Mathematisches Forschungsinstitut Oberwolfach

Report No. 21/2026

DOI: 10.4171/OWR/2026/21

Conic Linear Optimization for Computer-assisted Proofs

Organized by
Etienne de Klerk, Tilburg
Frank Vallentin, Cologne
Angelika Wiegele, Klagenfurt

17 May – 22 May 2026

Abstract. 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. Examples of conic linear optimization problems include semidefinite programming in the finite-dimensional case, and the generalised problem of moments in the infinite-dimensional case. Here, the underlying cones are the cone of positive semidefinite matrices, and the cone of positive Borel measures (with given support), respectively.

Since the 1980s, efficient algorithms have been developed for (certain) conic optimization problems. The use of computers makes it possible to prove inequalities with (near-) optimal constants or (near-) optimal weights which otherwise would be completely out of reach with classical pen and paper reasoning. Also the field of conic optimization has a powerful duality theory. This duality theory makes it relatively easy to verify rigorously the obtained numerical computer-assisted proofs without for instance using extensive case by case checking.

Thus there has been a flood of new 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.

This workshop was a follow-up on the (half) workshop in 2022 of the same name (Oberwolfach Workshop 2215b).

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

License: Unless otherwise noted, the content of this report is licensed under CC BY SA 4.0.

Introduction by the Organizers

The workshop *Conic linear optimization for computer-assisted proofs*, organized by Etienne de Klerk, Frank Vallentin, and Angelika Wiegele, was attended by 25 participants. The specializations of the participants included mathematical optimization, control theory, real algebraic geometry, and discrete geometry. Particular attention was devoted to the following themes:

- computer-assisted analysis of convergence proofs for first- and higher-order optimization methods;
- convergence rates and degree bounds for conic optimization certificates in polynomial optimization;
- applications of conic optimization to extremal problems in discrete geometry;
- and the use of conic optimization in algorithmic approximation theory.

Here we give a short overview of the talks given at the workshop, in chronological order.

Talks on Monday. François Glineur gave a featured talk on *Analysis of worst-case performance of iterative methods via conic optimization*. He reviewed the method of formulating the exact worst-case performance of suitable iterative optimization algorithms as semidefinite programs.

This theme was continued by Adrien Taylor in his talk *Building First-Order Algorithms from SDP*. He showed how semidefinite programming can be used to design first-order optimization algorithms that are optimal with respect to various performance measures. The talk concluded with a demonstration of the software package PEPit.

Sander Gribling gave a talk entitled *Revisiting the convergence rate of the Lasserre hierarchy for polynomial optimization over the hypercube*. He presented new bounds on the distance between the cone of polynomials that are nonnegative on the hypercube and the cone of sums of squares of bounded degree. These bounds were then used to derive improved convergence rates for the Lasserre hierarchy in polynomial optimization over the hypercube.

Juan C. Vera Lizcano spoke about *Improved degree bounds for Positivstellensätze over general compact semialgebraic sets*. He presented a framework for deriving positivity certificates on general compact semialgebraic sets by exploiting certificates on simpler sets.

In the afternoon, an open problem session was held. Problems were presented by Fernando de Oliveira Filho (intersection criteria for spherical regular triangles), Daniel Varga (applications of logical equivalences in discrete geometry), Didier Henrion (concrete, pathological univariate polynomial optimization problems), Daniel Brosch (local versus global nonnegativity criteria on graphons), and Máté Matolcsi (kissing numbers for hyperbolic norms and finding explicit, nice dual solutions for a linear program giving the geometric fractional chromatic number of a graph).

Talks on Tuesday. Renata Sotirov gave a featured talk entitled *Beyond binarity: Integer semidefinite programs*, in which she discussed extensions of semidefinite programming to semidefinite programming with integer variables, focusing in particular on variables taking values in $\{-1, 0, 1\}$.

Jason Altschuler spoke on *Time-varying stepsize hedging in convex optimization and min-max optimization.* He reviewed time-varying step-size choices that led to provable improved rates of convergence for the gradient descent method for strongly convex functions.

Georgina Hall spoke on *Sum of Squares Submodularity.* She presented tractable sufficient conditions for certifying submodularity based on sum-of-squares techniques and demonstrated their applications for instance to submodular regression and submodular maximization.

Tuesday afternoon was dedicated to presentations by the participating PhD students. Lucas ter Voert: *A higher-order Newton method for optimization with SOS-convex polynomial constraints*, Johannes Schmucker: *Exploiting Term Sparsity and Symmetries in Polynomial Optimization Problems*, and Jonas Britz: *Semidefinite hierarchies for diagonal unitary invariant quantum states.*

Talks on Wednesday. Wednesday's program focused on the use of conic optimization for extremal problems in discrete geometry.

Fernando de Oliveira began by a tutorial talk on *The use of conic optimization hierarchies in discrete geometry* providing the ground work for the talks of David de Laat on *Optimality and uniqueness of the D_4 root system* and Máté Matolcsi on *Linear programming bounds in geometric problems.*

De Laat demonstrated how the second level of the Lasserre hierarchy in polynomial optimization can be used to prove that the optimal kissing configuration in dimension 4 is unique up to orthogonal transformations. Matolcsi explained how Fourier-analytic linear programming techniques, combined with geometric-combinatorial arguments, can be used to resolve a conjecture of Erdős: the density of a measurable subset of the plane avoiding distance 1 is strictly less than 0.25.

In the afternoon participants enjoyed a scenic hike to St. Roman with Schwarzwälder Kirschtorte.

Talks on Thursday. Simon Foucart gave a featured talk about *Application of Optimization in Constructive Approximation: Some Personal Examples.* In this survey lecture, he presented three case studies illustrating the role of conic optimization in algorithmic approximation theory: multivariate Chebyshev polynomials, minimal projections, and computational optimal recovery. This theme was continued by Mioara Joldes who presented the talk *Guess-and-prove Newton-Krawczyk methods for multivariate Chebyshev approximation.* She described the following, typical work flow:

$$\text{semidefinite programming relaxation} \to \text{numerical solution} \\ \to \text{identify patterns} \to \text{proof.}$$

She used this work flow to arrive at results on polynomial approximation for different domains: the ball, the simplex, the cube, and the crosspolytope.

Mareike Dressler described in her talk *Conic Approaches to Nonnegativity with SONC and SAGE* alternative conic certificates for polynomial nonnegativity that are tailored to sparse polynomials, focusing on the SONC and SAGE frameworks.

In *Polarization of lattices: cold spots, spherical designs and linear programming*, Philippe Moustrou showed how linear programming bounds for spherical designs can be used to make progress on the lattice polarization problem for highly symmetric lattices. The lattice polarization problem is a continuous analogue of the lattice sphere covering problem.

The talk *Applications of the geometric fractional chromatic number to problems in geometry* by Daniel Varga continued the theme of Matolcsi's lecture. He discussed the linear programming formulation of the geometric fractional chromatic number and presented several applications to problems in discrete geometry.

Talks on Friday. Daniel Brosch gave a tutorial talk on *Exploiting Symmetry in Conic Optimization*. He surveyed several techniques for reducing the complexity of conic optimization problems by exploiting symmetry, including degree principles, orbit space reduction, and block diagonalization.

The remainder of the morning was devoted to a session on recent results and open problems. Monique Laurent reported that Jineon Baek, just a few days before the workshop resolved a conjecture of De Klerk and Pasechnik from 2002 on the exact convergence rate of a particular sum-of-squares hierarchy for computing the independence number of a graph. For doing this, the assistance of ChatGPT was the key. Frank Vallentin presented a polynomial-time algorithm for the weak membership problem in the cone of sums-of-squares polynomials. Monique Laurent discussed the algebraic boundary of the graph elliptope, showing that the elliptope $\mathcal{E}(G)$ is a spectrahedron (the feasibility set of a semidefinite program) if and only if the graph G is chordal. Didier Henrion presented recent work on exploiting low tensor-rank structure in polynomial optimization. He also described Juan Vera's solution to the open problem posed on Monday concerning a concrete pathological univariate polynomial optimization problem.

Workshop: Conic Linear Optimization for Computer-assisted Proofs

Table of Contents

François Glineur
Analysis of worst-case performance of iterative methods via conic optimization 1311

Adrien Taylor (joint with B. Goujaud, C. Moucer, F. Glineur, J. Hendrickx, A. Dieuleveut, and Y. Drori)
Designing first-order optimization algorithms via semidefinite programming 1314

Sander Gribling (joint with Etienne de Klerk, Juan Vera)
Squared polynomial approximation kernels for the hypercube: improved error bounds and implications for Lasserre hierarchies 1315

Juan C. Vera (joint with Olga Heijmans-Kuryatnikova, Luis F. Zuluaga)
Degree Bounds for Positivstellensätze over general semialgebraic sets . . . 1316

Renata Sotirov (joint with Frank de Meijer, Veronica Piccialli, Antonio M. Sudoso)
Beyond binarity: Integer semidefinite programming for ternary quadratic problems 1318

Jason M. Altschuler (joint with Pablo A. Parrilo and Henry Shugart)
Two Vignettes of Stepsize Hedging 1319

Georgina Hall (joint with Anna Deza)
Sum of Squares Submodularity 1321

Lucas ter Voert (joint with Etienne de Klerk)
A higher-order Newton method for optimization with SOS-convex polynomial constraints 1322

Johannes Schmucker (joint with Daniel Brosch)
Exploiting Term Sparsity and Symmetries in Polynomial Optimization Problems 1323

Jonas Britz (joint with Monique Laurent)
Semidefinite hierarchies for diagonal unitary invariant quantum states . 1326

Fernando Mário de Oliveira Filho
Conic programming in discrete geometry 1326

David de Laat
Optimality and uniqueness of the D_4 root system 1328

Máté Matolcsi
Linear programming bounds in geometric problems 1330

Simon Foucart
Application of Optimization in Constructive Approximation: Some Personal Examples 1332

Mioara Joldes (joint with Denis Arzelier, Florent Bréhard, Jean-Bernard Lasserre)
Guess-and-Prove Methods for Multivariate Chebyshev Approximation . . 1334

Mareike Dressler
Conic Approaches to Nonnegativity with SONC and SAGE 1336

Philippe Moustrou (joint with C. Bachoc, F. Vallentin, M. C. Zimmermann)
Polarization of lattices: cold spots, spherical designs and linear programming 1338

Dániel Varga
Applications of the geometric fractional chromatic number to problems in geometry 1341

Daniel Brosch
Exploiting Symmetry in Conic Optimization 1344

Abstracts

Analysis of worst-case performance of iterative methods via conic optimization

François Glineur

Performance estimation problems (PEPs) provide a systematic framework for computing exact worst-case guarantees of black-box iterative optimization methods over prescribed function classes. Given a method, an initialization condition, and a performance criterion, the goal is to maximize this criterion, i.e. find its worst-case value over all admissible functions and all initial points satisfying the initial condition. For instance, for N steps of gradient descent with a given stepsize, one may ask for the largest possible value of the final objective accuracy $f(x_N) - f(x_*)$ among all convex L-smooth objective functions f and all initial points satisfying $\|x_0 - x_*\| \leq R$, where x_* is a minimizer of function f. The problem is infinite-dimensional, since the function f itself is a variable.

1. Performance estimation of black-box methods with interpolation. We use the black-box nature of the method. First-order methods only use a finite list of oracle outputs at a list of iterates $\{x_i\}_{0 \leq i \leq N}$. We therefore introduce finite data

$$f_i = f(x_i), \qquad g_i = \nabla f(x_i), \qquad i \in I,$$

where typically $I = \{*, 0, \ldots, N\}$ and $g_* = 0$. We obtain a finite-dimensional formulation by removing the function f from the formulation, using interpolation conditions : we need to characterize when the finite list $(x_i, g_i, f_i)_{i \in I}$ can be interpolated by a function in a given class. For L-smooth convex functions (i.e. with an L-Lipschitz continuous gradient), the exact interpolation condition is

$$f_j \geq f_i + \langle g_i, x_j - x_i \rangle + \frac{1}{2L}\|g_i - g_j\|^2, \qquad i, j \in I. \tag{1}$$

Such necessary and sufficient interpolation conditions were the ingredient that turned the initial PEP approach of Drori and Teboulle [1] into exact finite-dimensional formulations for smooth convex and smooth strongly convex functions [2]. The optimization method under study is encoded in the PEP by expressing each iterate x_i as a fixed linear combination of x_0 and of the previously observed gradients $\{g_j\}_{0 \leq j \leq i-1}$. This covers many standard fixed-coefficient first-order methods, including nonconstant stepsize schedules.

2. Lifted conic semidefinite formulation. At this stage, the formulation is finite-dimensional but not yet convex, because scalar products between the vectors x_i and g_i appear in (1). The second idea is to use a Gram matrix lifting. Let matrix P contain the vectors $\{x_i\}_{0 \leq i \leq N}$ and $\{g_i\}_{0 \leq i \leq N}$ appearing in the PEP and let $G = P^T P \succeq 0$ be their Gram matrix. All scalar products become entries of G, hence the finite-dimensional PEP becomes a semidefinite program.

This formulation is dimension-free: every positive semidefinite matrix is the Gram matrix of some finite family of vectors, and the rank of an optimal G gives the dimension of the corresponding worst-case instance.

The solution of this SDP provides an exact worst-case rate and the corresponding worst-case function. The dual of this SDP is equally important, as its solution provides a proof of the worst-case rate expressed as a nonnegative linear combination of the interpolation conditions combined with sums-of-squares terms.

3. Sample results obtained with the PEP methodology.

3.1. One step on smooth strongly convex functions. Let f be μ-strongly convex and L-smooth. Then a step of gradient descent with the optimal stepsize leads to the following exact linear convergence rate, and this factor cannot be improved

$$x_+ = x - \frac{2}{L+\mu} g \quad \Rightarrow \quad \frac{\|x_+ - x_*\|}{\|x - x_*\|} \leq \left(\frac{L-\mu}{L+\mu}\right). \tag{2}$$

3.2. Constant-step gradient descent on smooth convex functions. For convex L-smooth functions, gradient descent with constant stepsize h/L satisfies the exact objective-residual rate

$$\sup_f \frac{f(x_N) - f(x_*)}{\frac{1}{2}L\|x_0 - x_*\|^2} = \max\left\{\frac{1}{2Nh+1}, (1-h)^{2N}\right\}, \qquad \text{for any } 0 \leq h \leq 2. \tag{3}$$

The two terms correspond to two different extremal mechanisms: a Huber-type instance and a quadratic instance, corresponding respectively to the first and second terms in the max. The history of the exact bound (3) is instructive. Drori and Teboulle conjectured it, and provided a proof for $h \leq 1$ in the original PEP paper [1], relying on a relatively unintuitive linear algebra argument. Teboulle and Vaisbourd [5] later give an elementary proof extending the range to $h \leq 3/2$. Finally, a recent preprint by Kim [6] proposes a proof for the full range $h \in [0, 2]$, where a key idea is to use interpolation inequalities involving nonconsecutive iterates (see also the work of Rotaru, Glineur and Patrinos [7] for a similar exact full-range result for the gradient-norm criterion $\frac{\|\nabla f(x_N)\|^2/(2L)}{f(x_0)-f(x_*)}$, also covering strongly convex, weakly convex and nonconvex smooth functions). Other extensions cover composite and proximal variants [3], and exact line-search gradient descent [4].

3.3. Subgradient method for convex functions with bounded subgradients. Consider the subgradient method with constant stepsizes $x_{k+1} = x_k - h(R/B)g_k$, under $\|g\| \leq B$ and $\|x_0 - x_*\| \leq R$. The following exact last-iterate guarantee is provided in Zamani and Glineur [8]

$$\frac{f(x_N) - f(x_*)}{BR} \leq \begin{cases} \left(\frac{1}{2}s_N^2 - N\right)h + \dfrac{1}{2s_N^2 h}, & h \geq s_N^{-2}, \\ 1 - Nh, & h \leq s_N^{-2}. \end{cases}$$

where the sequence $\{s_i\}$ is defined by $s_0 = 1$ and $s_{i+1} = s_i + \frac{1}{s_i}$. Optimizing the stepsize for this exact rate leads to $h^*(N) = \frac{1}{\sqrt{s_N^2(s_N^2 - 2N)}}$, providing the optimal

and exact rate for constant stepsize subgradient methods

$$f(x_N) - f(x_*) \leq BR\sqrt{1 - \frac{2N}{s_N^2}} \lesssim \frac{BR}{\sqrt{N+1}}\sqrt{1 + \frac{1}{4}\log N}.$$

Interestingly, proofs for those results, originally obtained from PEP, could be significantly simplified. The final argument relies on a weighted Jensen-type inequality obtained by tracking the distance from each iterate to a different reference point. An optimal subgradient method with linearly decreasing stepsizes

$$x_{k+1} = x_k - \frac{R}{B}\,\frac{(N+1-k)}{(N+1)^{3/2}}\,g_k \quad \Rightarrow \quad f(x_N) - f(x_*) \leq \frac{BR}{\sqrt{N+1}} \tag{4}$$

is also given in [8], which exactly matches the black-box lower bound for this class. The dependence of the stepsizes on the horizon is unavoidable: no single universal stepsize sequence can attain the rate in (4) simultaneously for all horizons.

4. Method design and open questions. The PEP methodology provides both a computational tool and a proof-discovery mechanism. Its primal side identifies extremal trajectories and worst-case instances, while its dual side provides proofs establishing the corresponding rate. A natural follow-up question consists in solving the design problem, where instead of analyzing a given method one would like to find the best method for a given class and performance criterion, e.g. optimize over stepsize schedules or momentum coefficients. However, the corresponding PEP formulation loses convexity in general, making it much less tractable. Branch-and-bound PEP methods provide one systematic route to globally optimized methods [9], and recent work of Kamri, Hendrickx and Glineur develops numerical approaches for designing optimized first-order algorithms by exploiting PEP structure [10]. A principled and tractable general theory for method design remains a central open question, together with interpolation theory for richer function classes and extensions beyond first-order fixed-step methods.

References

[1] Y. Drori and M. Teboulle, *Performance of first-order methods for smooth convex minimization: a novel approach*, Math. Program. **145** (2014), 451–482.

[2] A. B. Taylor, J. M. Hendrickx and F. Glineur, *Smooth strongly convex interpolation and exact worst-case performance of first-order methods*, Math. Program. **161** (2017), 307–345.

[3] A. B. Taylor, J. M. Hendrickx and F. Glineur, *Exact worst-case performance of first-order methods for composite convex optimization*, SIAM J. Optim. **27** (2017), 1283–1313.

[4] E. de Klerk, F. Glineur and A. B. Taylor, *On the worst-case complexity of the gradient method with exact line search for smooth strongly convex functions*, Optim. Lett. **11** (2017), 1185–1199.

[5] M. Teboulle and Y. Vaisbourd, *An elementary approach to tight worst case complexity analysis of gradient based methods*, Math. Program. **201** (2023), 63–96.

[6] J. Kim, *A proof of the exact convergence rate of gradient descent*, arXiv:2412.04427, 2025.

[7] T. Rotaru, F. Glineur and P. Patrinos, *Exact worst-case convergence rates of gradient descent: a complete analysis for all constant stepsizes over nonconvex and convex functions*, Math. Program., published online 2026, doi:10.1007/s10107-025-02313-1.

[8] M. Zamani and F. Glineur, *Exact convergence rate of the last iterate in subgradient methods*, SIAM J. Optim. **35** (2025), 2182–2201.

[9] S. Das Gupta, B. P. G. Van Parys and E. K. Ryu, *Branch-and-bound performance estimation programming: a unified methodology for constructing optimal optimization methods*, Math. Program. **204** (2024), 567–639.

[10] Y. Kamri, J. M. Hendrickx and F. Glineur, *Numerical design of optimized first-order algorithms*, arXiv:2507.20773, 2025.

Designing first-order optimization algorithms via semidefinite programming

Adrien Taylor

(joint work with B. Goujaud, C. Moucer, F. Glineur, J. Hendrickx, A. Dieuleveut, and Y. Drori)

In this talk, I presented the `PEPit` package [5], which implements performance estimation problems (PEPs) [1, 2, 3] for the convergence analysis of first-order optimization algorithms. PEPs provide a systematic framework for computing worst-case guarantees by formulating the analysis of an algorithm over a class of optimization problems as an optimization problem itself, which can often be expressed or relaxed as a semidefinite program.

I then discussed how this methodology can be used not only to analyze existing methods, but also to design new first-order algorithms with prescribed convergence guarantees [4]. In this approach, one searches over algorithm parameters so as to optimize a worst-case performance bound, thereby turning semidefinite programming into a tool for constructive algorithm design.

These ideas have led to the discovery of provably optimal methods, including OGM [7] for smooth convex minimization and ITEM [6] for smooth strongly convex minimization. The talk emphasized both the theoretical insight provided by the PEP framework and the practical role of `PEPit` as a flexible software tool for computer-assisted analysis and experimentation in first-order optimization.

References

[1] Y. Drori, M. Teboulle (2014). *Performance of first-order methods for smooth convex minimization: a novel approach*. Mathematical Programming 145(1)

[2] A. Taylor, J. Hendrickx, F. Glineur (2017). *Smooth strongly convex interpolation and exact worst-case performance of first-order methods*, Mathematical Programming 161(1).

[3] A. Taylor, J. Hendrickx, F. Glineur (2017). *Exact worst-case performance of first-order methods for composite convex optimization*, SIAM Journal on Optimization 27(3).

[4] Y. Drori, A. Taylor (2020). *Efficient first-order methods for convex minimization: a constructive approach*, Mathematical Programming 184.

[5] B. Goujaud, C. Moucer, F. Glineur, J. Hendrickx, A. Taylor, A. Dieuleveut (2024). *PEPit: computer-assisted worst-case analyses of first-order optimization methods in Python*, Mathematical Programming Computation 16(3).

[6] A. Taylor, Y. Drori (2023). *An optimal gradient method for smooth strongly convex minimization* , Mathematical Programming 199(1).

[7] D. Kim, J. Fessler (2016). *Optimized first-order methods for smooth convex minimization*, Mathematical Programming 159(1).

Squared polynomial approximation kernels for the hypercube: improved error bounds and implications for Lasserre hierarchies

Sander Gribling
(joint work with Etienne de Klerk, Juan Vera)

In this talk, based on [4], we consider the problem of approximating a polynomial f that is nonnegative on the hypercube $[-1,1]^n$ by a sum of squared polynomials. It is known since the work of Berg [2] that the cone of sums of squared polynomials is dense in the cone of nonnegative polynomials, in the coefficient ℓ_1-norm. Lasserre and Netzer gave an explicit proof of this result [5]: they showed that for every $\varepsilon > 0$ there exists an $r(f,\varepsilon)$ such that for all $r \geq r(f,\varepsilon)$ the polynomial $f + \varepsilon\left(1 + \sum_{j=1}^n x_j^{2r}\right)$ is a sum of squared polynomials. Their proof is however not constructive as the form of $r(f,\varepsilon)$ is not known.

Our main contribution is an explicit construction of a sequence of approximations in the coefficient ℓ_1-norm *of the Chebyshev basis*. The coefficient ℓ_1-norm in the Chebyshev basis is at most that of the monomial basis; we do not obtain a convergence rate in the ℓ_1-norm of the monomial basis. The Chebyshev basis however suffices for our main application, as we discuss later.

Our approach is based on the polynomial kernel method. Concretely, we propose a new family of (squared) polynomial approximation kernels for approximating nonnegative polynomials on the hypercube $[-1,1]^n$. Our kernels produce polynomial sums-of-squares of degree r, achieving an $O(\log^3 r/r^2)$ error in the ℓ_1-norm of the coefficients in the Chebyshev basis. This improves on the known error bound $O(1/r)$ from the literature [3].

As an application, which formed the main motivation of our work, we obtain an improved convergence rate for the Lasserre hierarchy for polynomial optimization on the hypercube, again improving a known rate by Baldi and Slot from $O(1/r)$ to $O(\log^3 r/r^2)$ [1].

References

[1] L. Baldi and L. Slot, *Degree bounds for Putinar's Positivstellensatz on the hypercube.* SIAM Journal on Applied Algebra and Geometry, **8(1)**:1–25, 2024.

[2] Christian Berg, *The multidimensional moment problem and semi-groups.* Proc. Symp. Appl. Math., **37**:110–124, 1980.

[3] S. Gribling, E. de Klerk, E., J. Vera, *Revisiting the convergence rate of the Lasserre hierarchy for polynomial optimization over the hypercube.* Optimization, 1–28 (2026).

[4] S. Gribling, E. de Klerk, E., J. Vera, *Squared polynomial approximation kernels for the hypercube: improved error bounds and implications for Lasserre hierarchies*, arXiv:2605.31496 (2026).

[5] J. B. Lasserre, T. Netzer, *SOS approximations of nonnegative polynomials via simple high degree perturbations.* Mathematische Zeitschrift, **256**:99–112, 2007.

Degree Bounds for Positivstellensätze over general semialgebraic sets

Juan C. Vera

(joint work with Olga Heijmans-Kuryatnikova, Luis F. Zuluaga)

Let $g_1, \ldots, g_m \in \mathbb{R}[x]$, and let $S_g := \{x \in \mathbb{R}^n : g_j(x) \geq 0,\ j = 1, \ldots, m\}$ be compact and nonempty. For a polynomial $p \in \mathbb{R}[x]$, consider $p_{\min} := \min_{x \in S_g} p(x)$. A standard way to approximate $p_{\min}$ is to construct algebraic certificates proving that

$$p - \lambda > 0 \text{ on } S_g.$$

Or, equivalently, for a given error tolerance $\varepsilon > 0$, to certifying positivity of $f := p - p_{\min} + \varepsilon$ on S_g. The quantitative question is therefore the following.

> Given $f > 0$ on S_g, how large must the degree of a Positivstellensatz certificate for f be?

Initial results in this direction focused on obtaining degree bounds for Positivstellensatz over simple sets. Specifically, these include the hypercube $[-1, 1]^n$ [4, 11], the binary cube $\{0, 1\}^n$ [16], the $(n-1)$-simplex $\Delta^{n-1} = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$ [9], the n-simplex $\Delta^n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i \leq 1\}$ [15], the unit sphere $S^{n-1} := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1\}$ [7, 9], and the unit ball $B^n = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 \leq 1\}$ [15]. Although progress has been made on deriving analogous degree bounds for Positivstellensatz over more general compact semialgebraic sets, such bounds are not as strong as the ones derived for simple sets.

We derive new degree bounds for Positivstellensätze over general compact semialgebraic sets. To accomplish this, we propose a *lift-and-project* approach for polynomial optimization. This approach leverages degree bounds for Positivstellensätze on the hypercube but differs fundamentally from the existing algebro-geometric reduction methodologies [13, 3, 2]. The approach is motivated by the results in [10]. It adopts principles from classical mixed-integer programming (MIP) in which the MIP is lifted into a higher-dimensional space by adding new variables. This variables capture interactions between original variables, and by imposing relaxed consistency conditions they allow to represent convex hulls more accurately. Projecting back to the original space stronger LP relaxations of the MIP are obtained (see, e.g., [1, 14]).

We begin by scaling S_g so it is contained in the hypercube $[-1, 1]^n$. Next, given $f \in \mathbb{R}[x]$ positive on S_g, we add m new variables $u_1, \ldots, u_m$, one for each constraint $g_j(x) \geq 0$, $j = 1, \ldots, m$. We construct $F \in \mathbb{R}[x, u]$ as a *lifted* version of f. This lifted polynomial is positive on $T := [-1, 1]^n \times [0, 1]^m$. The key aspect of this construction is that eliminating the new variables by the substitution $u = g(x)$ projects T onto S_g and $F(x, u)$ onto $f(x)$. More specifically, $S_g = \{x \in \mathbb{R}^n : (x, g(x)) \in T\}$, and $f(x) = F(x, g(x))$. By applying an effective Positivstellensatz over T, we obtain a certificate of nonnegativity for F over T with bounded degree. After substituting $u = g(x)$, we project this certificate to a certificate of the same type for the nonnegativity of f over S_g. For example, if $F \in Q(1 \pm x, u, 1 - u)_r$, the method yields $f \in Q(g)_{ar+b}$, where a and b depend only on g.

We apply the lift-and-project approach to obtain new degree bounds. Let $f > 0$ on S_g and L_g denote the Łojasiewicz exponent of g [12]. It is important to note that $L_g = 1$ holds under common regularity assumptions [2].

(1) We improve the current best Putinar's Positivstellensatz degree bound from $O((f_{\max,\Delta^n}/f_{\min})^{7L_g+3})$ [3] to $O((f_{\max,[-1,1]^n}/f_{\min})^{L_g})$.
(2) We obtain a $O((f_{\max,[-1,1]^n}/f_{\min})^{L_g})$ degree bound on Schmüdgen's Positivstellensatz. Notably, the degree bounds for Putinar's and Schmüdgen's Positivstellensätze over general semialgebraic sets match the same quadratic improvement as for the hypercube $[-1,1]^n$.
(3) We obtain $O((f_{\max,[-1,1]^n}/f_{\min})^{2L_g})$ degree bounds on Krivine-Stengle's Positivstellensatz and the extended-Handelman's Positivstellensatz. These results provide the first explicit degree bounds for linear optimization-based hierarchies over general compact semialgebraic sets that use nonnegative constants to certify nonnegativity. Interestingly, the degree bounds are similar to those for Putinar's and Schmüdgen's Positivstellensätze which use SOS to certify nonnegativity.

Our methodology offers a unified explanation for several apparently distinct degree estimates. For Putinar and Schmüdgen certificates, the method improves the known effective bounds over general compact semialgebraic sets. For Krivine–Stengle and extended Handelman certificates, it provides explicit degree bounds for linear programming-based positivity hierarchies. The framework narrows the gap between the well-understood case of boxes and the general compact semialgebraic sets that arise naturally in polynomial optimization.

References

[1] Balas, E., Ceria, S., and Cornuéjols, G. (1993a). A lift-and-project cutting plane algorithm for mixed 0–1 programs. *Mathematical Programming*, 58:295–324.

[2] Baldi, L. and Mourrain, B. (2023). On the effective Putinar's Positivstellensatz and moment approximation. *Mathematical Programming*, 200(1):71–103.

[3] Baldi, L., Mourrain, B., and Parusiński, A. (2025). On Łojasiewicz inequalities and the effective Putinar's Positivstellensatz. *Journal of Algebra*, 662:741–767.

[4] Baldi, L. and Slot, L. (2024). Degree bounds for Putinar's Positivstellensatz on the hypercube. *SIAM Journal on Applied Algebra and Geometry*, 8(1):1–25.

[5] de Klerk, E. and Laurent, M. (2010). Error bounds for some semidefinite programming approaches to polynomial minimization on the hypercube. *SIAM Journal on Optimization*, 20:3104–3124.

[6] de Klerk, E. and Vera, J. (2026). The link between 1-norm approximation and effective Positivstellensatze for the hypercube. *Numerical Algebra Control and Optimization*, 16:84–104.

[7] Fang, K. and Fawzi, H. (2021). The sum-of-squares hierarchy on the sphere and applications in quantum information theory. *Mathematical Programming*, 190(1):331–360.

[8] Gribling, S., de Klerk, E., and Vera, J. (2026). Revisiting the convergence rate of the lasserre hierarchy for polynomial optimization over the hypercube. *Optimization*, pages 1–28.

[9] Kirschner, F. and De Klerk, E. (2022). Convergence rates of RLT and Lasserre-type hierarchies for the generalized moment problem over the simplex and the sphere. *Optimization Letters*, 16(8):2191–2208.

[10] Kuryatnikova, O., Vera, J. C., and Zuluaga, L. F. (2024). Reducing nonnegativity over general semialgebraic sets to nonnegativity over simple sets. *SIAM Journal on Optimization*, 34(2):1970–2006.
[11] Laurent, M. and Slot, L. (2023a). An effective version of Schmüdgen's Positivstellensatz for the hypercube. *Optimization Letters*, 17(3):515–530.
[12] Łojasiewicz, S. (1959). Sur le probleme de la division. *Studia Mathematica*, 18(1):87–136.
[13] Schweighofer, M. (2004). On the complexity of Schmüdgen's Positivstellensatz. *Journal of Complexity*, 20(4):529 – 543.
[14] Sherali, H. D. and Adams, W. P. (2013). *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31. Springer Science & Business Media.
[15] Slot, L. (2022b). Sum-of-squares hierarchies for polynomial optimization and the Christoffel–Darboux kernel. *SIAM Journal on Optimization*, 32(4):2612–2635.
[16] Slot, L. and Laurent, M. (2023). Sum-of-squares hierarchies for binary polynomial optimization. *Mathematical Programming*, 197(2):621–660.

Beyond binarity: Integer semidefinite programming for ternary quadratic problems

Renata Sotirov
(joint work with Frank de Meijer, Veronica Piccialli, Antonio M. Sudoso)

Integer semidefinite programming (ISDP) may be seen as a generalization of integer linear programming, where the vector of variables is replaced by a positive semidefinite (PSD) matrix variable whose elements are constrained to be integers. Recent results show that integer semidefinite programs allow for more compact and substantially different formulations for several problem classes compared to traditional modeling approaches, and also enable the development of new algorithmic techniques [5, 6]. The resulting ISDP programs do not include $\mathcal{NP}$-hard rank constraints, but instead rely on an alternative notion of exactness in terms of integrality of PSD matrices. Although embedding an integer linear program into the space of diagonal integer PSD matrices is straightforward, constructing an ISDP for a quadratic integer program remains challenging due to the difficulty of representing general integer PSD matrices. However, a complete characterization of integer PSD matrices exists when the variables are binary or take values in $\{\pm 1\}$.

In this talk, we introduce an integer semidefinite program for the ternary quadratic problem (TQP), that is an optimization problem with quadratic objective and linear constraints, where the decision variables take values in $\{0, \pm 1\}$. The TQP belongs to the class of nonlinear integer programs, which have received increased attention in recent years [1]. We first list some known results on integer PSD matrices and then present new theoretical results of the set of ternary PSD matrices. One of the characterizations of the set of rank-one ternary PSD matrices induces a ternary semidefinite programming formulation for the TQP.

We further introduce the convex hull of all rank-one ternary PSD matrices and study its relation with the polytope induced by the triangle and non-standard split inequalities. We show that those sets do not coincide even in the dimension three.

We also show that unconstrained TQP corresponds to the max-cut problem when diagonal elements in the objective matrix are all non-positive.

To strengthen the SDP relaxation for the TQP obtained from our ISDP formulation, we derive valid inequalities. Since inequalities for binary problems do not directly apply to the ternary case, we derive valid cuts for the ISDP of the TQP using the Chvátal-Gomory procedure for integer semidefinite programs [6]. We also present split and non-standard split inequalities for the TQP that are obtained by exploiting an approach proposed by Letchford [7]. We strengthen our initial SDP relaxation by adding the derived inequalities and introduce a specialized B&B algorithm for solving nonconvex TQP instances. The nonconvexity refers to the quadratic objective, whose associated matrix is indefinite. Our results show that our B&B algorithm significantly outperform all other approaches on all instances and several TQP variations.

The talk is based on [4].

References

[1] Raymond Hemmecke, Matthias Köpe, Jon Lee, and Robert Weismantel. *Nonlinear integer programming.* In Michael Jünger, Thomas Liebling, Denis Naddef, George Nemhauser, William Pulleyblank, Gerhard Reinelt, Giovanni Rinaldi, and Laurence A. Wolsey, editors, 50 Years of Integer Programming 1958–2008: The Early Years and State-of-the-Art Surveys. Springer, Berlin, (2009).

[2] Christoph Buchheim, Marianna De Santis, Laura Palagi, and Mauro Piacentini. *An exact algorithm for nonconvex quadratic integer minimization using ellipsoidal relaxations*, SIAM J. Optim., **23** (2013), 1867–1889.

[3] Christoph Buchheim and Angelika Wiegele. *Semidefinite relaxations for non-convex quadratic mixed-integer programming*, Math. Program., **141** (2012), 435–452.

[4] Frank de Meijer, Veronica Piccialli, Renata Sotirov, Antonio M. Sudoso. *Beyond binarity: Semidefinite programming for ternary quadratic problems*, Preprint (2026), `https://arxiv.org/abs/2603.28979`.

[5] Frank de Meijer and Renata Sotirov, *On integrality in semidefinite programming for discrete optimization*, SIAM J. Optim, **34** (2024), 1071–1096.

[6] Frank de Meijer and Renata Sotirov, *The Chvátal-Gomory Procedure for Integer SDPs with Applications in Combinatorial Optimization*, Math. Program., **209** (2025), 323–395.

[7] Adam N. Letchford. *Integer quadratic quasi-polyhedra.* In Friedrich Eisenbrand and F. Bruce Shepherd, editors, Integer Programming and Combinatorial Optimization, pages 258–270. Springer Berlin Heidelberg (2010).

Two Vignettes of Stepsize Hedging

Jason M. Altschuler

(joint work with Pablo A. Parrilo and Henry Shugart)

In this talk, we describe how the theme of the workshop (conic linear optimization for computer-assisted proofs) has enabled us to dispel longstanding beliefs about the speed limit of gradient descent (GD) in convex optimization and min-max optimization. Complete versions of these results can be found in [1, 2, 4].

Concretely, we revisit the following foundational question in both contexts:

What is the optimal choice of stepsizes for GD?

Below, we first state these two results briefly and informally, and then we explain how the theme of the workshop plays a helpful role in their proofs.

Vignette 1: convex optimization. Let us begin with the simpler setting of convex optimization. Consider problems of the form $\min_{x\in\mathbb{R}^d} f(x)$ where f is convex and 1-smooth (i.e., ∇f is 1-Lipschitz). It is a celebrated fact that with an appropriate choice of stepsizes $\{\alpha_t\}$, the GD algorithm

$$x_{t+1} = x_t - \alpha_t \nabla f(x_t)$$

solves any such convex optimization problem to arbitrary accuracy from any initialization x_0. The mainstream approach—which can be found in any textbook on convex optimization—is to use constant stepsize schedules $\alpha_t \equiv \bar{\alpha}$ since this ensures that the error $f(x_n) - \min_x f(x)$ converges at the rate $O(1/n)$.

Is faster convergence possible by optimizing the stepsizes? Conventional wisdom says no. Surprisingly, however, the answer is yes. In this talk, we describe the silver stepsize schedule [1, 2], which is a time-varying, non-monotonic, fractal-like schedule that leads to faster convergence at the rate $O(1/n^{\log_2(1+\sqrt{2})}) \approx O(1/n^{1.27})$.

Vignette 2: min-max optimization. Next, consider problems of the form $\min_{x\in\mathbb{R}^d} \max_{y\in\mathbb{R}^d} f(x,y)$ where f is convex-concave and 1-smooth. This is the direct min-max analog of convex optimization. The natural analog of the GD algorithm is gradient-descent-ascent (GDA):

$$\begin{aligned} x_{t+1} &= x_t - \alpha_t \nabla_x f(x_t, y_t)\,, \\ y_{t+1} &= y_t + \beta_t \nabla_y f(x_t, y_t)\,. \end{aligned}$$

The idea is that for stepsizes $\alpha_t, \beta_t \geq 0$, the variable x attempts to descend while y attempts to ascend. However, a key stumbling block is that GDA fails to converge even in simple settings such as $\min_{x\in\mathbb{R}} \max_{y\in\mathbb{R}} xy$. This failure occurs for any stepsize schedule $\alpha_t, \beta_t \geq 0$.

Is there a missed opportunity? Surprisingly, again, the answer is yes. In this talk, we describe the slingshot stepsize schedule, which is time-varying, asymmetric, periodically negative, and enables convergence for unconstrained smooth convex-concave problems—the direct analog of smooth convex optimization [4].

Conic linear optimization for computer-assisted proofs. A common thread in these two vignettes is that the relevant progress is not visible one step at a time. Standard analyses of GD are inherently local: they prove a descent inequality for a single iteration (e.g., via the descent lemma) and then sum or telescope. This naturally suggests greedy prescriptions such as constant stepsizes, and is too shortsighted to detect the benefit of hedging across iterations.

The algorithmic opportunity comes from analyzing multi-step progress *holistically*. One should choose stepsizes whose iterations compensate for one another, even if some individual steps look suboptimal—or, in the min-max setting, even

perverse. Proving such statements requires global analyses that account for how different iterations may synergize and hedge between different worst-case objective functions. A useful starting point is the framework of performance estimation problems (PEP), pioneered by [5] and refined by [6]. Said succinctly, PEP provides a computer-assisted way of proving convergence rates of optimization algorithms by linearly combining valid inequalities for the relevant function class. This is where the theme of the workshop enters.

However, although PEP is helpful, it does *not* solve the problem at hand for several reasons. The most important reason is that PEP does not directly help with choosing stepsizes: optimizing the stepsize schedule appears to be a nonconvex problem. In this talk, I describe these challenges and our approaches for addressing this—based on a technique called "recursive gluing" in the setting of convex optimization, and based on constructing appropriate Lyapunov functions in the setting of min-max optimization. The interested reader can find more details and context in [1, 2, 4] or in the short expository article [3].

References

[1] J. M. Altschuler and P. A. Parrilo, Acceleration by Stepsize Hedging: Multi-Step Descent and the Silver Stepsize Schedule, *Journal of the ACM* **72**(2) (2025), Article 12, 1–38.

[2] J. M. Altschuler and P. A. Parrilo, Acceleration by Stepsize Hedging: Silver Stepsize Schedule for Smooth Convex Optimization, *Mathematical Programming* **213** (2025), 1105–1118.

[3] J. M. Altschuler and P. A. Parrilo, Stepsize Hedging: an Alternative Mechanism for Accelerating Gradient Descent, invited research highlight, *INFORMS Computing Society Newsletter*, to appear in 2026.

[4] H. Shugart and J. M. Altschuler, Negative Stepsizes Make Gradient-Descent-Ascent Converge, preprint, arXiv:2505.01423, 2025.

[5] Y. Drori and M. Teboulle, Performance of first-order methods for smooth convex minimization: a novel approach, *Mathematical Programming*, 145(1–2):451–482, 2014.

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

Sum of Squares Submodularity

Georgina Hall

(joint work with Anna Deza)

We introduce the notion of *t-sum of squares (sos) submodularity*, a hierarchy, indexed by t, of sufficient algebraic conditions for certifying submodularity of set functions. We show that, for fixed t, each level of the hierarchy can be verified via a semidefinite program of size polynomial in n, the size of the ground set of the set function. This is particularly relevant given existing hardness results around testing whether a set function is submodular [1]. We derive several equivalent algebraic characterizations of t-sos submodularity and identify submodularity-preserving operations that also preserve t-sos submodularity. We further present a complete classification of the cases for which submodularity and t-sos submodularity coincide, as well as examples of t-sos-submodular functions. We demonstrate the

usefulness of t-sos submodularity through three applications: (i) a new convex approach to submodular regression, involving minimal manual tuning; (ii) a systematic procedure to derive lower bounds on the submodularity ratio in approximate submodular maximization; and (iii) improved difference-of-submodular decompositions for difference-of-submodular optimization. Overall, our work builds a new bridge between discrete optimization and real algebraic geometry by connecting sum of squares-based algebraic certificates to a fundamental discrete structure, submodularity.

References

[1] Yves Crama. Recognition problems for special classes of polynomials in 0–1 variables. *Mathematical Programming*, 44:139–155, 1989. DOI: 10.1007/BF01587085.

A higher-order Newton method for optimization with SOS-convex polynomial constraints

Lucas ter Voert

(joint work with Etienne de Klerk)

It is well known that Newton's method converges locally quadratically to an optimal solution to an unconstrained optimization problem if the objective function is strongly convex and its Hessian is Lipschitz continuous. To obtain improved convergence properties, there has been extensive work on designing Newton methods of order higher than two by, e.g., Nesterov and Polyak [7], Nesterov [6], Doikov and Nesterov [3], Silina and Zhang [10], and Cartis and Zhu [2].

In 2024, Ahmadi, Chaudhry, and Zhang [1] designed such a method of which the time per iteration is polynomial in the number of variables and the local order of convergence equals the order of the Taylor expansion used, using *sum-of-squares (SOS)* techniques in polynomial optimization (see, e.g., [8, 5]). Given an iterate x, a semidefinite program (SDP) is solved to construct an *SOS-convex* approximation of a higher-order Taylor expansion of the objective function at the iterate x, and the next iterate x^+ is defined as the minimizer of this polynomial, which can be found by solving another SDP. It is well known that, under some assumptions, an SDP can be solved to arbitrarily small error in polynomial time (see, e.g. [11]).

We extended the method of [1] to optimization with SOS-convex polynomial constraints, and proved that it exhibits the same convergence properties as in the unconstrained case.

We consider a problem of minimizing a d times continuously differentiable function subject to SOS-convex polynomial constraints. A polynomial $p : \mathbb{R}^n \to \mathbb{R}$ is called *SOS-convex* if the polynomial $q(x, y) := y^\top \nabla^2 p(x)\, y$ is a *sum of squares (SOS)* (see [4]). Note that any SOS-convex polynomial is convex.

Given an iterate x, we solve the SDP proposed in [1] to construct an SOS-convex approximation of the dth-order Taylor expansion of the objective function at the iterate x. We then define the next iterate x^+ as the minimizer of this polynomial subject to the constraints. As in the unconstrained case, the next iterate x^+ can

be found by solving another SDP (see, e.g., [5, Theorem 3.3]). We note that the time per iteration is polynomial in the number of variables and constraints.

We assume that the objective function is strongly convex and that its dth derivative is Lipschitz continuous. Then, our algorithm is well defined in the sense that, for any iterate x, the next iterate x^+ exists and is unique. Our main result is that our method converges locally with order d to an optimal solution, as was already proven in [1] in the unconstrained case.

If we assume, additionally, that the gradients of the active constraints at the optimal solution are linearly independent—so that the *Karush–Kuhn–Tucker (KKT) conditions* are satisfied—and that strict complementary slackness is satisfied, then the set of active constraints at the optimal solution can be locally identified in one iteration (see, e.g., [9, Theorem 2.1]).

References

[1] A. A. Ahmadi, A. Chaudhry, and J. Zhang. Higher-order Newton methods with polynomial work per iteration. *Advances in Mathematics*, 452, 109808, 2024.

[2] C. Cartis and W. Zhu. Global convergence of high-order regularization methods with sums-of-squares Taylor models. *arXiv preprint arXiv:2404.03035*, 2024.

[3] N. Doikov and Y. Nesterov. Local convergence of tensor methods. *Mathematical Programming*, 193(1):315–336, 2022.

[4] J. W. Helton and J. Nie. Semidefinite representation of convex sets. *Mathematical Programming*, 122:21–64, 2010.

[5] J. B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM Journal on Optimization*, 19:1995–2014, 2009.

[6] Y. Nesterov. Implementable tensor methods in unconstrained convex optimization. *Mathematical Programming*, 186(1):157–183, 2021.

[7] Y. Nesterov and B. T. Polyak. Cubic regularization of Newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006.

[8] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000.

[9] S. M. Robinson. Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear-programming algorithms. *gramming*, 7(1):1–16, 1974.

[10] O. Silina and J. Zhang. An unregularized third order Newton method. *arXiv preprint arXiv:2209.10051*, 2022.

[11] L. Vandenberghe and S. Boyd. Semidefnite programming. *SIAM Review*, 38(1):49–95, 1996.

Exploiting Term Sparsity and Symmetries in Polynomial Optimization Problems

Johannes Schmucker

(joint work with Daniel Brosch)

Polynomial optimization problems (POPs) arise naturally in many areas of mathematics, engineering, control theory, combinatorics, quantum information, etc. A general POP can be written in the form

$$f^* = \inf_{x \in K} f(x),$$

where $f \in \mathbb{R}[x_1, \ldots, x_n]$ is a multivariate polynomial and the feasible set is given by

$$K = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\},$$

where $g_1, \ldots, g_m \in \mathbb{R}[x_1, \ldots, x_n]$.

In general, POPs are very difficult to solve and even approximating f^* is typically computationally challenging. Over the past three decades, major progress has nevertheless been achieved through convex relaxations based on *sums of squares* (SOS) on the primal and *sequences of moments* on the dual side. In particular, the *moment-SOS hierarchy* introduced by Lasserre [1] and Parillo [2] provides a systematic sequence of *semidefinite programming* (SDP) relaxations converging from below to the global optimum f^* under relatively mild assumptions.

While this hierarchy has been highly successful both theoretically and practically, the size of the resulting SDPs grows rapidly with the number of variables and/or the polynomial degree. For a relaxation order t, the corresponding SDP on the moment side involves moment and localizing matrices indexed by monomials up to degree t. For instance, the size of the moment matrix is $\binom{n+t}{t}$, and solving the SDP becomes prohibitive quickly even for moderately sized instances.

To address this issue, several approaches have been developed that exploit additional structure present in the polynomials f and $g_1, \ldots, g_m$. Two of the most important forms of structure are *sparsity* and *symmetry*.

Symmetry has long been recognized as a powerful tool for reducing computational costs in polynomial and semidefinite optimization [3, 4, 5]. Indeed, if a POP is invariant under the action of a finite group G, then the corresponding SDP relaxations inherit this symmetry. Using techniques from representation theory, the resulting matrices can be block-diagonalized into smaller blocks, leading to substantial computational savings.

Sparsity can appear in several different forms. One approach for the unconstrained case is based on the *Newton polytope* of a polynomial that can be used to restrict the monomials appearing in SOS decompositions, thereby reducing the size of the corresponding semidefinite programs [6]. Another notion of sparsity is *correlative sparsity* [7, 8]. In this framework, one studies the interaction graph of variables that are linked with each other. If only small subsets of variables interact, then the original large SDP can be decomposed into several smaller blocks indexed by maximal cliques of the graph.

More recently, Wang, Magron, and Lasserre introduced a hierarchy based on *term sparsity* [9, 10]. Rather than focusing on interactions between variables, term sparsity is particularly well suited for sparse polynomials, that is, polynomials whose support contains only few monomials. The main idea is to construct graphs associated with the moment matrix and the localizing matrices, where the vertices correspond to the indexing monomials and edges encode multiplicative relations between them to capture the sparsity pattern.

In this talk, we introduce a symmetric variant of the term sparsity hierarchy. Our aim is to simultaneously exploit both sparsity and symmetry in order to further reduce the complexity of moment-SOS relaxations. A key challenge is that

the original term sparsity hierarchy is not symmetry preserving, in the sense that chordal extensions do not, in general, preserve the automorphism group of the underlying term sparsity pattern graphs. To this end, we study structural properties of G-invariant chordal graphs and use these results to construct symmetry-preserving chordal extensions. For the case of symmetric, i.e., S_n-invariant, polynomials, we observe that the associated sparsity pattern graphs stabilize with respect to chordality as the number of variables n increases.

First computational results for symmetric polynomial optimization problems indicate significant reductions in block sizes and running times for the resulting SDPs when compared to approaches exploiting symmetry alone. At the same time, our experiments also illustrate the interesting interplay between sparsity and symmetry. In classical SOS hierarchies, the SDP relaxations typically become larger at higher relaxation orders in order to improve the bounds. In contrast, in the symmetry-adapted term sparsity hierarchy it may happen that higher levels of the hierarchy lead to smaller SDPs, due to the appearance of additional symmetries despite having less sparsity.

References

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

[2] P. A. Parillo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, PhD thesis, California Institute of Technology, 2000.

[3] K. Gatermann and P. A. Parillo, *Symmetry groups, semidefinite programs, and sums of squares*, Journal of Pure and Applied Algebra **192** (2002), 95–128.

[4] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre, *Exploiting symmetries in SDP-relaxations for polynomial optimization*, Mathematics of Operations Research **38** (2013), 122–141.

[5] F. Vallentin, *Symmetry in semidefinite programs*, Linear Algebra and its Applications **430** (2009), 360–369.

[6] B. Reznick, *Extremal PSD forms with few terms*, Duke Mathematical Journal **45** (1978), 363–374.

[7] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, *Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity*, SIAM Journal on Optimization **17** (2006), 218–242.

[8] J. B. Lasserre, *Convergent SDP-relaxations in polynomial optimization with sparsity*, SIAM Journal on Optimization **17** (2006), 822–843.

[9] J. Wang, V. Magron, and J. B. Lasserre, *TSSOS: A moment-SOS hierarchy that exploits term sparsity*, SIAM Journal on Optimization **31** (2019), 30–58.

[10] J. Wang, V. Magron, and J. B. Lasserre, *Chordal-TSSOS: A moment-SOS hierarchy that exploits term sparsity with chordal extension*, SIAM Journal on Optimization **31** (2020), 114–141.

Semidefinite hierarchies for diagonal unitary invariant quantum states

Jonas Britz
(joint work with Monique Laurent)

Quantum entanglement is a fundamental resource in quantum information theory, yet deciding whether a given bipartite quantum state is separable or entangled is computationally intractable in general. In this talk, we consider the Doherty–Parrilo–Spedalieri (DPS) hierarchy of semidefinite approximations for the separable cone in two structured settings: for states with diagonal unitary invariance, and for states with Bose symmetry.

For diagonal unitary invariant states, we show that the hierarchy can be block diagonalized, which leads to a substantially more efficient implementation. For Bose symmetric states, we characterize the dual hierarchy in terms of sums of squares of Hermitian complex polynomials, extending a known result from the unstructured case. In both settings, the completely positive cone CP, its dual COP, and their sums-of-squares-based approximations play a central role.

References

[1] J. Britz and M. Laurent, *Semidefinite hierarchies for diagonal unitary invariant bipartite quantum states*, arXiv:2512.06551 [math.OC], 2026.

Conic programming in discrete geometry

Fernando Mário de Oliveira Filho

Consider the following problems:

(P1) *Binary codes.* Given integers n, $d \geq 1$, what is the maximum cardinality $A(n,d)$ of a subset of $H_n = \{0,1\}^n$ such that the Hamming distance between any two distinct points in the set is at least d?

(P2) *Spherical codes.* Given an integer $n \geq 2$ and $\theta \in (0,\pi]$, what is the maximum cardinality $A(n,\theta)$ of a subset of the unit sphere $S^{n-1} = \{\, x \in \mathbb{R}^n : \|x\| = 1 \,\}$ such that the angular distance between any two distinct points in the set is at least θ?

(P3) *Witsenhausen's problem.* What is the maximum measure of a subset of S^{n-1} that does not contain pairs of orthogonal points?

(P4) *Sphere-packing problem.* What is the maximum density of a *sphere packing* in $\mathbb{R}^n$, that is, what is the maximum density of an arrangement of interior-disjoint unit balls in $\mathbb{R}^n$?

An *independent set* in a graph is a set of vertices not containing an edge; the *independence number* of a graph is the maximum cardinality of an independent set of the graph. Computing the independence number of a graph is a well-known NP-hard problem. All four problems above can be modeled as finding the independence number of a graph $G = (V,E)$:

(i) For (P1), take $V = H_n$ and say two distinct vertices are adjacent if the Hamming distance between them is less than d. Then $A(n, d)$ is the independence number of G.
(ii) For (P2), let $V = S^{n-1}$ and say two distinct vertices are adjacent if the angular distance between them is less than θ. Again, $A(n, \theta)$ is the independence number of G. Note that G is now an infinite graph.
(iii) For (P3), let $V = S^{n-1}$ and say two vertices are adjacent if they are orthogonal. Independent sets of G are now sets without pairs of orthogonal points. The independence number of G is now infinite; we work instead with the *measurable independence number*, namely the maximum measure of a measurable independent set.
(iv) For (P4), let $V = \mathbb{R}^n$ and say two distinct vertices x, $y \in \mathbb{R}^n$ are adjacent if $\|x - y\| < 2$. Independent sets of G correspond to the centers of the balls in a sphere packing. The independence number is again infinite, and we use the density of the associated sphere packing instead.

The linear programming bound of Delsarte [5] gives a linear program whose optimal value is an upper bound for the parameter $A(n, d)$ of (P1). Related is the linear programming bound of Delsarte, Goethals, and Seidel [6], which gives an upper bound for the parameter $A(n, \theta)$ of (P2), now via an infinite-dimensional linear program. An upper bound for the maximum in (P3) is provided by another infinite-dimensional linear program introduced by Bachoc, Nebe, Oliveira, and Vallentin [1], and the linear programming bound of Cohn and Elkies [3] gives an upper bound for the sphere-packing density, which was used by Viazovska et al. [4, 13] to determine the sphere-packing density in dimensions 8 and 24.

Moving beyond linear programming, Schrijver [12] used semidefinite programming to improve on Delsarte's bound for binary codes. Bachoc and Vallentin [2] developed the analogue semidefinite programming bound for spherical codes.

The *theta number*, a graph parameter introduced by Lovász [9] as the optimal value of a semidefinite programming problem, provides an upper bound for the independence number that can be computed in polynomial time. McEliece, Rodemich, and Rumsey [10] and Schrijver [11] showed that Delsarte's bound for binary codes is a simple strenghtening of the Lovász theta number. The same connection was made for the linear programming bound for spherical codes by Bachoc, Nebe, Oliveira, and Vallentin [1], by extending the Lovász theta number to the infinite graph described in (ii) above.

This connection shed new light into several classical bounds in discrete geometry. It led to the development of a systematic understanding of how bounds for geometrical problems can be set up, for instance by placing both the linear programming bound of Delsarte, Goethals, and Seidel and the semidefinite programming bound of Bachoc and Vallentin as steps in a hierarchy of stronger and stronger bounds, namely the Lasserre hierarchy for the independence number of a graph [8, 7]. This opened the way to many improved bounds and new results in geometry, allowing many tools from conic programming to be used for obtaining results in discrete geometry.

In this talk, I give an introduction to bounds for geometrical problems like (P1)–(P4), showing how the classical bounds mentioned above are derived from extensions of the Lovász theta number. I discuss how conic programming, for instance in the form copositive programming, polynomial optimization, or hierarchies like the Lasserre hierarchy, can be used to derive improved bounds for these problems, giving a short overview of the main results in the field.

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) 645–661.

[2] C. Bachoc and F. Vallentin, New upper bounds for kissing numbers from semidefinite programming, *J. Amer. Math. Soc.* 21 (2008) 909–924.

[3] H. Cohn and N. Elkies, New upper bounds on sphere packings. I, *Ann. of Math. (2)* 157 (2003) 689–714.

[4] H. Cohn, A. Kumar, S.D. Miller, D. Radchenko, and M.S. Viazovska, The sphere packing problem in dimension 24, *Ann. of Math. (2)* 185 (2017) 1017–1033.

[5] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Research Reports Supplements 1973 No. 10, Philips Research Laboratories, Eindhoven, 1973.

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

[7] D. de Laat, F.C. Machado, F.M. de Oliveira Filho, and F. Vallentin, k-point semidefinite programming bounds for equiangular lines, *Math. Program.* 194 (2022) 533–567.

[8] D. de Laat and F. Vallentin, A semidefinite programming hierarchy for packing problems in discrete geometry, *Math. Program.* 151 (2015) 529–553.

[9] L. Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory* 25 (1979) 1–7.

[10] R.J. McEliece, E.R. Rodemich, and H.C. Rumsey, The Lovász bound and some generalizations, *J. Combin. Inform. System Sci.* 3 (1978) 134–152.

[11] A. Schrijver, A comparison of the Delsarte and Lovász bounds, *IEEE Trans. Inform. Theory* 25 (1979) 425–429.

[12] A. Schrijver, New code upper bounds from the Terwilliger algebra and semidefinite programming, *IEEE Trans. Inform. Theory* 51 (2005) 2859–2866.

[13] M.S. Viazovska, The sphere packing problem in dimension 8, *Ann. of Math. (2)* 185 (2017) 991–1015.

Optimality and uniqueness of the D_4 root system

David de Laat

The kissing number problem asks for the largest number $k(n)$ of nonoverlapping unit spheres in $\mathbb{R}^n$ that can simultaneously touch a common unit sphere.

In two dimensions, the kissing number is six, and the optimal configuration is unique up to isometry. In dimension three, the kissing number is 12, but optimal configurations are far from unique. A rigid optimal configuration is obtained from the FCC lattice, while the vertices of the icosahedron provide a completely nonrigid configuration. In dimension four, the vertices of the D_4 root system yield a highly symmetric configuration of 24 points on the unit sphere S^3 with pairwise inner products in -1, $-\frac{1}{2}$, 0, and $\frac{1}{2}$. This gives a kissing configuration of size 24. Musin

showed in 2003 (published in 2008) that the kissing number is at most 24, proving that the configuration coming from the D_4 root system is optimal [6].

In this talk I discuss our proof that the D_4 root system is, up to isometry, the unique optimal kissing configuration in dimension four [1]. In particular, this implies that the D_4 root system is an optimal spherical code: it minimizes

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

over all sets C consisting of 24 points in the unit sphere S^3. This result also has consequences for universal optimality. Cohn, Conway, Elkies, and Kumar showed that the D_4 root system is not universally optimal and conjectured that no universally optimal configuration of 24 points in S^3 exists [5]. Since every optimal spherical code is isometric to D_4, this conjecture now follows immediately.

The classical Delsarte–Goethals–Seidel linear programming bound is a two-point bound obtained from positive definite kernels on the sphere [7]. It gives sharp upper bounds for the kissing number problem in dimensions 8 and 24, meaning that the upper bound coincides with the size of an actual kissing configuration. By using complementary slackness, this sharpness can be exploited to prove uniqueness of the optimal configurations in these dimensions [4, 3].

In 2008, Bachoc and Vallentin introduced the three-point bound on the kissing number [2]. It is strong enough to provide alternative proofs of $k(3) = 12$ and $k(4) = 24$ and gives improved upper bounds in many other dimensions. However, the three-point bound appears not to be sharp in dimension four and therefore cannot be used to prove uniqueness.

Instead, we work with the generalized Lasserre hierarchy for packing problems in discrete geometry [8]. In this framework one studies positive definite kernels indexed by independent sets of size at most t, where t is the level of the hierarchy. The first level reduces to the two-point bound, so we will use the second level. Harmonic analysis for the orthogonal group reduces the problem to a semidefinite program involving positive semidefinite matrix variables. The resulting optimization problem still contains infinitely many polynomial inequalities in the pairwise inner products of four points, but these constraints can be represented using sums-of-squares techniques.

After truncation, we obtain a large-scale semidefinite program. Numerical solutions are computed to high precision using a specialized interior-point solver for low-rank semidefinite programs [9]. With this solver, the semidefinite program can be solved in two weeks on 8 cores. Since floating-point solutions are neither exactly feasible nor exactly optimal, we subsequently apply a rounding procedure that reconstructs exact algebraic solutions from the numerical data. Rounding the numerical solution to an exact solution takes 4 hours. The resulting certificates can then be verified rigorously and independently of the numerical solver or rounding procedure.

This approach yields a sharp bound, which means we can now use complementary slackness. Using this it follows that any optimal 24-point kissing configuration in dimension four has only the inner products -1, $-\frac{1}{2}$, 0, and $\frac{1}{2}$. We use this to

show that any optimal kissing configuration is a root system, and using the classification of irreducible root systems, it follows that the only possibility is the D_4 root system itself.

References

[1] D. de Laat, N. M. Leijenhorst, and W. H. H. de Muinck Keizer, *Optimality and uniqueness of the D_4 root system*, preprint, 2024, arXiv:2404.18794.

[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.

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

[4] A. M. Odlyzko and N. J. A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Combin. Theory Ser. A **26** (1979), no. 2, 210–214.

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

[6] O. R. Musin, *The kissing number in four dimensions*, Ann. of Math. (2) **168** (2008), no. 1, 1–32.

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

[8] D. de Laat and F. Vallentin, *A semidefinite programming hierarchy for packing problems in discrete geometry*, Math. Program. **151** (2015), no. 2, 529–553.

[9] N. M. Leijenhorst and D. de Laat, *Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput. **16** (2024), 503–534.

Linear programming bounds in geometric problems

Máté Matolcsi

In my talk, I presented two different linear programming bounds that arise naturally in geometric optimization problems. The first is the well-known classical "Delsarte LP bound" (which has several other names in the literature), while the second is a novel approach introduced in [1] and [7], called the "geometric fractional chromatic number".

The Delsarte LP-bound is based on the fact that an autocorrerlation function of the form $1_A * 1_{-A}$ is positive definite (or, positive type, in another terminology). The first application of this bound was for bounding the size of binary codes [4], but I presented a more recent application in the problem of sphere packings in $\mathbb{R}^d$. Namely, assume for simplicity that the centers of the spheres must lie on an ε-grid, and the packing is periodic in all coordinate directions with some large period L. Let A denote the set of the centers of the spheres in $[0, L]^d$. The normalized autocorrelation function $f(x) = \frac{1}{|A|} 1_A * 1_{-A}$ has the following properties: $f(0) = 1$, $\sum f(x) = |A|$, $f|_{B(0,2)^c} = 0$, $f \geq 0$ and $\hat{f} \geq 0$. This leads to the primal linear program of maximizing $\sum f(x)$ subject to all the other conditions. The Delsarte LP bound arises as the dual of this LP, and for the problem of sphere packings it was first worked out in detail in [2]. Here we give a formulation for any body $K \subset \mathbb{R}^d$, presented in [5]:

Theorem 1. *Let $h : \mathbb{R}^d \to \mathbb{R}$ be continuous such that $h(0) = 1$, $h|_{(K-K)^c} \leq 0$, and $\hat{h} \geq 0$. Then $1/\int_{\mathbb{R}^d} h$ gives an upper bound on the density of translational packings by K.*

For example, the function $h = \frac{1}{|K|} 1_K * 1_{-K}$ is always admissible, so we get the trivial density bound $1/|K|$ (volume packing bound).

If K tiles $\mathbb{R}^d$ then of course the maximal density of a packing is $1/|K|$, and the LP bound above is tight. However, the following theorem in [5] shows that in any non-trivial situation, the LP bound is strictly better than the volume packing bound.

Theorem 2. *If K is a convex body which does not tile $\mathbb{R}^d$, then the Delsarte LP bound is always strictly better than the trivial bound $1/|K|$.*

The proof is based on the fact that extremal solutions of the primal LP and the dual LP necessarily weakly-tile with each other, and a convex body can only tile the space weakly if it tiles properly – a non-trivial result proved earlier in [6].

In the second part of the talk I presented the method of the "geometric fractional chromatic number" (GFCN), introduced in [1] and [7]. I explained the method in the context of 1-avoiding sets in the plane.

A set $A \subset \mathbb{R}^2$ is called 1-avoiding if $\|a - a'\| \neq 1$ for any $A, a' \in A$. What is the maximal possible upper density of a measurable 1-avoiding set A? It was conjectured by Erdős that the density is $< 1/4$. This conjecture was proven recently in [1] by a combination of the Delsarte LP and the GFCN LP.

Assume a 1-avoiding set A of density $\delta(A)$ is given, and let X be any finite set of points in the plane. We think of X as a "probe". We drop X randomly on the plane, and observe the intersection of it with A. For any $Y \subset X$, let $\gamma_{A,X}(Y)$ denote the probability that we observe the pattern Y as the intersection. If X has n points, we have 2^n patterns Y. We can then think of $\gamma_{A,X}(Y)$ as variables, and write up linear constraints for them. Namely, $\gamma_{A,X}(Y) \geq 0$, $\sum_Y \gamma_{A,X}(Y) = 1$, $\gamma_{A,X}(Y) = 0$ whenever Y contains two points at distance 1, and $\sum_{x \in Y} \gamma_{A,X}(Y) = \delta(A)$ for any fixed $x \in X$. Importantly, yet another constraint is given by the underlying geometric invariance of the problem: if $Y \equiv Z$ are geometrically congruent to each other (i.e. there exists a Euclidean transformation T such that $T(Y) = Z$), then $\sum_{Y \subset I} \gamma_{A,X}(I) = \sum_{Z \subset J} \gamma_{A,X}(J)$.

This leads to the linear program (GFCN) of maximizing $\sum_{x \in Y} \gamma_{A,X}(Y)$ subject to all the other conditions. Any finite configuration X leads to an upper bound on $\delta(A)$, and we can try to find sets X which give strong bounds. By a sophisticated computer search we have identified a set X of 27 points such that the GFCN LP testifies $\delta(A) \leq 1/4$, nearly proving Erdős' conjecture. However, we were not able to go below the threshold $1/4$ by using the GFCN method only.

The breakthrough $\delta(A) \leq 0.247$ was achieved in [1] by combining the Delsarte LP with the GFCN LP. This combination is made possible by the fact that for any

distance d appearing in X, the values $f(d)$ of the symmetrized autocorrelation function can be simultaneously expressed as a linear combination of the variables $\gamma_{A,X}(Y)$ and as a Fourier expansion involving Bessel functions.

The same appraoch was later carried out in [3] to yield the currently known best estimate $\delta(A) \leq 0.2953$ for the density of sets A in the unit sphere such that A avoids orthogonality, i.e. $\langle a, a' \rangle \neq 0$ for any $a, a \in A$.

References

[1] G. Ambrus, A. Csiszárik, M. Matolcsi, D. Varga, P. Zsámboki, *The density of planar sets avoiding unit distances*, Mathematical Programming **207** no. 1–2, (2024), 303–327.

[2] H. Cohn, N. Elkies, *New upper bounds on sphere packings. I.*, Ann. of Math. (2) **157** (2003), no. 2, 689–714.

[3] D. Czifra, Á. Dúcz, M. Matolcsi, D. Varga, P. Zsámboki, *Improved bounds for the double cap conjecture*, preprint `https://arxiv.org/abs/2605.28709`

[4] P. Delsarte, *Bounds for unrestricted codes, by linear programming*, Philips Res. Rep. **2** (1972), 272–289.

[5] M. Kolountzakis, N. Lev, M. Matolcsi, *The Turán and Delsarte problems and their duals*, preprint, `https://arxiv.org/abs/2510.10172`

[6] N. Lev, M. Matolcsi, *The Fuglede conjecture for convex domains is true in all dimensions*, Acta Math. **228** (2022), 385–420.

[7] M. Matolcsi, I. Z. Ruzsa, D. Varga, P. Zsámboki, *The fractional chromatic number of the plane is at least 4*, preprint, `https://arxiv.org/abs/2311.10069`

Application of Optimization in Constructive Approximation: Some Personal Examples

Simon Foucart

In this talk, modern optimization techniques are advocated as computational tools suited to attack several extremal problems from Approximation Theory. Three such problems, which had reached their limitations based on purely analytical approaches, are showcased.

First is the problem of minimal projections, which involves minimization over measures and exploits the moment method. Given a subspace V of a Banach space X, it consists in finding a projection from X onto V—a linear map P from X onto V such that $P(v) = v$ for all $v \in V$—that minimizes the operator norm $\|P\| = \sup\{\|P(x)\| : x \in X, \|x\| \leq 1\}$; see [6] for more details. With X being a space of continuous functions, the problem has an elegant analytical solution when V is the space of trigonometric polynomials of degree $\leq n$, say: it is the Fourier projection. But analytic solutions seem unattainable when V is the space of algebraic polynomials of degree $\leq n$, as strongly suggested by the case $n = 2$. Instead, we reveal how the problem can be solved computationally, as proposed in [5]. There, the line of attack boiled down to producing accurate upper bounds and lower bounds, the latter being the most challenging ones. They were obtained by transforming the optimization variables from signed Borel measures to sequences of trigonometric moments by relying on the discrete trigonometric moment problem.

Second is the problem of best approximation, which involves minimization over polynomials and exploits the sum-of-squares method. As a representative example, given a monomial $m_k(x) = x_1^{k_1} \cdots x_d^{k_d}$ of degree $n = |k_1| + \cdots + |k_d|$, we consider the problem of finding a best approximant to m_k by polynomials $p \in \mathcal{P}_{n-1}^d$ of degree $\leq n-1$ in d variables with respect to the L_∞-norm on some $\Omega \subseteq \mathbb{R}^d$. Except when $\Omega = [-1,1]^d$, this best approximation problem had only been solved analytically with small n and d for the ball and the simplex. In [1], we proposed a computational approach based on the program $\min\{c : c \pm (m_k - p) > 0 \text{ on } \Omega\}$ with optimization variables $p \in \mathcal{P}_{n-1}^d$ and $c \in \mathbb{R}$, Ω being a generic semialgebraic compact subset of $\mathbb{R}^d$. We obtained a sequence of upper bounds by relaxing the positivity constraints by sum-of-square constraints. Not only do these upper bounds converge to the sought-after value, but they do so in a finite number of steps, owing to the dual of the best approximation program—a maximization over measures—admitting an optimal solution which is atomic. It coincides with a notion called signature in Approximation Theory, actually preferred over best approximants themselves, since best approximants are typically not unique when $d > 1$ but they share a common signature. In the follow-up work [2], we further exploited our computational framework to discover—and then prove—the form of another multivariate generalization of Chebyshev polynomials, this time not understood as a specific monomial m_k minus its best approximant, but as a polynomial in $\mathcal{P}_n^d$ with smallest L_∞-norm under the constraint that its degree-n coefficients sum up to one.

Third is the problem of optimal recovery, which is quite relevant in Data Science. The goal here is to *learn* an object f living in a Banach space F—prototypically, f is a function—based on *a priori* information in the form $f \in \mathcal{K}$ for some model set $\mathcal{K} \subseteq F$ and on *a posteriori* information in the form $y = \Lambda(f)$ for some linear observation map $\Lambda : F \to \mathbb{R}^m$. For the specific task of learning $\Gamma(f)$, be it full approximation ($\Gamma = \mathrm{Id}$ is the identity) or prediction ($\Gamma = \gamma$ is a functional), the performance of a recovery procedure view as $\Delta : \mathbb{R}^m \to \mathrm{range}(\Gamma)$ is assessed in a worst-case setting through $\mathrm{wce}(\Delta) = \sup\{\|\Gamma(f) - \Delta(\Lambda(f))| : f \in \mathcal{K}\}$. Our aim is to find optimal recovery maps, i.e., Δ that minimize $\mathrm{wce}(\Delta)$. To this end, it would be helpful to know, say in the specific case of functionals $\gamma : F \to \mathbb{R}$, whether the fact that γ has a simple structure implies a similarly simple structure for an optimal recovery map $\delta : \mathbb{R}^m \to \mathbb{R}$. It had been known for a long time that this is the case for linear functionals γ and convex and symmetric model sets $\mathcal{K}$: one can find a linear functional $\delta^{\mathrm{opti}} : \mathbb{R}^m \to \mathbb{R}$ such that $\mathrm{wce}(\delta^{\mathrm{opti}}) \leq \mathrm{wce}(\delta)$ for all $\delta : \mathbb{R}^m \to \mathbb{R}$. This fundamental result has recently been extended to the situation (among others) where $\gamma(f)$ is the ℓth largest among N linear functionals. In [3], it was shown that there is an optimal recovery map of the form $\delta^{\mathrm{opti}}(y) = \sup_{|I|=\ell} \inf_{i \in I} \left(c_0^{(I,i)} + \sum_{k=1}^m c_k^{(I,i)} y_k\right)$, where I stands for a subset of $\{1, \ldots, N\}$. The main theoretical tool was a (seemingly novel) refinement of the Hahn–Banach dominated extension theorem. The fundamental result has been extended in another direction, too. Namely, in the scenario where F is a Hilbert space, it was shown in [4] that if $\gamma(f) = \langle A(f), f\rangle$ is a quadratic form and if

$\mathcal{K} = \{f \in H : \|T(f)\| \leq 1\}$ is an hyperellipse, then there is an optimal recovery map of the form $\delta^{\text{opti}}(y) = \langle By, y\rangle + \beta$. The main theoretical tool was a refinement of Sion minimax theorem. In both of these extensions, the existence results were complemented by computational constructions of the optimal recovery maps relying on various techniques from Optimization Theory.

It is noteworthy that, as a common thread to all three problems showcased above, one typically ends up having to solve semidefinite programs.

References

[1] M. Dressler, S. Foucart, M. Joldes, E. de Klerk, J. B. Lasserre, and Y. Xu, *Optimization-aided construction of multivariate Chebyshev polynomials*, Journal of Approximation Theory **305** (2025), 106116.

[2] M. Dressler, S. Foucart, M. Joldes, E. de Klerk, J. B. Lasserre, and Y. Xu, *Least multivariate Chebyshev polynomials on diagonally determined domains*, SIAM Journal on Applied Algebra and Geometry. To appear.

[3] S. Foucart, *Optimal algorithms for nonlinear estimation with convex models*, Forum of Mathematics, Sigma. In revision.

[4] S. Foucart, *Worst-case-optimal estimation of a quadratic form by quadratic functionals of its linear observations*. Submitted.

[5] S. Foucart and J. B. Lasserre, *Determining projection constants of univariate polynomial spaces*, Journal of Approximation Theory **235** (2018), 74–91.

[6] S. Foucart and L. Skrzypek, *Minimal projections: from classical theory to modern developments*, Surveys in Approximation Theory. In preparation.

Guess-and-Prove Methods for Multivariate Chebyshev Approximation

Mioara Joldes

(joint work with Denis Arzelier, Florent Bréhard, Jean-Bernard Lasserre)

The aim of this talk is to explain how numerical evidence, which arises quite naturally in multivariate Chebyshev approximation, can be turned into a rigorous mathematical proof. One starts from a compact semialgebraic set $\Omega \subset \mathbb{R}^m$, a function f, and a finite-dimensional approximation space V. The problem is to find an element $p \in V$ for which the uniform error $||f - p||_{\infty,\Omega}$ is as small as possible. This is a classical question, but in several variables the certification of optimality is much less transparent than in the univariate case.

In one variable, the alternation theorem gives a complete and robust optimality criterion [1]. A best approximant is characterized by an ordered set of alternating extreme errors, and this finite pattern is itself a certificate. In several variables there is no longer a canonical ordering of the extremal points, and the alternation picture must be replaced by its dual form. Optimality is then witnessed by a signed measure concentrated on the contact set of the error, with the property that it annihilates the approximation space [1, 2]. Thus the difficulty is not only to find a good candidate approximant, but to certify the whole dual configuration attached to it: the contact points, their signs, and the corresponding weights.

The approach described here is a guess-and-prove procedure. The guessing part is based on the optimization-aided framework developed in the previous work [3].

We consider best uniform approximation of monomials on the standard domains—simplexes, cross-polytopes, balls and hypercubes—by polynomials of prescribed lower degree. Then, moment and sum-of-squares relaxations are used to construct candidate multivariate Chebyshev polynomials. In practice, these relaxations reveal much more than a good polynomial: they suggest the value of the best error, the locations where the error is saturated, the signs of the saturation, and the way in which those points lie on the boundary of the domain. Starting from the predicted extremal configuration, we ask whether the finite data produced by the computation can be certified as an exact primal-dual optimality certificate.

We use Hettich's finite optimality conditions for constrained Chebyshev approximation [4]. The predicted certificate consists of contact points, signs, active constraints and positive weights. These data encode, respectively, the extremal set of the error, the error level, the faces of Ω on which stationarity is imposed, and the masses of the dual measure. In local coordinates on the corresponding faces, the equal-height conditions, the annihilation of the approximation space, and the first-order stationarity conditions form a square nonlinear system.

The first part of the proof is therefore a finite-dimensional certification of this Hettich system. Starting from the numerical configuration, we solve the associated nonlinear equations at high precision and then apply an interval Newton–Krawczyk argument [5]. This produces an interval box containing a unique exact solution of the system. Once the weights are certified to be positive, the corresponding signed measure is a genuine dual certificate. In particular, it gives a rigorous lower bound for the best approximation error, and hence proves that no element of the approximation space can have a smaller uniform error than the certified level.

The second part is global. One must prove that the candidate approximant never exceeds the certified error level on the whole domain. This step is tricky because the inequality is sharp precisely at the support points. A naive interval evaluation near such points gives no useful separation. Instead, the proof must use the local analytic structure of the error. Interior support points are treated by Taylor expansion and positivity of the relevant Hessian. Boundary points require coordinates adapted to the face and control of inward derivatives. Edges and vertices lead to lower-dimensional estimates. Once small neighborhoods of the saturated points have been handled in this way, the remaining compact set can be verified by a Bernstein-based subdivision algorithm, where the error is separated from the extremal level.

We illustrate the certification procedure on a representative simplex example, starting from the numerical signature produced by the relaxation method of [3]. The aim is to show how such numerical data can be converted into a rigorous optimality certificate. This is intended as a first step toward treating other examples considered in [3].

The example suggests a further algebraic question. Beyond numerical certification, the extremal configuration appears to have a structured algebraic nature, revealed after using the symmetries of the simplex and rewriting the contact equations in suitable parameters. It would be interesting to determine to what extent

this phenomenon is systematic for the signatures produced in [3]: more precisely, whether the numerical optimality data can often be replaced by a small system of exact algebraic equations and inequalities, whose solutions are then certified by interval methods.

The contribution should therefore be understood as an ongoing attempt to build a certification mechanism on top of the computational discovery method of [3].

References

[1] T. J. Rivlin, *An Introduction to the Approximation of Functions*, Dover, New York, 1981.

[2] I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Springer, Berlin, 1970.

[3] M. Dressler, S. Foucart, M. Joldes, E. de Klerk, J.-B. Lasserre and Y. Xu, *Optimization-aided construction of multivariate Chebyshev polynomials*, J. Approx. Theory **305** (2025), 106116.

[4] R. Hettich, *A Newton-method for nonlinear Chebyshev approximation*, Approximation Theory, Lecture Notes in Math. **556**, Springer, 1976, 222–236.

[5] S. M. Rump, *Verification methods: rigorous results using floating-point arithmetic*, Acta Numerica **19** (2010), 287–449.

Conic Approaches to Nonnegativity with SONC and SAGE

Mareike Dressler

Deciding nonnegativity of polynomials is a classical problem in real algebraic geometry and a cornerstone of nonlinear polynomial optimization. In recent years, several interrelated approaches to certifying nonnegativity of sparse polynomials and signomials (weighted sums of exponentials composed with linear functionals) have emerged.

In this talk, we introduce *sums of nonnegative circuit (SONC) polynomials*, a sparse alternative to sums of squares (SOS) certificates. A polynomial

$$f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in S} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} + d\mathbf{x}^{\boldsymbol{\beta}},$$

where $\mathbf{x}^{\boldsymbol{\alpha}} := x_1^{\boldsymbol{\alpha}_1} x_2^{\boldsymbol{\alpha}_2} \cdots x_n^{\boldsymbol{\alpha}_n}$, is called a *circuit polynomial* if $S \subseteq 2\mathbb{N}^n$ is affinely independent, $\boldsymbol{\beta} \in \mathbb{N}^n$ lies in the relative interior of $\operatorname{conv}(S)$, and $c_{\boldsymbol{\alpha}} > 0$. Their nonnegativity can be characterized explicitly via the arithmetic-geometric inequality [1]. SONC decompositions yield certificates of nonnegativity and are independent of the classical SOS approach, see e.g. [2].

Next, we discuss key geometric and structural properties of SONC polynomials, many of which sharply contrast with their SOS counterparts. In particular, the property of being SONC is not preserved under polynomial multiplication or affine transformation of variables [3]. However, a unique feature of the SONC approach is its *sparsity-preserving nature*, making it particularly well-suited for high-degree sparse polynomials, see e.g. [4].

Afterwards, we provide an overview of the resulting optimization framework. Membership in the SONC cone can be decided via *relative entropy programming*

(REP) [2]. The SONC cone and its dual also admit approaches via *geometric programming*, *second order cone programming*, *linear programming*, or *dual power cone optimization* via column generation.

Finally, we explain how these ideas extend to signomial optimization through (conditional) *sums of arithmetic-geometric exponentials (SAGE)*. In [5], we introduce signomial rings and show how this concept and a newly presented signomial Positivstellensatz lead to a novel convex relaxation hierarchy of lower bounds of signomial and polynomial optimization. We now outline the main ideas of [5]:

Let $\mathcal{A} \subseteq \mathbb{R}^n$ be a distinguished finite ground set containing the origin. To every $\boldsymbol{\alpha} \in \mathcal{A}$ we associate a "monomial" basis function $e^{\boldsymbol{\alpha}} \colon \mathbb{R}^n \to \mathbb{R}$ that takes values $e^{\boldsymbol{\alpha}}(\mathbf{x}) = \exp\langle \boldsymbol{\alpha}, \mathbf{x}\rangle$. A *signomial* supported on a finite set $A \subseteq \mathbb{R}^n$ is a real-linear combination $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha}\in A} c_{\boldsymbol{\alpha}} e^{\boldsymbol{\alpha}}(\mathbf{x})$. The *signomial ring* $\mathbb{R}[\mathcal{A}]$ is the $\mathbb{R}$-algebra generated by the basis functions $\{e^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathcal{A}}$. Signomials that are nonnegative on a convex set X and have at most one negative term are called *X-AGE* functions. A signomial that can be written as a sum of X-AGE functions is termed *X-SAGE*. We prove the following Positivstellensatz, which is the first signomial Positivstellensatz to leverage conditional SAGE in the presence of nonconvex constraints and to permit irrational exponents.

Theorem 1 (Positivstellensatz for X-SAGE)**.** *Consider a compact convex set X, signomials $g_1, \ldots, g_m \in \mathbb{R}[\mathcal{A}]$, and $K = \{\mathbf{x} \in X \,:\, g_i(\mathbf{x}) \geq 0 \;\; \text{for all } i = 1, \ldots, m\}$. If $f \in \mathbb{R}[\mathcal{A}]$ is **positive** on K, then there is an identity*

$$\left(\textstyle\sum_{\boldsymbol{\alpha}\in\mathcal{A}} e^{\boldsymbol{\alpha}}\right)^r f \;=\; \lambda_0 + \textstyle\sum_{i=1}^m \lambda_i \cdot g_i,$$

for X-SAGE $\lambda_0 \in \mathbb{R}[\mathcal{A}]$, posynomials $\lambda_i \in \mathbb{R}[\mathcal{A}]$, and $r \in \mathbb{N}$.

Note that the representation involves an explicit identity which is affine in f and the "unknown" signomials λ_0, λ_i. Notably, neither X nor K (nor their images under exponential maps) is required to be semialgebraic.

To leverage this representation theorem algorithmically, we introduce an artificial notion of $\mathcal{A}$-degree for signomials, which, unlike polynomials, possess no intrinsic degree. This yields a complete hierarchy of REP-computable lower bounds for signomial optimization. To our knowledge, this is the first completeness result for minimizing arbitrary signomials over compact convex sets subject to finitely many signomial inequalities. It is also the first completeness result for a hierarchy based on conditional SAGE in the presence of nonconvex constraints. To conclude, we provide numerical examples illustrating the hierarchy and its performance on a reaction network problem.

There is no shortage of intriguing open questions. To name just a few:

- How to choose the "best" signomial ring $\mathbb{R}[\mathcal{A}]$ or ground set $\mathcal{A}$?
- Quantitative complexity and degree bounds for the Positivstellensatz.
- Detailed study of the signomial moment theory and upper bounds, as initiated in [5].

References

[1] S. Iliman and T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, Res. Math. Sci., 3:1–35, 2016.

[2] M. Dressler, S. Iliman and T. de Wolff, *A Positivstellensatz for Sums of Nonnegative Circuit Polynomials*, SIAM J. Appl. Algebra Geom., 1(1):536–555, 2017.

[3] M. Dressler, A. Kurpisz and T. de Wolff, *Optimization over the Boolean Hypercube via Sums of Nonnegative Circuit Polynomials*, Found. Comput. Math., 2(22):365-387, (2022).

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

[5] M. Dressler and R. Murray, *Algebraic Perspectives on Signomial Optimization*, SIAM J. Appl. Algebra Geom., 6 (2022), no. 4, pp. 650–684.

Polarization of lattices: cold spots, spherical designs and linear programming

Philippe Moustrou

(joint work with C. Bachoc, F. Vallentin, M. C. Zimmermann)

Let $L \subseteq \mathbb{R}^n$ be an n-dimensional lattice in Euclidean space, and let $f\colon [0,\infty) \to \mathbb{R}$ be a nonnegative function. The *polarization* of the lattice L with respect to the potential function f is defined by

$$\mathcal{P}(f,L) = \min_{z\in\mathbb{R}^n} p(f,L,z) \ \text{ with } \ p(f,L,z) = \sum_{x\in L} f(\|x-z\|^2).$$

The *max-min polarization problem* for lattices then consists in maximizing $\mathcal{P}(f,L)$ over n-dimensional lattices L of a given point density.

As the sphere covering problem is the inhomogeneous counterpart of the sphere packing problem, the max-min polarization problem can similarly be viewed as the inhomogeneous variant of the *potential energy minimization problem* which has been extensively studied [6]. Due to its bilevel structure, the max-min polarization problem is significantly more difficult than the potential energy minimization problem. Only very few exact results are known.

We restrict our attention to the *Gaussian core model*, where the potential functions are of the form $f_\alpha(r) = e^{-\alpha r}$, so that

$$p(f_\alpha, L, z) = \sum_{x\in L} e^{-\alpha\|x-z\|^2}.$$

This choice of potential functions is natural due to a theorem by Bernstein, see [7, Theorem 9.16], which states that every completely monotonic function $f : (0,\infty) \to \mathbb{R}$, that is, f is C^∞ and $(-1)^k f^{(k)} \geq 0$ for all $k \geq 0$, can be expressed as

$$f(r) = \int_0^\infty e^{-\alpha r}\, d\mu(\alpha),$$

for some measure μ on $[0,\infty)$. In particular, when a lattice is optimal for the max-min polarization problem for every f_α, with $\alpha > 0$, then the lattice is called *universally optimal for polarization*.

Recently, the max-min polarization problem for two-dimensional lattices was solved. Bétermin, Faulhuber, and Steinerberger [3] proved that the hexagonal lattice—the A_2 root lattice—is universally optimal for polarization among all two-dimensional lattices with the same point density.

To solve this bilevel optimization, a thorough understanding of the inner minimization problem is required. In the two-dimensional case, Baernstein II [2] analyzed the situation when $L = A_2$ is the hexagonal lattice. For every positive α, he showed that $z \mapsto p(f_\alpha, L, z)$ attains its minimum at a deep hole c of the lattice. A *deep hole* of a lattice is a point c that maximizes the distance to the nearest lattice points: The point c satisfies

$$\|x - c\| = \max_{z \in \mathbb{R}^n} \min_{x \in L} \|x - z\|.$$

In his proof, Baernstein II used the maximum principle for the heat equation. He considers $p(f_\alpha, L, z)$ as the solution of the heat equation for the point z, where unit heat sources are located at the lattice points. Then for fixed α, which represents the reciprocal of time, minimizing $p(f_\alpha, L, z)$ amounts to finding the temperature at the coolest points in the plane. Inspired by this physical interpretation, we propose to call the minimizers of $z \mapsto p(f_\alpha, L, z)$ *cold spots*. We say that a cold spot is *universal*, if it is a cold spot for every $\alpha > 0$, and thus for every completely monotonic potential.

Now Baernstein's result states that for the hexagonal lattice, the deep holes are universal cold spots. However, he also shows that for two-dimensional lattices, other than the hexagonal lattice, deep holes and cold spots generally differ.

Finding the cold spots is, in general, a difficult problem. Currently no computational method to exactly determine cold spots exist. While one can compute partial sums of the series $p(f_\alpha, L, z)$ and approximate local minima numerically, obtaining rigorous performance guarantees would require significant additional work.

In the limiting case of very steep potential functions, as $\alpha \to \infty$, one naturally expects that the max-min polarization converges to the lattice sphere covering problem

$$\min_{L \subseteq \mathbb{R}^n} \max_{z \in \mathbb{R}^n} \min_{x \in L} \|x - z\|,$$

where the minimization is over n-dimensional lattices of given point density.

This is indeed the case, as shown by Bétermin and Petrache [4, Theorem 1.5]: They proved that as $\alpha \to \infty$ the cold spots of L tend to deep holes. More precisely, they showed that there exists a deep hole c of L such that, for any $z \in \mathbb{R}^n$, there exists a threshold α_z such that for all $\alpha > \alpha_z$,

$$\sum_{x \in L} e^{-\alpha \|x - z\|^2} \geq \sum_{x \in L} e^{-\alpha \|x - c\|^2}.$$

In [1], we consider the following question: Under which conditions do cold spots *coincide* with deep holes, not just in the limit, but already for finite α? This is known to occur for the hexagonal lattice. We call a cold spot *stable*, if it is a cold

spot for every $\alpha \geq \alpha_0$ for some finite α_0. One expects that stable cold spots are deep holes for lattices whose deep holes are "highly symmetric".

More precisely, consider the *inhomogeneous shells* around the point z:

$$L(z,r) = \{x \in L : \|x - z\| = r\}.$$

We are interested in the case that all these nonempty, inhomogeneous shells carry spherical designs, where we say that a finite set X on the sphere of radius r with center z is a *spherical M-design* if every polynomial p of total degree at most M has the same average over the sphere as over the set X.

First, we show that a point can be critical for every α large enough only if its shells are spherical 1-design, and that in this case, these points are critical for every $\alpha > 0$. Then, we show that for lattices with a unique kind of deep hole c, when the shells around c are spherical 2-designs, then c is a stable cold spot. We furthermore prove that these conditions are satisfied by any root lattice, as well as E_6^*, E_7^*, the Coxeter-Todd lattice in dimension 12 and the Barnes-Wall lattice in dimension 16.

In our proof we combine two different bounds for inhomogeneous Gaussian lattice sums by a covering argument. The first bound is from Bétermin, Petrache [4], which is helpful to bound $p(f_\alpha, L, z)$ for points z far away from deep holes. The second bound is new and comes from an application of linear programming bounds for spherical designs. With this bound one can show that deep holes are local minimizers for $z \mapsto p(f_\alpha, L, z)$. Furthermore, it also yields very explicit, quantitative information about the inhomogeneous Gaussian lattice sum in the neighborhood of c, which is needed for the covering argument. By this we get reasonable estimates for the threshold α_0. For example, we shall show that for $L = E_8$ already $\alpha_0 = 23$ suffices and for $L = D_4$ already $\alpha_0 = 5$ suffices.

As a future direction of research, we consider the case where α goes to to 0. In that situation, the *Poisson Summation Formula* allows to rewrite the problem as the minimization problem

$$\min_{z \in \mathbb{R}^n} \sum_{y \in L^*} \cos(2\pi z \cdot y) e^{-t\|y\|^2}$$

where L^* is the dual lattice of L, and a small α in the previous formulation now corresponds to a large t. Using the geometric properties of root lattices, we expect to show that there exists t_0 such that for any $t > t_0$, the minimizers are again the deep holes. With a good estimate on t_0, combined with the previous method, one could even hope for proving that the deep holes are universal cold spots for root lattices.

References

[1] C. Bachoc, P. Moustrou, F. Vallentin, C. M. Zimmermann Polarization of lattices: Stable cold spots and spherical designs to appear in Forum Mathematics, Sigma

[2] A. Baernstein II, A minimum problem for heat kernels of flat tori, pp. 227–243 in: Extremal Riemann surfaces (San Francisco, CA, 1995), volume 201 of Contemporary Mathematics, AMS, 1997.

[3] L. Bétermin, M. Faulhuber, S. Steinerberger, A variational principle for Gaussian lattice sums, arXiv:2110.06008 [math.CA]
[4] L. Bétermin, M. Petrache, Dimension reduction techniques for the minimization of theta functions on lattices, J. Math. Phys. 58, 071902 (2017).
[5] S.V. Borodachov, D.P. Hardin, E.B. Saff, Discrete energy of rectifiable sets, Springer, 2019.
[6] H. Cohn, A. Kumar, Universally optimal distribution of points on spheres, J. Amer. Math. Soc. 20 (2007), 99–148.
[7] B. Simon, Convexity: An Analytic Viewpoint, Cambridge University Press, 2011.

Applications of the geometric fractional chromatic number to problems in geometry

Dániel Varga

For a given ambient space and a finite forbidden geometric configuration, the problems considered in this talk ask for the largest possible density of a measurable set avoiding that configuration. Following [1], we refer to such an extremal quantity as independence density, a measurable generalization of the independence ratio of a graph or hypergraph. The most well-studied case is the upper density of measurable 1-avoiding sets, that is, sets A such that $\|x-y\| \neq 1$ for any $x, y \in A$. As a different example, one may ask for the upper density of measurable subsets of the plane that contain no three points forming a unit equilateral triangle.

The geometric fractional chromatic number (GFCN), introduced in [2] in the context of 1-avoiding sets in $\mathbb{R}^2$, is a linear programming method for obtaining strong bounds on a wide range of independence densities. Here, we present several results obtained using this method, either on its own or in combination with the classical Delsarte LP method.

Elsewhere in this volume, coauthor Máté Matolcsi defines the GFCN linear program formally. Here, we limit ourselves to an informal introduction.

The GFCN method derives independence-density upper bounds by testing hypothetical independent sets against finite point configurations, which we call probes. Given an independent set A and a finite probe X, we place a congruent copy of X at random in the ambient space and record which probe points land in A. Each placement determines a subset $Y \subseteq X$ of observed probe points, thereby inducing a probability distribution on the subsets of X.

For a fixed probe X, the $2^{|X|}$ observation probabilities obey explicit linear constraints that are valid for every independent set A. First, they form a probability mass function. Second, observations containing a forbidden configuration have probability zero, since they would contradict the independence of A. Third, geometric invariance gives further constraints: any two congruent subconfigurations $Y, Z \subseteq X$ induce a linear relation among the observation probabilities. Finally, the density of A is itself a linear function of these probabilities. This yields a linear program whose objective is the density of A. Since every independent set produces a feasible point, the maximum gives an upper bound on the independence density.

The remaining challenge is to find probes whose linear programs yield strong bounds for a given independence density. In our applications, these probes were

found by large-scale computer search, using an enhanced beam search in earlier work and a variant of Monte Carlo Tree Search more recently.

For the spaces considered here, let $m_1(S)$ denote the supremum of upper densities of 1-avoiding sets in S.

Theorem 1 ([3]). *There is a 27-vertex probe G_{27} in the plane demonstrating that*

$$m_1(\mathbb{R}^2) \leq 1/4.$$

Let r be the infimum of the independence ratios of planar unit-distance graphs. In [3] we use the above G_{27} probe and exploit the amenability of the group of Euclidean transformations on $\mathbb{R}^2$ to construct an infinite family of planar unit-distance graphs with independence ratios arbitrarily close to $1/4$. This appears to be the first improvement on the classical bound $r \leq 2/7$ witnessed by the 7-vertex Moser spindle. The construction marks an advance on a question posed by Paul Erdős in 1987, asking whether $r \geq 1/4$.

The GFCN method can be readily combined with the classical Delsarte LP bound, yielding improved density bounds. For a measurable set $A \subset S$, one considers its autocorrelation function $f_A(d)$, the probability that two random points at mutual distance d both lie in A. $f_A(d)$ is of positive type. For distances d appearing in X, $f_A(d)$ can be simultaneously expressed as a linear combination of the observation probabilities and as a Fourier expansion with nonnegative coefficients. This gives additional linear constraints to the GFCN LP. With these, we can settle the conjecture by Erdős that $m_1(\mathbb{R}^2) < 1/4$:

Theorem 2 ([2]).

$$m_1(\mathbb{R}^2) \leq 0.247.$$

The best current bound, unpublished, is $m_1(\mathbb{R}^2) < 0.242$, found using an improved graph search algorithm by Pál Zsámboki.

Witsenhausen's problem on the sphere [5], also known as Kalai's double-cap conjecture [6], asks for the largest density of a measurable subset of $\mathbb{S}^d$ containing no orthogonal pair. In dimension 2, the natural double-cap construction gives the lower bound $1 - 1/\sqrt{2} \approx 0.2929$. The upper bound remained at the trivial value of $1/3$ for 40 years after the question was posed. The previous best upper bound, 0.2977, is due to Bekker et al. [7], using optimization-hierarchy methods. Using the GFCN method with the Delsarte constraints, our computer search identified a 33-vertex probe yielding:

Theorem 3 ([8]). *Every measurable subset of $\mathbb{S}^2$ containing no orthogonal pair has density at most* 0.2953.

Moving to hypergraph independence densities, let us now consider measurable sets in the plane avoiding unit equilateral triangles. A simple construction with alternating parallel stripes gives density $1/2$. It is reasonable to conjecture that this is optimal. This quantity was considered by Castro-Silva et al. [9], but their focus is on higher-dimensional variants, giving the bound 0.64355 for the planar case. Our computer search yields a 15-vertex probe proving the following, without using the Delsarte constraints:

Theorem 4 (unpublished)**.** *Every measurable subset of* $\mathbb{R}^2$ *containing no unit equilateral triangle has density at most* $9/16 = 0.5625$.

An analogous question on the sphere is the maximal density $m_\perp(\mathbb{S}^2)$ of a measurable $A \subset \mathbb{S}^2$ such that no three elements of A form an orthonormal basis. This quantity has appeared previously in [11, 9, 12, 10]. Its relationship to so-called non-Kochen-Specker sets in quantum mechanics provides further motivation to the question [10]. Our bound in dimension two appears to be the first improvement over the 2/3 bound obtained from the trivial averaging argument:

Theorem 5 (unpublished)**.** $m_\perp(\mathbb{S}^2) \leq 7/11$.

The discussion above has presented the GFCN method in the context of independence density bounds. However, probes can inspect measurable colorings, not just independent sets. So far, our attempts to utilize this idea have only yielded a GFCN-based reproof of Falconer's classical theorem, using a 15-vertex probe:

Theorem 6 ([13])**.** *The measurable chromatic number* $\chi_m(\mathbb{R}^2)$ *of the plane is at least* 5.

We aim to improve this to the new bound $\chi_m(\mathbb{R}^2) \geq 6$, but achieving this seems to require new ideas. A promising direction is conic programming, which may provide a more general way to strengthen the GFCN method.

References

[1] D. Castro-Silva, *Geometrical sets with forbidden configurations*, preprint, `https://arxiv.org/abs/2102.10018`.

[2] G. Ambrus, A. Csiszárik, M. Matolcsi, D. Varga, and P. Zsámboki, *The density of planar sets avoiding unit distances*, Math. Program. **207** (2024), 303–327. `https://arxiv.org/abs/2207.14179`

[3] M. Matolcsi, I. Z. Ruzsa, D. Varga, and P. Zsámboki, *The fractional chromatic number of the plane is at least* 4, arXiv:2311.10069, 2025. `https://arxiv.org/abs/2311.10069`

[4] H. T. Croft, *Incidence incidents*, Eureka **30** (1967), 22–26.

[5] H. S. Witsenhausen, *Spherical sets without orthogonal point pairs*, Amer. Math. Monthly **81** (1974), no. 10, 1101–1102.

[6] G. Kalai, *Some old and new problems in combinatorial geometry I: around Borsuk's problem*, in *Surveys in Combinatorics 2015*, London Math. Soc. Lecture Note Ser. 424, Cambridge Univ. Press, Cambridge, 2015, pp. 147–174.

[7] B. Bekker, O. Kuryatnikova, F. M. de Oliveira Filho, and J. C. Vera, *Optimization hierarchies for distance-avoiding sets in compact spaces*, Trans. Amer. Math. Soc. **378** (2025), no. 8, 5573–5595.

[8] D. Czifra, Á. Dúcz, M. Matolcsi, D. Varga, and P. Zsámboki, *Improved bounds for the double cap conjecture*, arXiv:2605.28709, 2026. `https://arxiv.org/abs/2605.28709`

[9] D. Castro-Silva, F. M. de Oliveira Filho, L. Slot, and F. Vallentin, *A recursive Lovász theta number for simplex-avoiding sets*, Proc. Amer. Math. Soc. **150** (2022), no. 8, 3307–3322.

[10] T. Williams and A. Constantin, *Maximal non-Kochen-Specker sets and a lower bound on the size of Kochen-Specker sets*, Phys. Rev. A **111** (2025), no. 1, 012223.

[11] A.-M. Perez, *On generalizations of the double cap conjecture*, Research Science Institute report, 2019.

[12] D. Zakharov, *Spherical sets avoiding orthonormal bases*, preprint, `https://arxiv.org/abs/2310.06821`.

[13] K. J. Falconer, *The realization of distances in measurable subsets covering* $\mathbb{R}^n$, J. Combin. Theory Ser. A **31** (1981), 184–189.

Exploiting Symmetry in Conic Optimization

Daniel Brosch

Many optimization problems arising in combinatorics, discrete geometry, coding theory, and polynomial optimization possess large symmetry groups. Hierarchies of semidefinite programming relaxations, such as the moment-SOS hierarchy, often yield very strong bounds on such problems, but grow too quickly to be computed beyond the first level. Exploiting the problem's inherent structure, such as symmetries or sparsity, can mitigate this. There are three methods of symmetry reduction I consider here: the (half-)degree principle, orbit space reductions and block-diagonalization based on Artin-Wedderburn theory. For a more thorough survey on the topic I refer to the recent chapterby Moustrou, Riener and Verdure [1].

1. The half-degree principle

The first approach applies to optimization problems involving symmetric polynomials. Assume we want to compute $\inf_{x\in\mathbb{R}^n} f(x)$, where $f \in \mathbb{R}[x]^{S_n}$ is invariant under the standard action of S_n on variables $x_1, \dots, x_n$. It is clear that if $x^\star$ is a minimizer of f, then so is $\mathrm{sort}(x^\star)$, the point we obtain by sorting the coefficients by value. But we can assume even more: the minimizer only has a finite number of distinct components.

Theorem 1 (Timofte [3], 2003. See also Riener [2], 2012.)**.** *Let $f \in \mathbb{R}[x]^{S_n}$ be of degree d and let $k := \max\{2, \lfloor \frac{d}{2} \rfloor\}$. Then $f(x) \geq 0$ for all $x \in \mathbb{R}^n$ if and only if $f(x) \geq 0$ for all x with at most k distinct components. If f has a minimizer, then f has a minimizer with at most k distinct components.*

Thus a problem in n variables can be reduced to a collection of lower-dimensional problems depending only on k parameters.

2. Orbit space reduction

A more general reduction method is based on invariant theory. Let a finite group G act linearly on $\mathbb{R}^n$. If an optimization problem is invariant under this action, then the objective function depends only on the orbit of a point x.

Theorem 2 (Hilbert's finiteness theorem)**.** *Let G be a finite group. Then the invariant subalgebra $\mathbb{R}[x]^G$ is generated by finitely many homogenous invariants, i.e. there is a finite set of invariant homogenous polynomials $\pi_1, \dots, \pi_m \in \mathbb{R}[x]^G$ such that every invariant polynomial $f \in \mathbb{R}[x]^G$ can be written as a polynomial in $\pi_1, \dots, \pi_m$. If I is the ideal of algebraic relations between the π_i, then*

$$\mathbb{R}[x]^G \simeq \mathbb{R}[\pi_1, \dots, \pi_m]/I.$$

The associated Hilbert map

$$\Pi : \mathbb{R}^n \to \mathbb{R}^m, \qquad x \mapsto (\pi_1(x), \ldots, \pi_m(x))$$

identifies points lying in the same orbit and provides coordinates on the orbit space. We call the image of Π the *orbit space* $\mathbb{R}^n//G$. Consequently, a G-invariant optimization problem can be reformulated as an optimization problem over the image of the orbit space

$$\begin{array}{lcl} \inf\limits_{x\in\mathbb{R}^n} \ f(x) & = & \inf\limits_{z\in\mathbb{R}^m} \ \tilde{f}(z) \\ \text{s.t. } g_i(x) \geq 0 \quad \text{for } i = 1, \ldots, m, & & \text{s.t. } \tilde{g}_i(z) \geq 0 \quad \text{for } i = 1, \ldots, m, \\ & & \qquad z \in \mathbb{R}^m//G. \end{array}$$

where $f = \tilde{f}(\pi_1, \ldots, \pi_m)$ and the $g_i = \tilde{g}_i(\pi_1, \ldots, \pi_m)$ are G-invariant polynomials.

The main challenge is to describe the image of the Hilbert map. For finite groups and, the orbit space can be characterized using results of Procesi and Schwarz [4].

Theorem 3 (Procesi and Schwarz[4])**.** *Let $G \subseteq \mathrm{GL}_n(\mathbb{R})$ be a compact matrix group, and let $\Pi = (\pi_1, \ldots, \pi_m)$ be the fundamental invariants of G. Then the orbit space is given by polynomial inequalities,*

$$\mathbb{R}^n//G = \Pi(\mathbb{R}^n) = \{z \in \mathbb{R}^m \mid J(z) \succcurlyeq 0, z \in V(I)\},$$

where I is the ideal of relations of $\pi_1, \ldots, \pi_m$ and

$$J = \left(\sum_{k=1}^{n} \frac{\partial \pi_i}{\partial x_k} \frac{\partial \pi_j}{\partial x_k} \right)_{1 \leq i,j \leq m} \in (\mathbb{R}[x]^G)^{m \times m}.$$

3. Block-diagonalization

The third approach is representation-theoretic and is especially powerful for semidefinite programming and sum-of-squares optimization. Here I again refer to the recent chapter [1], as well as the seminal paper on exploiting symmetries for polynomial optimization by Gatermann and Parrilo [5].

Suppose a semidefinite program is invariant under the action of a finite group G, i.e. there exists a matrix representation $\phi\colon G \to \mathrm{GL}_n(\mathbb{C})$ such that if X is feasible, then so is $\phi(g)^{-1} X \phi(g)$. By averaging over the group, one may assume that an optimal solution is itself G-invariant. The corresponding matrix variable therefore belongs to the commutant algebra of the group action, that is, the algebra of matrices commuting with every representing matrix of G

$$\mathrm{Com}(\phi(G)) = \{T \in \mathbb{C}^{n\times n} \text{ such that } T\phi(g) = \phi(g)T \text{ for all } g \in G\}.$$

The key observation is that this commutant algebra is semisimple. By the Artin–Wedderburn theorem it decomposes as a direct product of matrix algebras. After a suitable change of basis, every matrix in the commutant algebra becomes block diagonal, with one block for each irreducible representation occurring in the group action.

As a consequence, a large positive-semidefinite constraint can be replaced by several much smaller positive-semidefinite constraints. If the original matrix has

dimension N, the reduction often replaces one $N \times N$ matrix by a collection of blocks whose total size is dramatically smaller. This frequently turns otherwise intractable semidefinite programs into computationally feasible ones.

The practical computation of the block decomposition can be carried out, in the general finite group case, using character theory and the decomposition of the underlying representation into irreducible components. For many groups, such as S_n and O_n the reduction can be computed significantly more efficient by exploiting the well-understood representation theory of S_n and harmonic analysis of O_n.

4. Conclusion

The degree principle, orbit space reduction, and block-diagonalization represent three complementary approaches to exploiting symmetry. The degree principle reduces the number of distinct coordinate values that need to be considered. Orbit space reduction replaces the original variable space by a lower-dimensional space of invariants. Block-diagonalization uses representation theory to simplify convex relaxations, particularly semidefinite programs.

References

[1] P. Moustrou, C. Riener, and H. Verdure, *Symmetries in Polynomial Optimization*, in: M. Kočvara, B. Mourrain, and C. Riener (eds.), *Polynomial Optimization, Moments, and Applications*, Springer Optimization and Its Applications.

[2] C. Riener, *On the degree and half-degree principle for symmetric polynomials*, J. Pure Appl. Algebra **216** (2012), no. 4, 850–856.

[3] V. Timofte, *On the positivity of symmetric polynomial functions. Part I: General results*, Journal of Mathematical Analysis and Applications **284** (2003), 174–190.

[4] C. Procesi and G. Schwarz, *Inequalities defining orbit spaces*, Inventiones Mathematicae **81** (1985), 539–554.

[5] K. Gatermann and P. A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, Journal of Pure and Applied Algebra **192** (2004), 95–128.

Reporter: Lucas ter Voert

Participants

Dr. Jason Altschuler
Department of Statistics and Data Science, Wharton School
University of Pennsylvania
265 South 37th Street
Philadelphia PA 19104
UNITED STATES

Jonas Britz
Centrum Wiskunde & Informatica (CWI)
1090 GB Amsterdam
NETHERLANDS

Dr. Daniel Brosch
Institut für Mathematik
Universität Klagenfurt
Universitätsstr. 65-67
9020 Klagenfurt
AUSTRIA

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
2628 CD Delft
NETHERLANDS

Dr. Fernando Mário de Oliveira Filho
Dept. of Mathematics and Computer Science
Delft University of Technology
Postbus 5
2628 CD Delft
NETHERLANDS

Dr. Mareike Dressler
School of Mathematics and Statistics
University of New South Wales
Sydney NSW 2052
AUSTRALIA

Prof. Dr. Simon Foucart
Department of Mathematics
Texas A & M University
College Station TX 77843-3368
UNITED STATES

Prof. Dr. François Glineur
UCLouvain / ICTM / INMA
Avenue Georges Lemaître, 4
P.O. Box bte L4.05.01
B-1348 Louvain-la-Neuve
BELGIUM

Dr. Sander Gribling
Tilburg University
5037 AB Tilburg Cedex 13
NETHERLANDS

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

Dr. Mioara Joldes
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

Prof. Máté Matolcsi
HUN-REN Alfred Renyi Institute of
Mathematics
Realtanoda u. 13-15
1053 Budapest
HUNGARY

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

Johannes Schmucker
Institut für Mathematik
Universität Klagenfurt
Universitätsstr. 65-67
9020 Klagenfurt
AUSTRIA

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

Dr. Adrien Taylor
National Institute for Research in
Computer Science and Control
INRIA - CS 42112
47 Rue Barrault
75013 Paris Cedex
FRANCE

Lucas ter Voert
Tilburg University
Fatimastraat 59
5021 AM Tilburg
NETHERLANDS

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

Dr. Daniel Varga
Alfred Renyi Institute of Mathematics
Hungarian Academy of Sciences
P.O. Box 127
1364 Budapest
HUNGARY

Prof. Dr. Juan C. Vera Lizcano
Department of Econometrics
Tilburg University
Room K 509
P.O. Box 90153
5000 LE Tilburg
NETHERLANDS

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