at 60 km/hr passes within a minimum distance of 10 cm to a tree). The distance estimation problem is converted into a infinite-dimensional Linear Program (LP) in occupation measures based on existing work in optimal control [3], peak estimation [4] and optimal transport [5]. This LP involves an initial measure (initial state), a peak measure (time and state), an occupation measure (time and state), and a joint measure (state and point on unsafe set). Under mild conditions (e.g. Lipschitz dynamics, compact sets) conditions, the LP has the same objective value as the true distance of closest approach.

The moment-Sum of Squares hierarchy is used to obtain a sequence of lower bounds obtained through solving Linear Matrix Inequalities (LMIs) in increasing size [6], and these lower bounds will converge to the true minimal distance as the degree approaches infinity under an Archimedean assumption. The size of the largest positive semidefinite matrix constraint (moment matrix) grows as $\binom{2n+d}{d}$ in an n-state d-degree problem. This large matrix of size $\binom{2n+d}{d}$ may be reduced to n instances of matrices with size $\binom{n+1+d}{d}$ in the case where the distance function is separable (e.g. squared L_2 distance) through the application of correlative sparsity [7]. Near-optimal trajectories that achieve the minimal distance may be recovered if the solved moment matrices obey rank constraints [8].

The distance estimation problem can be modified to accommodate dynamics with uncertainty [9]. Piecewise (norm) distance functions (e.g. L_1, L_3, and L_∞ distances) may be treated using the theory of polyhedral liftings [10]. Safety of shapes traveling in an evolving orientation along trajectories may be assured by bounding the set-set distance between points on the shape and points on the unsafe set with a sequence of LMIs. Such a set-set program involves a coordinate transformation between the orientation of the shape and local coordinates on the shape to the global coordinates for which distance is measured. Set-set distance

estimation adds a new shape measure (orientation and local coordinate) and push-forward transformations to the LP and LMIs.

References

[1] J. Miller and M. Sznaier, "Bounding the Distance to Unsafe Sets with Convex Optimization," 2021, arXiv:2110.14047.

[2] S. Prajna and A. Jadbabaie, "Safety Verification of Hybrid Systems Using Barrier Certificates," in *International Workshop on Hybrid Systems: Computation and Control*. Springer, 2004, pp. 477–492.

[3] R. Lewis and R. Vinter, "Relaxation of optimal control problems to equivalent convex programs," *Journal of Mathematical Analysis and Applications*, vol. 74, no. 2, pp. 475–493, 1980.

[4] M. J. Cho and R. H. Stockbridge, "Linear Programming Formulation for Optimal Stopping Problems," *SIAM J. Control Optim.*, vol. 40, no. 6, pp. 1965–1982, 2002.

[5] C. Villani, *Optimal Transport: Old and New*. Springer Science & Business Media, 2008, vol. 338.

[6] J. B. Lasserre, *Moments, Positive Polynomials And Their Applications*, ser. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009.

[7] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity," *SIOPT*, vol. 17, no. 1, pp. 218–242, 2006.

[8] J. Miller, D. Henrion, and M. Sznaier, "Peak Estimation Recovery and Safety Analysis," *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 1982–1987, 2020.

[9] J. Miller, D. Henrion, M. Sznaier, and M. Korda, "Peak Estimation for Uncertain and Switched Systems," 2021.

[10] M. Yannakakis, "Expressing combinatorial optimization problems by Linear Programs," *Journal of Computer and System Sciences*, vol. 43, no. 3, pp. 441–466, 1991.

Stability and performance verification of dynamical systems controlled by neural networks: algorithms and complexity

Milan Korda

This talk discussed several contributions to stability and performance verification of nonlinear dynamical systems controlled by neural networks as depicted in Figure 1.

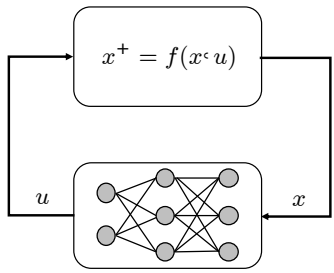

Figure 1. Nonlinear dynamical system controlled by a neural network.

First, we show that the stability and performance of a polynomial dynamical system controlled by a neural network with semialgebraically representable activation functions (e.g., ReLU) can be certified by convex semidefinite programming. The result is based on the fact that the semialgebraic representation of the activation functions and polynomial dynamics allows one to search for a Lyapunov function using polynomial sum-of-squares methods. Second, we remark that even in the case of a linear system controlled by a neural network with ReLU activation functions, the problem of verifying asymptotic stability is undecidable. Finally, under additional assumptions, we establish a converse result on the existence of a polynomial Lyapunov function for this class of systems. Numerical results with on examples of state-space dimension up to 50 and neural networks with several hundred neurons and up to 30 layers demonstrate the method. The content of the talk is based on the work [1] which is itself based on the older work [2] treating the stability of optimization-based controllers (e.g., model predictive control) by viewing them as difference inclusions with semialgebraic right-hand side.

Specifically, we consider discrete-time dynamical systems of the form

$$x^+ = f(x,u) \tag{1}$$

with $x \in \mathbb{R}^n$ being the state, $x^+ \in \mathbb{R}^n$ the successor state, $u \in \mathbb{R}^m$ the control input and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ a polynomial transition mapping. The goal is the verify the closed-loop stability and performance of system (1) when controlled by a neural network controller $u = \boldsymbol{\psi}(x)$. That is, the object of interest is the system

$$x^+ = f(x, \boldsymbol{\psi}(x)), \tag{2}$$

where $\boldsymbol{\psi}$ is a neural network of the form

$$\boldsymbol{\psi}(x) = W_N(\ldots \rho_2(W_2\rho_1(W_1x + b_1) + b_2)\ldots) + b_N \tag{3}$$

for some weight matrices W_i and bias vectors b_i. The activation functions ρ_i, applied componentwise on the output of each layer, are assumed to be semialgebraic; this is satisfied, e.g., for the ReLU, leaky ReLU or the saturation function[1]. The semialgebraicity of the activation functions implies that the graph of the function $\boldsymbol{\psi}$ can be expressed as

$$\begin{aligned} \operatorname{graph}_{\boldsymbol{\psi}} = \{(x,u) \mid \exists\, \lambda \in \mathbb{R}^{n_\lambda} \ \text{ s.t. } \ & g(x,u,\lambda) \geq 0, \\ & h(x,u,\lambda) = 0\} \end{aligned} \tag{4}$$

for some vectors of polynomials g and h and lifting variables λ associated to the semialgebraic functions ρ_i in $\boldsymbol{\psi}$. We recall that the graph of a function $\boldsymbol{\psi} : \mathbb{R}^n \to \mathbb{R}^m$ is a subset of $\mathbb{R}^{n+m}$ defined as

$$\operatorname{graph}_{\boldsymbol{\psi}} = \big\{\big(x, \boldsymbol{\psi}(x)\big) \mid x \in \mathbb{R}^n\big\}.$$

[1]The saturation function is typically applied at the output layer in order to enforce satisfaction of bounds on the control. Other activation functions such as tanh or sigmoids are not semialgebraic and hence cannot be treated using the presented approach without a further approximation.

Since for each x, the control input u satisfies $u = \boldsymbol{\psi}(x)$, it follows that $(x, u) \in \text{graph}_{\boldsymbol{\psi}}$ and hence also

$$u \in \mathbf{K}_x, \tag{5}$$

where the set $\mathbf{K}_x$ is given by

$$\mathbf{K}_x = \big\{u \mid \exists \lambda \in \mathbb{R}^{n_\lambda} \text{ s.t. } g(x, u, \lambda) \geq 0,\ h(x, u, \lambda) = 0\big\}.$$

We note that in this case, the set $\mathbf{K}_x$ is a singleton although the approach of [2] that this work is based on applies to non-singleton sets $\mathbf{K}_x$ as well.

Example (ReLU) Consider the single-neuron network with a ReLU activation function, i.e.,

$$\boldsymbol{\psi}(x) = \text{ReLU}(w^\top x + b) = \max(w^\top x + b, 0)$$

for some vector of weights $w \in \mathbb{R}^n$ and a bias $b \in \mathbb{R}$. The graph of the function $y = \text{ReLU}(z) = \max(0, z)$ is given by

$$\text{graph}_{\text{ReLU}} = \big\{(z, y) \mid y \geq z,\ y \geq 0,\ y(y - z) = 0\big\}.$$

Substituting $w^\top x + b$ for z and u for y , it follows that the set $\mathbf{K}_x$ is given by

$$\mathbf{K}_x = \{u \mid u \geq w^\top x + b,\ u \geq 0,\ u(u - w^\top x - b) = 0\}.$$

We note that in this case, no lifting variables λ are needed.

References

[1] Korda, Milan. Stability and performance verification of dynamical systems controlled by neural networks: algorithms and complexity. arXiv preprint arXiv:2102.02273 (2021).

[2] Korda, Milan, and Colin N. Jones. Stability and performance verification of optimization-based controllers. Automatica 78 (2017): 34-45.

Reporter: Felix Kirschner

Participants

Daniel Brosch
Department of Econometrics & Operations Research
Tilburg University
P. O. Box 90153
5000 LE Tilburg
NETHERLANDS

Valentin Dannenberg
Fachbereich Mathematik
Universität Rostock
Universitätsplatz 1
18055 Rostock
GERMANY

Prof. Dr. Etienne de Klerk
Department of Econometrics and Operations Research
Tilburg School of Economics and Management
Tilburg University
Warandelaan 2
P.O. Box 90153
5000 LE Tilburg
NETHERLANDS

David de Laat
Delft University of Technology
P. O. Box 5031
2600 GA Delft
NETHERLANDS

Dr. Fernando Mario de Oliveira Filho
Dept. of Mathematics and Computer Science
Delft University of Technology
P.O. Box Postbus 5
2600AA Delft
NETHERLANDS

Dr. Maria Dostert
Department of Mathematics
KTH
10044 Stockholm
SWEDEN

Dr. Elisabeth Gaar
Institut für Produktions- und Logistikmanagement
Johannes Kepler Universität Linz
Altenberger Str. 69
4040 Linz
AUSTRIA

Dr. Konstantin Golubev
Greifenseestrasse 11
8050 Zürich
SWITZERLAND

Prof. Dr. Georgina Hall
INSEAD Fountainebleau
PMLS 0.16
Boulevard de Constance
77305 Fontainebleau
FRANCE

Prof. Dr. Didier Henrion
LAAS - CNRS
Université de Toulouse
7, avenue du Colonel Roche
31400 Toulouse
FRANCE

Felix Kirschner
Department of Econometrics
Tilburg University
P. O. Box 90153
5000 LE Tilburg
NETHERLANDS

Milan Korda
LAAS-CNRS
7, Avenue du Colonel Roche
31077 Toulouse Cedex 4
FRANCE

Prof. Dr. Monique Laurent
Centrum Wiskunde & Informatica (CWI)
Postbus 94079
1090 GB Amsterdam
NETHERLANDS

Dr. Victor Magron
LAAS - CNRS
Équipe MAC, bureau E 48
7, avenue du Colonel Roche
31031 Toulouse Cedex
FRANCE

Jared Miller
Department of Electrical and Computer Engineering
Northeastern University
429 Dana Research Center, 100 Forsyth St
02115 Boston MA 02115-5000
UNITED STATES

Philippe Moustrou
Institut de Mathématiques de Toulouse
Université Toulouse Jean Jaurès
5 Allée Antonio Machado
31058 Toulouse
FRANCE

Dr. Simone Naldi
Institut XLIM
Université de Limoges
123, avenue Albert Thomas
87060 Limoges Cedex
FRANCE

Dr. Dmitrii V. Pasechnik
Department of Computer Science
Oxford University
Wolfson Building
Parks Road
Oxford OX1 3QD
UNITED KINGDOM

Prof. Dr. Mohab Safey El Din
LIP 6, PolSys Team
Sorbonne Université
Boite courrier 169
bureau 321
4 place Jussieu
75252 Paris Cedex 05
FRANCE

Melanie Siebenhofer
Institut für Mathematik
Alpen-Adria-Universität Klagenfurt
Universitätsstraße 65-67
9020 Klagenfurt
AUSTRIA

Prof. Dr. Renata Sotirov
Dept. of Econometrics & Operations Research
Tilburg University
P.O. Box 90153
5000 LE Tilburg
NETHERLANDS

Andreas Spomer
Department Mathematik/Informatik
Universität zu Köln
Weyertal 86-90
50931 Köln
GERMANY

Dr. Adrien Taylor
National Institute for Research in Computer Science and Control
INRIA - CS 42112
2, Rue Simone Iff
75020 Paris Cedex
FRANCE

Prof. Dr. Frank Vallentin
Department Mathematik/Informatik
Universität zu Köln
Weyertal 86-90
50931 Köln
GERMANY

Prof. Dr. Angelika Wiegele
Institut für Mathematik
Alpen-Adria-Universität Klagenfurt
Universitätsstraße 65-67
9020 Klagenfurt
AUSTRIA