Mathematisches Forschungsinstitut Oberwolfach

Report No. 35/2024

DOI: 10.4171/OWR/2024/35

Polynomial Optimization for Nonlinear Dynamics: Theory, Algorithms and Applications

Organized by
Giovanni Fantuzzi, Erlangen
David Goluskin, Victoria
Jean-Bernard Lasserre, Toulouse

28 July – 2 August 2024

Abstract. This workshop focused on using computational tools of polynomial optimization to deduce information about nonlinear dynamical systems, including systems governed by ordinary or partial differential equations. This approach sits at the interface of various research areas, requiring combinations of applied nonlinear dynamics and control theory, polynomial optimization, real algebraic geometry, partial differential equations, and variational analysis. The workshop brought together researchers in these different areas to share recent advances and to build the connections required for further progress.

Mathematics Subject Classification (2020): 90C23, 35A15, 93D30, 49-XX, 37-XX.

Introduction by the Organizers

The workshop *Polynomial optimization for nonlinear dynamics: theory, algorithms, and applications* was organized by Giovanni Fantuzzi (Erlangen, Germany), David Goluskin (Victoria, Canada) and Jean-Bernard Lasserre (Toulouse, France). The workshop was attended by 43 onsite participants, including 12 who were PhD students or postdocs. Participants came from universities widely distributed across Europe and North America, as well as from Australia, Chile, China and Israel. Several hours were devoted to research talks on all five mornings and on three of the evenings. Several hours after lunch, as well as the hours after dinner, were kept open for free discussion. This discussion time was put to full

use, with almost all participants proactively grouping themselves in various ways and scattering to blackboards throughout the facilities.

The topic of the workshop was a fast-growing family of computational methods that use polynomial optimization to deduce information about nonlinear dynamical systems. These dynamical systems are often, but not always, governed by nonlinear ODEs or PDEs, and the results obtained with help from polynomial optimization typically surpass what other existing methods can give. The mathematical progress needed to further improve, generalize, and analyze these optimization-based methods lies at the interface of several different research communities. These communities include applied nonlinear dynamics and control theory, polynomial optimization, real algebraic geometry, and analysis of PDEs and variational problems. A major success of the workshop was to bring these disparate communities together and strengthen the ties between them, which were facilitated by four one-hour overview talks and solidified by many impromptu discussions.

The four overview talks were designed both to be pedagogical and to highlight open challenges. David Goluskin opened the workshop with a high-level view of the various topics to be discussed and their interrelation, with emphasis on which areas call for further work. Didier Henrion gave an introduction to the language of occupation measures and explained how to formulate optimal control problems in terms of psuedomoments of measures, which are related by convex duality to polynomial optimization problems subject to sum-of-squares constraints. Monique Laurent surveyed what is known about convergence rates for upper and lower bounds on global minima of polynomials, which laid the groundwork for subsequent discussions of convergence rates for the polynomial optimization problems arising in the study of dynamical systems. Giovanni Fantuzzi gave a unified view on efforts to generalize methods from finite to infinite dimensions, such as going from minimizing polynomials to minimizing integral functionals, or going from ordinary to partial differential equations.

In addition to the overview talks, there were 17 research talks, and seven junior participants gave three-minute lightning talks in order to raise their visibility and introduce their research areas to the other participants. The timing of the research talks was flexible, but most were nominally allotted a half hour, and a few were allotted more time. Speakers were selected based on which topics the organizers considered to most likely stimulate discussion during the workshop. Corbinian Schlosser and Matteo Tacchi-Bénard presented recent advances in proving convergence rates for moment-sum-of-squares hierarchies, building on Monique Laurent's overview and taking steps towards generalization of such convergence rates to dynamical systems. Going beyond the ODE case, Elizabeth Carlson and Emilia Fridman presented applications of convex programming to particular PDEs, while Jared Miller addressed stochastic effects and hybrid systems. Federico Fuentes explained how to apply polynomial optimization techniques to approximate global optimizers of integral variational problems, while Ian Tobasco described challenging problems from materials science as a possible new application area for the methods of the workshop. Amir Ali Ahmadi and Mario Sznaier spoke about how

to learn dynamical systems with control, and Jason Bramburger showed how polynomial optimization methods can be applied when one has data a dynamical system rather than its governing equations. Pablo Parrilo discussed a version of the shortest-path problem for a network connecting convex sets, with an eye towards control. Heng Yang showed computational advances for using polynomial optimization to solve optimal control problems, and Michael Stingl kindly presented work by Michal Kočvara on ill-conditioning of computations that arise from polynomial optimization. Jie Wang presented ways to refine convergent hierarchies of polynomial optimization problems in the real and complex cases. Finally, turning to algebraic aspects of polynomial optimization and its generalizations, Philipp Di Dio discussed how evolution of polynomials under the heat equation or related operators may preserve or add positivity properties, while Mareike Dressler and Timo de Wolff spoke about conditions implying nonnegativity that differ from the more widely used sum-of-squares conditions.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-2230648, "US Junior Oberwolfach Fellows".

Workshop: Polynomial Optimization for Nonlinear Dynamics: Theory, Algorithms and Applications

Table of Contents

Abstracts

Polynomial optimization methods for dynamics: the standard approach and the need to go beyond it

David Goluskin

This overview talk explained how the different research areas represented at the meeting fit together, and it emphasized where further advances are most needed. The topic of the meeting is a family of methods that all loosely follow a certain template. The steps below describe the most standard version of this template.

(1) Begin with a particular ODE or discrete-time dynamical system on state space $\mathbb{R}^n$.
(2) Formulate sufficient conditions wherein finding an auxiliary function $V : \mathbb{R}^n \to \mathbb{R}$ subject to certain inequalities will imply a statement about the dynamical system. Alternatively, such statements can be formulated on the dual side in terms of measures.
(3) Assuming all relevant expressions are polynomial, inequality constraints can be strengthened into polynomial sum-of-squares (SOS) constraints. The result is an SOS optimization problem. Successively enlarging the polynomial spaces for the auxiliary function V and other polynomial expressions gives a hierarchy of SOS optimization problems. On the dual side, there is a corresponding hierarchy of moment problems.
(4) An SOS-moment problem is converted into a semidefinite program. The particular semidefinite program depends on the choice of polynomial basis – the monomial basis is the standard choice.
(5) The semidefinite program is solved using a primal-dual interior-point algorithm, whose output implies a statement about the dynamical system being studied.

I listed types of various statements about dynamical systems that can be produced using the standard template. For an explicit example of how this template works, I described how SOS optimization can give upper bounds on the maximum infinite-time average of a chosen quantity among all trajectories in a chosen set, including theoretical guarantees that the SOS hierarchies converge to the maximum time average.

The rest of the talk described open challenges, both for studying methods that follow the standard template and for going beyond these methods. First, within the standard template there is room for more such methods giving different types of statements about dynamical systems, and there is a need for stronger convergence theory. Convergence theorems are lacking when dynamics are not restricted *a priori* to a compact region of phase space, and convergence guarantees currently lack rates of convergence as polynomial degrees are raised in the SOS hierarchies. Second, computational implementations following the standard template often suffer from prohibitively poor numerical conditioning, which motivates modifying the template in various ways. One could modify step 5 by using a different algorithm

to solve semidefinite programs. One could modify step 4 by choosing different polynomial bases, and it is an open challenge to define what is meant by optimal bases for different SOS problems, let alone to find those bases. One could also modify step 3 by using other non-SOS nonnegativity certificates. Third, these methods must be generalized to infinite-dimensional state spaces, as when going from dynamical ODEs to PDEs.

Performance Analysis of Sum-Of-Squares Hierarchies for Polynomial Optimization

Monique Laurent

We consider the polynomial optimization problem, of the form

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

with $f, g_j \in \mathbb{R}[x]$ multivariate polynomials. Such problems are in general nonlinear, nonconvex hard problems, already for simple semialgebraic sets S like the ball, the unit sphere, the hypercube or the simplex. Indeed, they can capture NP-hard problems like testing convexity of a (quartic) polynomial, deciding whether an integer sequence can be partitioned, or computing the maximum cardinality of an independent set in a graph.

Hierarchies of upper and lower bounds can be constructed using the cone Σ of sums of squares of polynomials. For an integer $r \in \mathbb{N}$ define the parameters

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

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

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

Here, μ is a positive Borel measure with support S and we set $g_0 = g_\emptyset = 1$, $g_J = \prod_{j\in J} g_j$. The set of polynomials that admit a representation $\sum_{j=0}^m \sigma_j g_j$ (with $\sigma_j \in \Sigma$ and $\deg(\sigma_j g_j) \leq 2r$) forms the quadratic module $\mathcal{Q}(g)_{2r}$ generated by the g_j's, truncated at degree $2r$. The set of polynomials of the form $\sum_{J\subseteq[m]} \sigma_J g_J$ (with $\sigma_J \in \Sigma$ and $\deg(\sigma_J g_J) \leq 2r$) forms the pre-ordering $\mathcal{O}(g)_{2r}$ generated by the g_j's, truncated at degree $2r$. Since $\mathcal{Q}(g)_{2r} \subseteq \mathcal{O}(g)_{2r}$ and any polynomial in $\mathcal{O}_{2r}(g)$ is nonnegative over S, we have

$$f_{(r),\mathcal{Q}} \leq f_{(r),\mathcal{O}} \leq f_{\min} \leq f^{(r)}.$$

Moreover, since sums of squares of polynomials can be modelled using semidefinite programs, each of the parameters $f^{(r)}$, $f_{(r),\mathcal{Q}}$, $f_{(r),\mathcal{O}}$ can be modelled (and computed) using semidefinite optimization.

Assume S is compact. Then, the upper bounds $f^{(r)}$ converge asymptotically to $f_{\min}$ [7]. Based on Schmüdgen's Positivstellensatz, the lower bounds $f_{(r),\mathcal{O}}$ are also known to converge asymptotically to $f_{\min}$. The same holds for the (weaker)

lower bounds $f_{(r),\mathcal{Q}}$ under an Archimedean condition on the algebraic description of S, based on Putinar's Positivstellensatz (see [6]).

In this lecture we discuss the state-of-the-art results concerning the quality of the upper and lower bounds, namely how fast the error ranges $f^{(r)} - f_{\min}$, $f_{\min} - f_{(r),\mathcal{O}}$ and $f_{\min} - f_{(r),\mathcal{Q}}$ tend to 0 as a function of the order r of the relaxation. For both upper and lower bounds the results can be roughly divided into two categories: for 'general' sets S, and for 'special sets' S that admit some nice symmetry structure (like the ball, the unit sphere, the hypercube or the simplex). Different proof techniques are used for each category and, quite naturally, stronger results can be shown for special sets. In addition, there is an intimate link between the analysis for the upper and lower bounds.

We refer to the recent paper [9] for an overview of the state-of-the-art results and for detailed references to the literature.

Convergence analysis of the upper bounds. The most general result is when the set S is a convex body, or a semialgebraic set with a dense interior (selecting the Lebesgue measure for μ); then, one can show $f^{(r)} - f_{\min} = O\big(\frac{\log^2 r}{r^2}\big)$ [12]. The analysis has two key steps: (1) reduce the search to a univariate sum of squares s and then set $\sigma(x) = s(f(x))$, (2) select the univariate sum of squares s as a tight approximation of the Dirac delta at an extremity of an interval (using so-called needle polynomials).

A stronger analysis in $O\big(\frac{1}{r^2}\big)$ can be shown for special sets S: for the interval $[-1,1]$ (with $d\mu(x) = (1-x^2)^{-1/2}dx$ in [4], with $d\mu(x) = (1-x^2)^{\lambda}$ and $\lambda \geq -1/2$ in [12]), for the ball and the simplex in [12], and for the sphere in [5]. The starting point is expressing the parameter $f^{(r)}$ as the smallest eigenvalue of the matrix $A_f = (\int_S f(x)p_\alpha(x)p_\beta(x)d\mu(x))_{|\alpha|,|\beta|\leq r}$, where $\{p_\alpha\}$ are orthonormal polynomials with respect to the inner product provided by the measure μ on S. In the univariate case, when $K = [-1,1]$ and $f(x) = x$, the matrix A_x coincides with the (tridiagonal) Jacobi matrix, whose entries are given by the three-term recurrence satisfied by the polynomials p_k, and whose eigenvalues are the roots of the polynomial p_{r+1}. Hence, in this case, $f^{(r)}$ coincides with the least root of p_{r+1}, known to be in the order $-1 + \Omega\big(\frac{1}{r^2}\big)$ for Jacobi type measures $(1-x^2)^{\lambda}$ ($\lambda > -1$). This is the key ingredient used in [4] for the analysis in the case when $S = [-1,1]^n$. Additional reductions are needed to extend the analysis to other sets. This includes an integration trick in order to extend the analysis to the sphere in [5] and looking at the local shape around a global minimizer to extend the analysis to the ball, the simplex and 'round' convex bodies in [12].

Convergence analysis of the lower bounds. The best general result has been obtained recently and shows a convergence rate in $O\big(\frac{1}{r^c}\big)$ for the bounds $f_{(r),\mathcal{Q}}$ [1], providing an impressive improvement on an earlier result in [10] (with a logarithmic dependence on r). The analysis is technically involved, and combines a variety of techniques, including the Łojasiewicz inequality, approximation theory tools, and a reduction to the analysis for the case of the hypercube $[-1,1]^n$ (from [8]).

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

(P1) $\mathbf{K}_r$ preserves the constant polynomial: $\mathbf{K}_r 1 = 1$,
(P2) if $p \geq 0$ on S, then $\mathbf{K}_r p \in \mathcal{O}(g)_{2r}$,
(P3) $\|\mathbf{K}_r^{-1} f - f\|_\infty \leq \epsilon$.

In (P3) we assume f is scaled and translated so that its minimum (resp., maximum) over S is equal to 0 (resp., 1). It is not difficult to see that one can conclude $f_{\min} - f_{(r),\mathcal{O}} \leq \epsilon$. The challenge then is constructing a linear operator $\mathbf{K}_r$ having these properties. Property (P3) says (roughly) that $\mathbf{K}_r$ should be close to the identity operator when acting on polynomials of degree at most $d = \deg(f)$. The approach is to select a polynomial kernel $K_r(x,y) \in \mathbb{R}[x,y]_{2r,2r}$ and then to define the corresponding linear operator $\mathbf{K}_r$ that acts by convolution: for $p \in \mathbb{R}[x]$, define $\mathbf{K}_r p(x) = \int_S p(y) K_r(x,y) d\mu(y)$.

Consider first the case of the hypercube $K = [-1,1]^n$, equipped with the Chebyshev product measure $d\mu(x) = \prod_{i=1}^n (1-x_i^2)^{-1/2} dx_i$. Consider the multivariate Chebyshev polynomials $T_\alpha = \prod_{i=1}^n T_{\alpha_i}$ where T_k are the univariate Chebyshev polynomials. Then the T_α are orthogonal w.r.t. the inner product given by μ. Define the multivariate polynomial kernel $K_r(x,y) = \prod_{i=1}^n K_r(x_i,y_i)$, where the univariate kernel is of the form $K_r(x_i,y_i) = 1 + 2\sum_{k=1}^r \lambda_k^r T_k(x_i) T_k(y_i)$ (also known as the Jackson kernel), and the scalars λ_k^r are selected so that $0 \leq 1 - \lambda_k^r \leq \frac{\pi^2 d^2}{(r+2)^2}$ for $0 \leq k \leq d$ and $K_r(x_i,y_i) \geq 0$ on $[-1,1]^2$. By construction, the corresponding operator $\mathbf{K}_r$ satisfies (P1) and, using Markov-Lukacz, it also satisfies (P2). Finally, the selection of the scalars λ_k^r ensures that (P3) holds with $\epsilon = O\left(\frac{1}{r^2}\right)$, thus giving the analysis in $O\left(\frac{1}{r^2}\right)$ (as in [8]).

The same analysis in $O\left(\frac{1}{r^2}\right)$ is shown for the sphere in [3], and for the ball and the simplex in [11]. It however relies on a different class of polynomial kernels, constructed by exploiting the symmetry structure of the sets under consideration. The idea (pioneered in [3]) is construct the kernel $K_r(x,y)$ as a suitable perturbation of the Christoffel-Darboux kernel, as $K_r(x,y) = \sum_{k=0}^{2r} \lambda_k \sum_{|\alpha|=k} p_\alpha(x) p_\alpha(y)$, for appropriate scalars λ_k. Setting $\lambda_0 = 1$ ensures (P1) holds. The key (harmonic analysis) fact is that the sum $\sum_{|\alpha|=k} p_\alpha(x) p_\alpha(y)$ can be expressed in terms of an associated univariate polynomial, when dealing with the sphere, the ball or the simplex.

In the case of the unit sphere, this univariate polynomial is a Gegenbauer polynomial G_k evaluated at $x^T y$. Then, one selects $K_r(x,y) = s(x^T y)$, where $s(t) = \sum_{k=0}^{2r} \lambda_k G_k(t)$, and one searches for scalars λ_k so that s is a sum of squares (which ensures (P2) holds) and (P3) holds. Interestingly, the search for this univariate s boils down to an instance of the upper bounds $g^{(r)}$ for a suitably defined

univariate polynomial g on the interval $[-1, 1]$. As a consequence, it follows from the earlier results on the upper bounds that a suitable s can be found, which gives a suitable polynomial kernel $K_r(x, y)$ and thus a suitable kernel operator $\mathbf{K}_r$.

For the ball and the simplex, a similar approach can be followed, however with more technical details as the summation formulas for $\sum_{|\alpha|=k} p_\alpha(x)p_\alpha(y)$ (given in [13, 14]) are more involved. The full details can be found in [11].

Showing an improved convergence analysis for the lower bounds $f_{(r),\mathcal{Q}}$ for symmetric sets remains widely open. A recent result is an analysis in $O(1/r)$ for the case of the hypercube shown in [2].

References

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

[2] L. Baldi and L. Slot, *Degree bounds for Putinar's Positivstellensatz on the hypercube*, SIAM J. Appl. Algebra Geom., **8(1)** (2024).

[3] K. Fang and H. Fawzi, *The sum-of-squares hierarchy on the sphere, and applications in quantum information theory*, Math. Program., Ser. A, **190** (2021), 331–360.

[4] E. de Klerk and M. Laurent, *Worst-case examples for Lasserre's measure-based hierarchy for polynomial optimization on the hypercube*, Math. Oper. Res., **45(1)** (2020), 86–98.

[5] E. de Klerk and M. Laurent, *Convergence analysis of a Lasserre hierarchy of upper bounds for polynomial minimization on the sphere*, Math. Program., **193** (2022), 665–685.

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

[7] J. B. Lasserre, *A new look at nonnegativity on closed sets and polynomial optimization*, SIAM J. Optim., **21(3)** (2010), 864–885.

[8] M. Laurent and L. Slot, *An effective version of Schmüdgen's Positivstellensatz for the hypercube*, Optim. Lett., **17** (2023), 515–530.

[9] M. Laurent and L. Slot, *Convergence rates of sums of squares hierarchies for polynomial optimization*, arXiv:2408.04417 (2024).

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

[11] L. Slot, *Sum-of-squares hierarchies for polynomial optimization and the Christoffel-Darboux kernel*, SIAM J. Optim., **32(4)** (2022), 2612–2635.

[12] L. Slot and M. Laurent, *Improved convergence analysis of Lasserre's measure-based upper bounds for polynomial minimization on compact sets*, Math. Program., **193** (2022), 831–871.

[13] Y. Xu, *Summability of Fourier orthogonal series for Jacobi weight functions on the simplex in $\mathbb{R}^d$*, Proc. Amer. Math. Soc., **126(10)** (1998), 3027–3036.

[14] Y. Xu, *Summability of Fourier orthogonal series for Jacobi weight on a ball in $\mathbb{R}^d$*, Trans. Amer. Math. Soc., **351(6)** (1999), 2439–2458.

A Framework for Computing Convergence Rates for the Moment-SoS Hierarchy

Corbinian Schlosser
(joint work with Matteo Tacchi-Bénard, Alexey Lazarev)

The wide application of the moment-sum-of-squares (SoS) has its origins in polynomial optimization problems (POP). Those are of the form

$$\text{(POP)} \qquad f^* := \begin{array}{ll} \min\limits_{x\in\mathbb{R}^n} & f(x) \\ \text{s.t.} & x \in K \end{array}$$

where f is a given polynomial in n variabels, the set $K \subset \mathbb{R}^n$ is given by

$$K := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i = 1, \dots, m\}$$

for given polynomials $p_1, \dots, p_m$ in n variables. The starting point of the moment-SoS hierarchy lies in the following simple reformulation of POP as the following *linear* – but infinite dimensional – optimization problem (LP)

$$f^* = \begin{array}{ll} \sup\limits_{s\in\mathbb{R}} & s \\ \text{s.t.} & f - s \geq 0 \text{ on } K. \end{array} \tag{1}$$

For the moment-SoS hierarchy, the non-negativity constraint in (1) is replaced by an SoS constraint. This is done as follows: Let $\mathrm{SoS}[x]$ be the set of sums of squares of polynomials and, for $\ell \in \mathbb{N}$, the truncated quadratic module $\mathcal{Q}_\ell$ is defined as

$$\mathcal{Q}_\ell := \{\sigma_0 + \sum_{i=1}^{m} \sigma_i p_i : \sigma_0, \dots, \sigma_m \in \mathrm{SoS}[x], \deg \sigma_0, \deg \sigma_i p_i \leq 2\ell \text{ for } i = 2, \dots, m\}.$$

The moment-SoS hierarchy for POP reads, for $\ell \in \mathbb{N}$ consider

$$f_\ell^* := \begin{array}{ll} \sup\limits_{s\in\mathbb{R}} & s \\ \text{s.t.} & f - s \in \mathcal{Q}_\ell. \end{array} \tag{2}$$

Under mild compactness assumptions, the seminal works [1, 2] applied the celebrated Putinar's Positivstellensatze [3] to show $f_\ell^* \nearrow f^*$ as $\ell \to \infty$. It took until [4] that first general quantitative convergence results were established. Those are so-called effective Positivstellensätze which quantify $\ell \in \mathbb{N}$ for which a given polynomial f belongs to the truncated quadratic module $\mathcal{Q}_\ell$.

Naturally, the question on convergence rates transfers from POP to the variety of different applications of the moment-SoS hierarchy, see [5] for examples. However, only a few works [6, 7] have addressed it so far. In this talk, we outline a framework for computing convergence rates for the following general class of infinite-dimensional LPs and their moment-SoS hierarchy (on the right)

$$\mathrm{d}^* := \begin{array}{ll} \sup\limits_{p\in\mathbb{R}[y]} & T(p) \\ \text{s.t.} & \mathcal{A}\,p - h \geq 0 \text{ on } K \end{array} \quad , \quad \mathrm{d}_\ell^* := \begin{array}{ll} \sup\limits_{p\in\mathbb{R}[y]_\ell} & T(p) \\ \text{s.t.} & \mathcal{A}\,p - h \in \mathcal{Q}_\ell. \end{array} \tag{3}$$

where K is given as before, $T : \mathbb{R}[y] \to \mathbb{R}$ a linear operator, $\mathcal{A} : \mathbb{R}[y] \to \mathbb{R}[x]$ a linear operator between the polynomial ring in variables y respectively x, and h a given polynomial. It is readily verified that (1) and (2) are of the form (3).

Our proposed method [8] consists of the following steps:

1. Take a minimizer $p^\star$ of the LP (3).
2. If no minimizer exists, take a feasible polynomial p_ϵ with almost optimal cost.
3. Perturb $p^\star$ (resp. p_ϵ) into a strictly feasible polynomial $\hat{p}$ with $T(\hat{p}) \leq \mathrm{d}^\star + \delta$.
4. Apply effective Positivstellensätze, such as [9], to show that $\hat{p}$ is feasible for moment-SoS hierarchy at some level $\ell \in \mathbb{N}$.
5. Relate the approximation error δ and the hierarchy level ℓ to derive a convergence rate.

When no polynomial minimizer exists, it is still sometimes possible to enlarge the search space in the infinite dimensional LP in (3) so that a (non-polynomial) minimizer $\bar{p}$ exists. For those cases, in Step 2, we suggest choosing the polynomial p_ε as a polynomial approximation of $\bar{p}$ where quantitative properties are inferred via Jackson-type inequalities.

For Step 3, we assume a Slater condition being satisfied for the LP in (3).

We illustrate our method by showing how it encompasses the existing work [7] on volume computation of semialgebraic sets. By leveraging a state-of-the-art Positivstellensatz [9], we improve their logarithmic convergence rate to a polynomial one. This talk is complimented by the talk by Matteo Tacchi-Bénard.

References

[1] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., **11(3)** (2001), 796–817.

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

[3] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J., **42(3)** (1993), 969–984.

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

[5] D. Henrion, M. Korda, and J. B. Lasserre, *The Moment-SOS Hierarchy: Lectures in Probability, Statistics, Computational Geometry, Control and Nonlinear PDEs*, Vol. 4, World Scientific, 2020.

[6] M. Korda, D. Henrion, and C. N. Jones, *Convergence rates of moment-sum-of-squares hierarchies for optimal control problems*, Syst. Control Lett., **100** (2017), 1–5.

[7] M. Korda and D. Henrion, *Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets*, Optim. Lett., **12** (2018), 435–442.

[8] C. Schlosser, M. Tacchi, and A. Lazarev, *Convergence rates for the moment-SOS hierarchy*, arXiv preprint arXiv:2402.00436 (2024).

[9] L. Baldi and B. Mourrain, *On the effective Putinar's Positivstellensatz and moment approximation*, Math. Program., **200(1)** (2023), 71–103.

Computing Volumes with Sums-Of-Squares

Matteo Tacchi-Bénard

The problem of computing the volume of a compact set is known to be very difficult in general. Even finding the exact volume of a convex polytope is a computational challenge [1], while the volume of generic convex sets requires very sophisticated methods to be reliably approximated [2, 3]. When the considered set is not even convex, the task becomes even harder, reducing the possibilities to Monte-Carlo or discretization techniques. However, if instead of being convex the set at hand is assumed to be semialgebraic, then some powerful tools can be used to approximate its volume [4, 5]. More specifically, [5] leverages the moment-SoS hierarchy [6] to compute the volume of the set

$$\mathbf{S} = \{\boldsymbol{x} \in \mathbb{R}^n \ : \ g(\boldsymbol{x}) \geq 0\},$$

where $g \in \mathbb{R}[\boldsymbol{x}]$ is a polynomial. In this setting, assuming that $\mathbf{S} \subset \mathbf{B} = [-1,1]^n$, the moment-SoS hierarchy is implemented to approximate the volume $\mathrm{vol}(\mathbf{S})$ as the solution of the following Generalized Moment Problem (GMP) and its dual:

$$\begin{aligned} \mathrm{vol}(\mathbf{S}) = \max \int \mu & \qquad\qquad \mathrm{vol}(\mathbf{S}) = \inf \int_{[-1,1]^n} w \\ \text{s.t. } \mu \in \mathcal{M}(\mathbf{S})_+ & \qquad\qquad \text{s.t } w|_{\mathbf{S}} - 1 \in C^0(\mathbf{S})_+ \\ \overline{\mu} \in \mathcal{M}(\mathbf{B})_+ & \qquad\qquad w|_{\mathbf{B}} \in C^0(\mathbf{B})_+ \\ \forall w \in \mathbb{R}[\boldsymbol{x}], \int w\ \mu + \int w\ \overline{\mu} = \int_{[-1,1]^n} w & \qquad\qquad w \in \mathbb{R}[\boldsymbol{x}], \end{aligned} \tag{1a, 1b}$$

where for a compact set $\mathbf{K} \subset \mathbb{R}^n$, $\mathcal{M}(\mathbf{K})_+$ denotes the cone of Radon measures on $\mathbf{K}$, dual to the cone $C^0(\mathbf{K})_+$ of nonnegative continuous functions on $\mathbf{K}$, with respect to the topology of uniform convergence. Then, the moment-SoS hierarchy consists in replacing $C^0(\mathbf{K})_+$ with the corresponding truncated quadratic module $\mathcal{Q}(\mathbf{K})_r$ (see an example description of $\mathcal{Q}(\mathbf{S})_r$ in (4a) below) for increasing values of the relaxation order r. The challenge for its application to problem (1) is that, while the primal (1a) admits the Lebesgue measure supported on $\mathbf{S}$ as its minimizer, constraints qualification does not hold: the dual (1b) has neither polynomial nor continuous minimizer. Instead, it only admits minimizing sequences of polynomials that approximate the (discontinuous) indicator $\mathbb{1}_{\mathbf{S}}$ of $\mathbf{S}$ from above in the $L^1(\mathbf{B})$ topology, resulting in a Gibbs phenomenon in numerical computations. To cope with that difficulty, in [7], redundant constraints were added to the primal (1a), that have no effect on the infinite dimensional GMP but sharply tighten its moment relaxations, in order to speed up the practical convergence of the corresponding hierarchy. More precisely, in [8] we formalized the following "Stokes constraints" to be added to (1a):

$$\exists \nu \in \mathcal{M}(\partial\mathbf{S})_+; \quad \forall v \in C^2(\mathbf{S}), \quad \int \Delta v\ \mu = -\int \nabla g^\top \nabla v\ \nu, \tag{2}$$

which essentially is a variation of the divergence theorem, also known as Stokes', Gauss' or even Green's formula (∇g representing an inward pointing normal vector to $\partial\mathbf{S}$ and ν a scaled Hausdorff measure on $\partial\mathbf{S}$). Moreover, we proved that assuming that $\mathbf{S}$ has a smooth boundary $\partial\mathbf{S}$ and adding the above constraints to (1a) resulted in existence of a minimizer on the corresponding dual problem, which then reads as follows:

$$\begin{aligned} \mathrm{vol}(\mathbf{S}) = \min \int_{[-1,1]^n} & w \\ \text{s.t } & w|_{\mathbf{S}} + \Delta v - 1 \in C^0(\mathbf{S})_+ \\ & w \in C^0(\mathbf{B})_+ \\ & \left(\nabla g^\top \nabla v\right)\big|_{\partial\mathbf{S}} \in C^0(\partial\mathbf{S})_+ \\ & v \in C^2(\mathbf{S}). \end{aligned} \tag{3}$$

The proof of existence of a minimizer for (3) leverages the fact that its optimality conditions boil down to a Poisson PDE with Neumann boundary conditions, for which existence and regularity of solutions is well-documented. The resulting optimal $(v^\star, w^\star)$ are then approximated with appropriate polynomials with the moment-SoS hierarchy, with no Gibbs phenomenon occurring as there is no discontinuity anymore. While qualitatively explaining the spectacular practical improvement in the accuracy of the resulting SoS programming approximations, this reasoning lacked a quantitative analysis, which we recently presented in [9], leveraging new results on effective Positivstellensätze (P-sätze) [10]. Namely, under mild assumptions, [10] states that a degree d polynomial $p \in \mathbb{R}_d[\boldsymbol{x}]$ that is positive on $\mathbf{S}$ can be represented under the form

$$p(\boldsymbol{x}) = \sigma_0(\boldsymbol{x}) + g(\boldsymbol{x}) \cdot \sigma_1(\boldsymbol{x}) \tag{4a}$$

with $\sigma_0, \sigma_1 \in \mathrm{SoS}[\boldsymbol{x}]$ (polynomial sums of squares) and $\max(\deg(\sigma_0), \deg(g \cdot \sigma_1)) \leq 2r$ for any level of SoS representation in the hierarchy $r \in \mathbb{N}$ satisfying the following inequality:

$$r \geq \gamma\, d^{3.5nL} \left(\frac{\max\{|p(\boldsymbol{x})| \ : \ \boldsymbol{x} \in [-1,1]^n\}}{\min\{p(\boldsymbol{x}) \ : \ \boldsymbol{x} \in \mathbf{S}\}} \right)^{2.5nL} \tag{4b}$$

for some constants $\gamma, L > 0$ depending on the geometry of $\mathbf{S}$ (a similar result holds for $\mathbf{B}$ and $\partial\mathbf{S}$ instead of $\mathbf{S}$).

Then, [9] proposes the following procedure to compute a convergence rate of the SoS programming approximations of (1b) (resp. (3)):

(1) "Effective Weierstraß theorem": Approximate the optimizer $\mathbb{1}_{\mathbf{S}}$ (resp. $(v^\star, w^\star)$) on $[-1,1]^n$ with degree d polynomials w_d (resp. and v_d) using Jackson-type inequalities in the appropriate metric

(2) "Effective Slater condition": Perturb w_d (resp. and v_d) into a strictly feasible polynomial $\tilde{w}_d = w_d + \theta_d \hat{w}$ (resp. and $\tilde{v}_d = v_d + \theta_d \hat{v}$) of degree d for some well chosen inward pointing polynomial $\hat{w}$ (resp. and $\hat{v}$)

(3) "Effective Putinar P-satz": Apply (4) to the positive polynomials $\tilde{w}_d$ on $\mathbf{B}$ and $p_d(\boldsymbol{x}) = \tilde{w}_d(\boldsymbol{x}) - 1$ on $\mathbf{S}$ (resp. and $p_d(\boldsymbol{x}) = \tilde{w}_d(\boldsymbol{x}) + \Delta\tilde{v}_d(\boldsymbol{x}) - 1$ on $\mathbf{S}$ and $q_d(\boldsymbol{x}) = \nabla g(\boldsymbol{x})^\top \nabla \tilde{v}_d(\boldsymbol{x})$ on $\partial\mathbf{S}$)
(4) Conclusion: Tune the value of parameter θ_d to invert (4b) into an approximation error / convergence rate.

Denoting by $W_r^\star$ (resp. $V_r^\star$) the optimal value of the level $r \in \mathbb{N}$ SoS programming approximation of (1b) (resp. (3)), the above procedure returns the following convergence rates:

$$W_r^\star - \mathrm{vol}(\mathbf{S}) \in \underset{r\to\infty}{\mathcal{O}}\left(r^{-\frac{1}{6c}}\right) \tag{5a}$$

$$V_r^\star - \mathrm{vol}(\mathbf{S}) \in \underset{r\to\infty}{\mathcal{O}}\left(r^{-\frac{1}{(2.5+\varepsilon)c}}\right) \quad \forall \varepsilon > 0 \tag{5b}$$

for a fixed constant $c > 0$ depending only on the dimension and geometry of $\mathbf{S}$. In other words, the addition of Stokes constraints results in more than squaring the convergence speed of the moment-SoS hierarchy.

References

[1] B. Büeler, et al., *Exact volume computation for polytopes: a practical study*, in G. Kalai and G. M. Ziegler (eds.), *Polytopes - Combinatorics and Computation*, DMV Seminar, Birkhäuser, **29** (2000), 131–154.
[2] B. Cousins and S. Vempala, *A practical volume algorithm*, Math. Program. Comput., **8** (2016), 133–160.
[3] Z. B. Zabinsky and R. L. Smith, *Hit-and-Run methods*, in S. Gass and M. C. Fu (eds.), *Encyclopedia of Operations Research and Management Science*, (2013).
[4] P. Lairez, et al., *Computing the volume of compact semi-algebraic sets*, Int. Symp. Symb. Alg. Comput., (2019).
[5] D. Henrion, et al., *Approximate volume and integration for basic semialgebraic sets*, SIAM Rev., **51** (2009), 722–743.
[6] D. Henrion, et al., *The Moment-SOS hierarchy*, Series on Optimization and its Applications, World Scientific, (2020).
[7] J. B. Lasserre, *Computing Gaussian and exponential measures of semi-algebraic sets*, Adv. Appl. Math., **91** (2017), 137–163.
[8] M. Tacchi, et al., *Stokes, Gibbs and volume computation of semi-algebraic sets*, Discrete Comput. Geom., **69** (2023), 260–283.
[9] C. Schlosser, et al., *Convergence rates for the moment-SoS hierarchy*, arXiv preprint arXiv:2402.00436 (2024).
[10] L. Baldi and B. Mourrain, *On the effective Putinar's Positivstellensatz and moment approximation*, Math. Program., **200** (2023), 71–103.

From POP to POC with the Moment-SOS Hierarchy

Didier Henrion

The purpose of this tutorial talk was to explain how the moment-sums of squares (mom-SOS) hierarchy can solve globally polynomial optimal control (POC), i.e. optimal control problems with polynomial data and semi-algebraic state and control constraints.

The first part was a quick introduction to mom-SOS for polynomial optimization (POP) (i.e. without differential equations and control), emphasizing the primal-dual linear formulation on measures and continuous functions, and explaining how the moment resp. positive polynomial cones can be approximated with semi-definite representable pseudo-moments resp. polynomial SOS cones.

The second part described mom-SOS for POC, emphasizing the similarities and differences with POP, introducing relaxed controls (aka Young measures) and the Liouville equation on occupation measures modeling the optimal transport (and conservation) of mass along the controlled flow, culminating with the L1 norm convergence results to the lower semi-continuous value function.

Moment-SOS Methods for Variational Problems and PDEs

Giovanni Fantuzzi

Moment-sum-of-squares (moment-SOS) hierarchies are a powerful approach to solve polynomial optimization problems (POPs) on compact subsets of $\mathbb{R}^n$ [23, 34] and to predict the dynamics of systems governed by ordinary differential equations (ODEs) with polynomial vector fields (see, e.g., [36, 25, 35, 19, 11, 22, 39, 13, 38]). This motivated a number of recent attempts to extend moment-SOS hierarchies to infinite-dimensional problems from the calculus of variations, as well as to dynamic problems constrained by time-dependent partial differential equations (PDEs). Advances have appeared independently in the optimal control, fluid mechanics, and applied analysis communities, often using dual frameworks and with little cross-fertilization. In this talk (and this extended abstract) we reviewed some of these advances, in order to elucidate similarities, differences, and open questions.

1. Moment-SOS methods for variational problems

The first part of the talk discussed applications of polynomial optimization to static problems in the calculus of variations, which take the form

$$f^* := \min_{\substack{u:\Omega\to\mathbb{R}^m \\ u=g \text{ on } \partial\Omega}} \int_\Omega f(x,u(x),Du(x))\,dx. \tag{1}$$

Here, $\Omega \subset \mathbb{R}^n$ is an open and bounded Lipschitz domain, $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ and $g : \Omega \times \mathbb{R}^m$ are given polynomial functions, and $Du : \Omega \to \mathbb{R}^{m\times n}$ is the usual matrix of partial derivatives of u. We assume a minimizer to exist in a suitable Sobolev space, which is true under standard quasiconvexity, coercivity and growth assumptions on f [9, §3.2.6]. We do not however assume that the problem is convex. We seek to find the minimum value f^* and an associated optimal map u^* using polynomial optimization.

1.1. **Discretization-relaxation strategies.** One approach to attack problem (1) is to first discretize it into a large but *sparse* finite-dimensional POP, and then approximate minimizers of this POP using sparse moment-sum-of-squares (moment-SOS) hierarchies. This strategy was pioneered in [27, 29, 28], where finite differences were applied to generalizations of (1) including PDE constraints. If finite

differences are replaced by finite elements, moreover, then moment-SOS hierarchies produce convergent approximations to minimizer $u^*$for variational problems constrained only by boundary conditions and with a *unique* global minimizer [12]. It remains an open problem to remove this restrictive assumption and extend the convergence analysis to variational problems with PDE constraints.

1.2. **Null Lagrangians and occupation measures.** A key limitation of the discretize-and-relax approach reviewed above is that it does not give certified lower bounds on f^*. For this reason, following ideas from [40, 41, 1, 2], in [7] polynomial optimization was combined with *null Lagrangians* to bound f^* from below. Set $\mathcal{X} := \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$. A continuous function $\Psi : \mathcal{X} \to \mathbb{R}$ is a *null Lagrangian*, written as $\Psi \in \mathrm{NL}$, if there exists a vector field $\varphi : \mathcal{X} \to \mathbb{R}^n$ such that $\Psi(x, u(x), Du(x)) = \nabla \cdot \varphi(x, u(x), Du(x))$ for every Sobolev function u. By the divergence theorem, this means that the integral functional

$$L_\Psi[u] = \int_\Omega \Psi(x, u(x), Du(x))\, dx \tag{2}$$

depends only on the boundary values of u. If Ψ is a null Lagrangian and u satisfies the boundary condition $u = g$ on $\partial\Omega$, therefore, one can write

$$\int_\Omega f(x, u, Du)\, dx = \int_\Omega (f - \Psi)(x, u, Du)\, dx + L_\Psi[g]$$

and deduce that

$$f^* \geq \sup_{\substack{\Psi\in\mathrm{NL} \\ f-\Psi\geq 0 \text{ on } \mathcal{X}}} \int_\Omega \Psi(x, g, Dg)\, dx. \tag{3}$$

The maximum on the right-hand side of this inequality can be estimated from below using polynomial optimization when f is polynomial, the integration domain Ω is a basic semialgebraic set, the optimization is restricted to polynomial Ψ, and the inequality $f - \Psi \geq 0$ on $\mathcal{X}$ is strengthened into a sum-of-squares constraint.

Independently, [21] (see also [4] for an earlier formulation of the same framework with no numerical implementation) proposed to lift problem (1) into a linear minimization problem over *occupation measures*. In the present context, these are measures supported on $\mathcal{X}$ such that

$$\langle \Psi, \mu \rangle = \int_\Omega \Psi(x, g(x), Dg(x))\, dx \tag{4}$$

for every null Lagrangian Ψ, where angled brackets denote integration against a measure. Condition (4) holds for all measures μ obtained by pushing forward the Lebesgue measure on Ω by maps $x \mapsto (x, u(x), Du(x))$ associated to Sobolev functions u that satisfy $u = g$ on $\partial\Omega$. Thus,

$$f^* \geq \inf_{\substack{\mu\in\mathcal{M}(\mathcal{X}) \\ \langle\Psi,\mu\rangle = L_\Psi[g]\ \forall \Psi\in\mathrm{NL}}} \langle f, \mu \rangle. \tag{5}$$

If the constraint $\langle \Psi, \mu \rangle = L_\Psi[g]$ is imposed only for polynomial null Lagrangians, the minimization problem on the right-hand side of this inequality is a generalized

moment problem. Its optimal value, and thus f^*, can then be estimated from below using standard moment-SOS hierarchies.

Although they have been proposed independently, the maximization over null Lagrangians in (3) and the minimization over occupation measures in (5) are essentially equivalent. Precisely, the weak duality inequality

$$\inf_{\substack{\mu\in\mathcal{M}(\mathcal{X})\\ \langle\Psi,\mu\rangle=L_\Psi[g]\ \forall\Psi\in\mathrm{NL}}} \langle f,\mu\rangle \geq \sup_{\substack{\Psi\in\mathrm{NL}\\ f-\Psi\geq 0 \text{ on } \mathcal{X}}} \int_\Omega \Psi(x,g,Dg)\,dx \tag{6}$$

is always true, while strong duality (meaning the two sides of this inequality are in fact equal) holds under mild conditions [14]. This is true even for variational problems more general than (1), where the map u is subject to additional pointwise or integral constraints.

What remains to be determined are necessary and sufficient conditions ensuring that the lower bounds in (3) and (5) are equal to the global minimum f^* of (1). So far, equality is known to hold for convex variational problems [14, 20] and some nonconvex ones [7], but not in general [14]. A complete understanding of the nonconvex case, however, is still lacking.

2. Moment-SOS methods for PDE dynamics

Moment-SOS hierarchies have also been used to study dynamical systems governed PDEs. Particular problems that have been considered include approximating PDE solutions for a given initial condition [26, 18], proving the stability of equilibria [17, 8, 15], and bounding time averages [37, 16].

One way to approximate PDE solutions with polynomial optimization is to view the PDE as a constraint in a static variational problem over the PDE's space-time domain, and apply the moment-SOS methods described in §1. Occupation measure frameworks for such problems are sharp for linear PDEs [5] and for nonlinear scalar hyperbolic PDEs [26, 6]. For other classes of PDEs, including parabolic ones, computations give good results [41, 1, 2] but no rigorous theory is available.

A different line of attack is to view a time-dependent PDE on a time interval $[0,T]$ as an abstract evolution equation

$$\frac{d}{dt}u(t) = F(u(t)) \tag{7}$$

on a separable Hilbert space $\mathcal{H}$. To make sense of this, one usually introduces a countable basis $\varphi_0,\varphi_1,\dots$ for $\mathcal{H}$, a Banach space $\mathcal{V}$ with continuous embeddings $\mathcal{V}\hookrightarrow\mathcal{H}\hookrightarrow\mathcal{V}'$ (primes denote dual spaces), and views (7) as the infinite set of equations

$$\frac{d}{dt}\langle\varphi_k,u(t)\rangle = \langle\varphi_k,F(u(t))\rangle \quad \forall k=0,1,\dots \tag{8}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $\mathcal{V}$ and $\mathcal{V}'$. We call this process a *Galerkin expansion* and the time-dependent quantities $\langle\varphi_k,u(t)\rangle$ the *Galerkin coefficients*. For example, if (7) is a second-order parabolic PDE on the periodic interval $\Omega=(-\pi,\pi)$, one can take $\mathcal{H}$ as the space of periodic functions in $H^1(\Omega)$,

$\mathcal{V}$ as the space of periodic functions in $L^2(\Omega)$, and $\varphi_k(x) = \mathrm{e}^{ikx}$, so (8) describes the evolution of the time-dependent Fourier coefficients of the PDE solution u.

Having posed the PDE as an "infinite-dimensional ODE", one can attempt to mimic the established moment-SOS hierarchies for ODEs. The evolution of the Galerkin coefficients is thus characterized using the moments of an occupation measure supported on $[0,T] \times \mathcal{H}$,

$$\int_{[0,T]\times\mathcal{H}} t^{\alpha_0} \langle \varphi_{k_1}, u\rangle^{\alpha_1} \cdots \langle \varphi_{k_m}, u\rangle^{\alpha_m} \, d\mu(t,u), \tag{9}$$

and similarly defined moments of a *terminal measure* μ_T that is supported on $\mathcal{H}$ and encodes the PDE solution at the final time T. On the dual side, one uses functionals that depend polynomially on time and on the Galerkin coefficients. This approach was recently followed by [18], who provide sufficient conditions for infinite sequences to be moments of a measure-valued solution μ of the original PDE, and previously by [37], who use functionals of the Galerkin coefficients to characterize *weak stationary statistical solutions*—a type of measure-valued PDEs solutions that arise when studying time averages. It must be noted, however, that measure-valued PDE solutions need not coincide with more traditional notions of solutions; we refer to the work by [37] for further discussion.

The idea of studying PDE solutions using functionals of the Galerkin coefficients also underpins recent computational works aimed at proving stability and bounding time averages in fluid mechanics [17, 8, 15, 16]. The distinguishing feature of these works is that they attempt to address a key problem with the Galerkin expansion (8): the equation for a Galerkin coefficient, in general, depends on *infinitely many* other coefficients. Implementations of moment-SOS hierarchies that simply truncate the equations to a finite number of Galerkin coefficients, therefore, only produce results that apply to a finite-dimensional Galerkin *projection* of the original PDE, rather than the PDE itself. This issue affects the computations by [18] and some of those by [16], where the convergence of results for increasingly large Galerkin projections is observed but not rigorously proved. To avoid projection errors, [17] expand the PDE solution as

$$u(t) = \sum_{k=1}^{N} \langle \varphi_k, u(t)\rangle \, \varphi_k + v(t) \tag{10}$$

and replace (8) with the N equations for the Galerkin coefficients $\langle \varphi_k, u(t)\rangle$, plus an equation for the norm $\|v\|$ of the expansion "tail". Terms in these equations that depend on the unknown function v are estimated in terms of $\langle \varphi_1, u\rangle, \ldots, \langle \varphi_N, u\rangle$ and $\|v\|$, leading to a finite-dimensional but *uncertain* ODE system of $N+1$ equations. Any statements about the dynamics of this uncertain system applies automatically to the original PDE dynamics. This approach has provided novel and highly nontrivial results in fluid mechanics [15, 16], but relies on PDE-dependent estimates whose quality can considerably affect the results. There are also no

proofs that the uncertain system reduces to the original PDE as $N \to \infty$, although computations suggest that it does. Rigorous convergence results remain to be established.

References

[1] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, *Dissipation inequalities for the analysis of a class of PDEs*, Automatica, **66** (2016), 163–171.

[2] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, *Safety verification for distributed parameter systems using barrier functionals*, Sys. Control Lett., **108** (2017), 33–39.

[3] E. Aranda and R. J. Meziat, *The method of moments for some one-dimensional, non-local, non-convex variational problems*, J. Math. Anal. Appl., **382(1)** (2011), 314–323.

[4] R. Awi and W. Gangbo, *A polyconvex integrand; Euler-Lagrange equations and uniqueness of equilibrium*, Arch. Ration. Mech. Anal., **214(1)** (2014), 143–182.

[5] D. Bertsimas and C. Caramanis, *Bounds on linear PDEs via semidefinite optimization*, Math. Program., **108(1, Ser. A)** (2006), 135–158.

[6] C. Cardoen, S. Marx, A. Nouy, and N. Seguin, *A moment approach for entropy solutions of parameter-dependent hyperbolic conservation laws*, arXiv:2307.10043 [math.OC], (2023).

[7] A. Chernyavsky, J. J. Bramburger, G. Fantuzzi, and D. Goluskin, *Convex relaxations of integral variational problems: pointwise dual relaxation and sum-of-squares optimization*, SIAM J. Optim., **33(2)** (2023), 481–512.

[8] S. I. Chernyshenko, P. Goulart, D. Huang, and A. Papachristodoulou, *Polynomial sum of squares in fluid dynamics: a review with a look ahead*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., **372(2020)** (2014), 20130350, 18.

[9] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Springer New York, (2008).

[10] J. J. Egozcue, R. Meziat, and P. Pedregal, *From a nonlinear, nonconvex variational problem to a linear, convex formulation*, Appl. Math. Optim., **47(1)** (2003), 27–44.

[11] G. Fantuzzi, D. Goluskin, D. Huang, and S. I. Chernyshenko, *Bounds for deterministic and stochastic dynamical systems using sum-of-squares optimization*, SIAM J. Appl. Dyn. Syst., **15(4)** (2016), 1962–1988.

[12] G. Fantuzzi and F. Fuentes, *Global minimization of polynomial integral functional*, arXiv:2305.18801, (2023).

[13] G. Fantuzzi and D. Goluskin, *Bounding extreme events in nonlinear dynamics using convex optimization*, SIAM J. Appl. Dyn. Syst., **19(3)** (2020), 1823–1864.

[14] G. Fantuzzi and I. Tobasco, *Sharpness and non-sharpness of occupation measure bounds for integral variational problems*, arXiv:2207.13570, (2023).

[15] F. Fuentes, D. Goluskin, and S. I. Chernyshenko, *Global stability of fluid flows despite transient growth of energy*, Phys. Rev. Lett., **128(20)** (2022), Paper 204502, 6.

[16] D. Goluskin and G. Fantuzzi, *Bounds on mean energy in the Kuramoto–Sivashinsky equation computed using semidefinite programming*, Nonlinearity, **32(5)** (2019), 1705–1730.

[17] P. J. Goulart and S. I. Chernyshenko, *Global stability analysis of fluid flows using sum-of-squares*, Phys. D, **241(6)** (2012), 692–704.

[18] D. Henrion, M. Infusino, S. Kuhlmann, and V. Vinnikov, *Infinite-dimensional moment-sos hierarchy for nonlinear partial differential equations*, arXiv:2305.18768, (2023).

[19] D. Henrion and M. Korda, *Convex computation of the region of attraction of polynomial control systems*, IEEE Trans. Automat. Control, **59(2)** (2014), 297–312.

[20] D. Henrion, M. Korda, M. Kružík, and R. Rios-Zertuche, *Occupation measure relaxations in variational problems: the role of convexity*, arXiv:2303.02434, (2023).

[21] M. Korda, D. Henrion, and J. B. Lasserre, *Moments and convex optimization for analysis and control of nonlinear PDEs*, in E. Trelat and E. Zuazua (eds.), Handbook of Numerical Analysis – Numerical Control: Part A, chapter 10, 339–366, (2022).

[22] J. Kuntz, M. Ottobre, G. B. Stan, and M. Barahona, *Bounding stationary averages of polynomial diffusions via semidefinite programming*, SIAM J. Sci. Comput., **38(6)** (2016), A3891–A3920.
[23] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., **11(3)** (2001), 796–817.
[24] J. B. Lasserre, *Convergent SDP-relaxations in polynomial optimization with sparsity*, SIAM J. Optim., **17(3)** (2006), 822–843.
[25] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, *Nonlinear optimal control via occupation measures and LMI-relaxations*, SIAM J. Control Optim., **47(4)** (2008), 1643–1666.
[26] S. Marx, T. Weisser, D. Henrion, and J. B. Lasserre, *A moment approach for entropy solutions to nonlinear hyperbolic PDEs*, Math. Control Relat. Fields, **10(1)** (2020), 113–140.
[27] M. Mevissen, M. Kojima, J. Nie, and N. Takayama, *Solving partial differential equations via sparse SDP relaxations*, Pac. J. Optim., **4(2)** (2008), 213–241.
[28] M. Mevissen, J. B. Lasserre, and D. Henrion, *Moment and SDP relaxation techniques for smooth approximations of problems involving nonlinear differential equations*, IFAC Proc. Vol., **44(1)** (2011), 10887–10892.
[29] M. Mevissen, K. Yokoyama, and N. Takayama, Solutions of polynomial systems derived from the steady cavity flow problem, *Proc. 2009 Int. Sympos. Symb. Algebr. Comput.*
[30] R. Meziat and D. Patiño, Exact relaxations of non-convex variational problems, *Optim. Lett.*, **2(4)** (2008), 505–519.
[31] R. Meziat, T. Roubíček, and D. Patiño, Coarse-convex-compactification approach to numerical solution of nonconvex variational problems, *Numer. Funct. Anal. Optim.*, **31(4-6)** (2010), 460–488.
[32] R.J. Meziat and J. Villalobos, Analysis of microstructures and phase transition phenomena in one-dimensional, non-linear elasticity by convex optimization, *Struct. Multidiscip. Optim.*, **32(6)** (2006), 507–519.
[33] R. Meziat, Analysis of two dimensional nonconvex variational problems, in X.Y. Liqun Qi, Koklay Teo (eds.), *Optimization and Control with Applications*, Springer-Verlag (2005), 393–405.
[34] P.A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Math. Program. B*, **96(2)** (2003), 293–320.
[35] P.A. Parrilo, Polynomial optimization, sums of squares, and applications, in *Semidefinite Optimization and Convex Algebraic Geometry*, vol. 13 of *MOS-SIAM Ser. Optim.*, SIAM, Philadelphia, PA (2013), 47–157.
[36] S. Prajna, Barrier certificates for nonlinear model validation, *Automatica J. IFAC*, **42(1)** (2006), 117–126.
[37] R.M.S. Rosa and R.M. Temam, Optimal minimax bounds for time and ensemble averages for the incompressible Navier-Stokes equations, *Pure Appl. Funct. Anal.*, **7(1)** (2022), 327–355.
[38] C. Schlosser and M. Korda, Converging outer approximations to global attractors using semidefinite programming, *Automatica J. IFAC*, **134** (2021), Paper 109900, 9 pp.
[39] I. Tobasco, D. Goluskin, and C.R. Doering, Optimal bounds and extremal trajectories for time averages in nonlinear dynamical systems, *Phys. Lett. A*, **382(6)** (2018), 382–386.
[40] G. Valmorbida, M. Ahmadi, and A. Papachristodoulou, Semi-definite programming and functional inequalities for distributed parameter systems, in *53rd IEEE Conference on Decision and Control*, IEEE (2014).
[41] G. Valmorbida, M. Ahmadi, and A. Papachristodoulou, Stability analysis for a class of partial differential equations via semidefinite programming, *IEEE Trans. Automat. Control*, **61(6)** (2016), 1649–1654.
[42] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity, *SIAM J. Optim.*, **17(1)** (2006), 218–242.

Studying Stability of a Shear Flow using Sum-of-Squares Polynomial Optimization

Elizabeth Carlson

Determining nonlinear stability of steady states for complex dynamical systems are notoriously difficult problems, even for the seemingly simplest cases. For example, it is expected that the standard steady state shear profile, 2D planar Couette flow, is globally stable for all Reynolds numbers, however the state-of-the-art analysis is decades old and only proves the stability for relatively low Reynolds numbers. In recent years, a promising computational approach uses polynomial sum-of-squares optimization to find Lyapunov functions based on low-mode projections onto an orthogonal basis of $L^2 \cap H^1$. We will present a few extensions on existing work in this area.

Going beyond ODE Systems with the Moment-SOS Hierarchy

Jared Miller

The moment-SOS hierarchy is a computational tool that successively outer approximates an infinite-dimensional linear program in measures (or continuous functions) into a sequence of semidefinite programs [1]. The moment-SOS hierarchy has applications in the realm of dynamical systems, such as in synthesis of Lyapunov functions, optimal control [2], bounding long-time averages [3], peak estimation [4], and reachable set estimation [5].

1. Optimal Control

The general template for an optimal control problem (OCP) is:

$$V^*(\cdot) = \inf_u \int_{t_0}^{T} [\text{Running cost}(t)]dt + \text{Terminal Cost} \tag{1a}$$

$$\text{Dynamics are obeyed} \tag{1b}$$

$$u \text{ is a valid control} \tag{1c}$$

$$\text{State constraints (initial, running, final)} \tag{1d}$$

For a continuous-time ordinary differential equation (ODE), the OCP in (1) (with state space X, terminal set X_T, input class $\mathcal{U}$, running cost J, and terminal cost J_T) is

$$V^*(t_0, x_0) = \inf_u \int_{t_0}^{T} J(t, x(t), u(t))dt + J_T(x(T)) \tag{2a}$$

$$\dot{x} = f(t, x(t), u(t)) \qquad \forall t \in [t_0, T] \tag{2b}$$

$$u(t) \in \mathcal{U} \qquad \forall t \in [t_0, T] \tag{2c}$$

$$x(t_0) = x_0,\ x(T) \in X_T \tag{2d}$$

The objective value term $V^*(t_0, x_0)$ is the *value function* of the OCP. The Value function satisfies the Hamilton-Jacobi Bellman equations [6]:

$$V^*(T, x) = J_T(x) \text{ in } X_T \tag{3a}$$

$$\inf_u (\partial_t + f \cdot \nabla_x) V(t, x, u) + J(t, x, u) = 0 \tag{3b}$$

The value function V^* can be discontinuous. Linear-Program methods for optimal control relax the inequalities in (3) into inequalities, forming an optimization problem with respect to an initial distribution $\mu_0 \in \mathcal{M}_+(X)$. When the admissible input class $\mathcal{U}$ is the set of functions whose graph is restricted to $[0, T] \times U$ (for some set U), the inequality program for (3) is

(4a)

$$d^*(t_0, \mu_0) = \inf_v \int_{x_0} v(t_0, x_0) d\mu_0(x_0)$$

$$v(T, x) \leq J_T(x) \qquad \forall x \in X_T \tag{4b}$$

$$(\partial_t + f \cdot \nabla_x) v(t, x, u) + J(t, x, u) \geq 0 \quad \forall (t, x, u) \in [0, T] \times X \times U \tag{4c}$$

$$v \in C^1([0, T] \times X). \tag{4d}$$

Any function v satisfying the constraints of (4) is a point-wise subvalue function: $v(t_0, x_0) \leq V^*(t_0, x_0)$. Under conditions of compactness, lower-semicontinuity of J, and convexity in cost and dynamics [7], problems (4) and (3) will take equal values. Application of the moment-SOS hierarchy follows by restricting v as $v \in \mathbb{R}[t, x]$ and replacing each functional inequality constraint in (4) with a Putinar Positivestellensätz verification of nonnegativity [8].

The linear programming method can be extended to non-ODE dynamical systems. In particular, this presentation focused on three main classes:

(1) Discrete-time systems
(2) Hybrid systems
(3) Stochastic systems

2. Discrete-Time systems

The dynamical law obeyed by trajectories of a dynamical system is described by a generator, which returns the incremental change of a test function observed along dynamical trajectories. As a point of comparison, the continuous-time and discrete-time generators are

$$\text{Continuous-Time:} \quad \dot{x} = f(t, x, u) \qquad v \mapsto (\partial_t + f \cdot \nabla_x) v \tag{5}$$

$$\text{Discrete-Time:} \quad x_+ = R(t, x, u) \qquad v \mapsto v(t + 1, R(t, x, u)) - v(t, x) \tag{6}$$

Substitution of the generator in (6) instead of the Lie derivative in (4c) will render (4) the formulation for a discrete-time OCP. If the dynamics functions f, R are both polynomials, then the following degree bounds are observed

$$\deg(\partial_t + f \cdot \nabla_x) v = \deg v + \deg f - 1$$

$$\deg(v(t + 1, R(t, x, u)) - v(t, x)) = \deg v * \deg R$$

If instead a new lifting variable $\tilde{x} = R(t, x, u)$ is added [9], a lower polynomial degree is observed at the cost of increasing the number of variables (t, x, u) to $(t, x, \tilde{x}, u)$:

$$\deg(v(t+1, \tilde{x}) - v(t, x)) = \deg v$$

3. Hybrid Systems

A hybrid system [10] is a dynamical system that exhibits continuous dynamics with discrete transitions. Examples include walker robots, gearboxes in cars, sampled-data systems, and billiards. Dynamics evolve in a sequence of *modes* modes $\ell \in 1..L$. The modes are connected together by arcs $e = (\ell, \ell')$ defining allowable transitions. Continuous-time dynamics evolve in each mode according to

$$\dot{x}_\ell(t) = f_\ell(t, x(t), u(t)) \tag{7a}$$

Until a *guard surface* S_e is encountered. Upon hitting a guard surface, the system executes a discrete transition based on the reset map R_e at (e.g. ground)

$$x_{\mathrm{dst}(e)}(t^+) = R_e(t, x_{\mathrm{src}(e)}(t), u_e(t)). \tag{7b}$$

Further assumptions of transversality and determinism of switching are usually necessary to rule out Zeno or ambiguous executions.

An optimal control problem posed over hybrid dynamics with stage costs J_ℓ, terminal costs J_T, and reset costs G_e has an LP formulation with per-mode variables $\{v_\ell\}$ [11]:

$$d^*(t_0, x_0, \ell_0) v_{\ell_0}(t_0, x_0) \tag{8a}$$

$$(\partial_t + f_\ell \cdot \nabla_x) v_\ell(t, x_\ell, u) + J_\ell(t, x_\ell, u) \geq 0 \tag{8b}$$

$$J_{T\ell}(x_\ell) \geq v_\ell(T, x_\ell) \tag{8c}$$

$$G_e(t, x_{\mathrm{src}(e)}, u_e) \geq v_{\mathrm{src}(e)}(t, x_{\mathrm{src}(e)}) - v_{\mathrm{dst}(e)}(t, R_e(t, x_{\mathrm{src(e)}, u_e})) \tag{8d}$$

$$\forall \ell \in 1..L : \; v_\ell \in C^1([0, T] \times X_\ell). \tag{8e}$$

4. Stochastic Systems

A stochastic process is a time-indexed collection of probability distributions $\{\mu_t\}$ [12] related together by a dynamical law. The dynamical law is represented by a *generator* with increment Δt. The generator is a linear operator $\mathcal{L}_\tau$ such that the following relation holds for all test functions:

$$\mathcal{L}_\tau v = \lim_{\tau' \to \tau} (\langle v(t + \tau', x), \mu_{t+\tau'} \rangle - \langle v(t, x), \mu_t \rangle)/\tau'. \tag{9}$$

The Lie derivative in (4c) is the instance of a generator for the ODE with time increment $\Delta t = 0$. Other examples of generators include an expression for a Stochastic Differential Equation:

$$\mathcal{L} v(t, x) = \partial_t v + f(t, x) \cdot \nabla_x v + \frac{1}{2} g(t, x)^T \left(\nabla^2_{xx} v \right) g(t, x). \tag{10}$$

The law and generator for a discrete-time random process with parameter λ drawn from a distribution ξ is

$$x[t+\Delta t] = f(t, x[t], u[t], \lambda[t]), \qquad \lambda[t] \sim \xi \tag{11}$$

$$\mathcal{L}_{\Delta t} v = \frac{\int_\Lambda v(t+\Delta t, f(t,x,u,\lambda))d\xi(\lambda) - v(t,x}{\Delta t}. \tag{12}$$

Substituting these generators into the expression in (4) results in a formulation for a stochastic control program that minimizes the expectation of the costs J and J_T. Similarly, stochastic dynamics can be added in a hybrid setting through the application of continuous-time generators $(\mathcal{L}_\ell)$ and discrete-transition generators (R_ℓ) [13].

References

[1] J.B. Lasserre, *Moments, Positive Polynomials and Their Applications*, Imperial College Press, London, 2009.

[2] D. Henrion, J.B. Lasserre, and C. Savrognan, Nonlinear optimal control synthesis via occupation measures, in *47th IEEE Conference on Decision and Control*, (2008), 4749–4754.

[3] I. Tobasco, D. Goluskin, and C.R. Doering, Optimal bounds and extremal trajectories for time averages in nonlinear dynamical systems, *Phys. Lett. A*, **382(6)** (2018), 382–386.

[4] G. Fantuzzi and D. Goluskin, Bounding extreme events in nonlinear dynamics using convex optimization, *SIAM J. Appl. Dyn. Syst.*, **19(3)** (2020), 1823–1864.

[5] D. Henrion and M. Korda, Convex computation of the region of attraction of polynomial control systems, *IEEE Trans. Automat. Control*, **59(2)** (2014), 297–312.

[6] D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction*, Princeton University Press, Princeton, NJ, 2011.

[7] R.M. Lewis and R.B. Vinter, Relaxation of optimal control problems to equivalent convex programs, *J. Math. Anal. Appl.*, **74(2)** (1980), 475–493.

[8] M. Putinar, Positive polynomials on compact semi-algebraic sets, *Indiana Univ. Math. J.*, **42(3)** (1993), 969–984.

[9] J. Miller and M. Sznaier, Analysis and control of input-affine dynamical systems using infinite-dimensional robust counterparts, *arXiv preprint*, arXiv:2112.14838 (2023).

[10] R. Goebel, R.G. Sanfelice, and A.R. Teel, Hybrid dynamical systems, *IEEE Control Syst. Mag.*, **29(2)** (2009), 28–93.

[11] P. Zhao, S. Mohan, and R. Vasudevan, Optimal control of polynomial hybrid systems via convex relaxations, *IEEE Trans. Automat. Control*, **65(5)** (2019), 2062–2077.

[12] C. Gardiner, *Stochastic Methods*, 4th ed., Springer, Berlin, 2009.

[13] O. Costa and F. Dufour, A linear programming formulation for constrained discounted continuous control for piecewise deterministic Markov processes, *J. Math. Anal. Appl.*, **424(2)** (2015), 892–914.

Wrinkling of confined elastic shells

Ian Tobasco

We discuss a problem in nonlinear elasticity, regarding the wrinkling of a confined elastic shell. In the absence of external forces, a shell prefers to be curved. One can exploit this to make a wrinkle pattern, by forcing a shell to reside nearby a plane. Panel (A) of Figure 1 displays two experiments and corresponding simulations in which a square is cut out from a spherical or saddle shell, and placed on a water

bath. Capillary forces cause the shells to adhere to the bath, and buoyancy forces attempt to pull them flat, making a wrinkle pattern. Figure 2 shows the patterns that arise by altering the cut out shape.

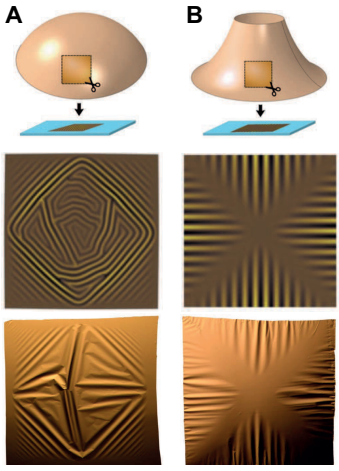
Can these wrinkle patterns be predicted? The papers [2, 1] explain that the answer is "yes", and in fact using a simple set of geometric rules. For one, the wrinkles of negatively curved shells follow the paths of quickest exit from the cut out shape ($\kappa < 0$ in the figures). Such paths are line segments that meet the boundary at right angles, and meet each other at the *medial axis* drawn in the $\kappa < 0$ half of Figure 2 in white. Surprisingly, the medial axis also explains the patterns of positively curved shells ($\kappa > 0$). The explanation is in Panel (B) of Figure 1. Most points p on the medial axis have exactly two closest boundary points, called q and r in the top half of the figure. The line segments pq and pr lie along the negatively curved wrinkles. The line segment qr lies along the wrinkles of the "reciprocal" positively curved shell, obtained by sending $\kappa \to -\kappa$ while fixing the cut out shape. In the special case that the point p on the medial axis has three or more closest boundary points (q, r, s, t in the bottom half of the figure), the wrinkles are "disordered". Experimentally, such regions give wrinkles that rearrange between trials, as well as in response to perturbations. While we do not yet know how precisely characterize such disordered wrinkles, the sets where they occur are completely determined.

These simple rules for the wrinkling of confined shells are rigorously derived in [2, 1] as statements about minimizers of an energy functional. The starting point is a "geometrically linearized" model for a floating shallow shell. Suppose the initial cut out shell is the graph of a function $p(x)$ defined for $x = (x_1, x_2)$ belonging to a reference domain $\Omega \subset \mathbb{R}^2$. Then, its deformed configuration can be found by minimizing the energy (written here with Poisson's ratio $\nu = 0$ for simplicity)

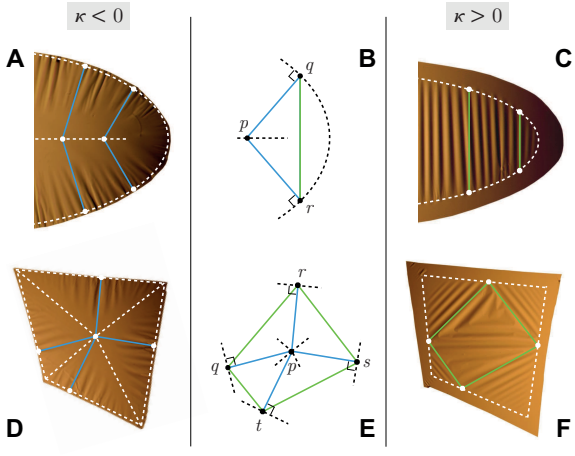
$$(1) \quad E(u,w) = \int_\Omega \frac{1}{2}|e(u) + \frac{1}{2}\nabla w \otimes \nabla w - \frac{1}{2}\nabla p \otimes \nabla p|^2 + \frac{b}{2}|\nabla\nabla w|^2 + \frac{k}{2}|w|^2 \, dx.$$

Here, $u = (u_1(x), u_2(x))$ and $w(x)$ capture the displacements of the shell parallel and perpendicular to the water bath, and $e(u)$ is the symmetric gradient of u given by $e_{ij}(u) = (\partial_i u_j + \partial_j u_i)/2$ for $i = 1, 2$. The first two terms in (1) are the elastic energy. The last term accounts for buoyancy, which prefers $w \to 0$. The (non-dimensional) parameters b and k are proportional to the squared shell thickness and the gravitational acceleration of the water, respectively. Taking $b \to 0$ corresponds to studying a sequence of ever thinner shells; taking $k \to \infty$ ensures the shells are pulled nearly flat. Similar energy functionals have been used in the surrounding literature (see, e.g., [3, 4] or [5] for a general introduction).

Given the energy (1), the question is to characterize its minimizers as $b \to 0$ and $k \to \infty$. To state a theorem, we work in the asymptotic regime $k^{-3/2} \ll b \ll k^{-1}$ and assume that $p \in W^{2,\infty}(\Omega)$ where $\Omega \subset \mathbb{R}^2$ is Lipschitz and strictly star-shaped.



(A)



(B)

FIGURE 1. Elastic shells wrinkle upon a water bath [1]. (A) shows squares cut from a spherical or saddle shell. Finite element simulations and physical experiments appear in that order. (B) depicts our rules for the wrinkles of positively ($\kappa > 0$) and negatively ($\kappa < 0$) curved shells. The top half addresses ordered wrinkles. The bottom half locates the disordered parts. Medial axes and boundaries used for the predictions are dotted.

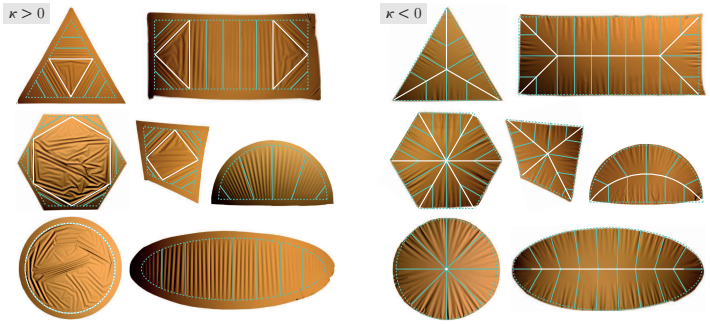


FIGURE 2. Predicted wrinkle patterns overlaid on experiments of floating shells [1]. On the left are positively curved shells ($\kappa > 0$); on the right are negatively curved shells ($\kappa < 0$). Regions with cyan lines are predicted to have orderly wrinkles with peaks and troughs along the lines. Regions absent these lines (bordered in white, $\kappa > 0$) exhibit a statistical response. Dotted lines show the ideal shapes used for the predictions.

Then, the minimum energies obey

$$\lim \frac{\min E_{b,k}}{2\sqrt{bk}} = \min_{\substack{u_{\text{eff}}(x) \\ e(u_{\text{eff}}) \leq \frac{1}{2}\nabla p \otimes \nabla p}} \int_\Omega \frac{1}{2}|\nabla p|^2 - \int_{\partial\Omega} u_{\text{eff}} \cdot \hat{\nu}\, ds. \tag{2}$$

In-plane displacements $u_{b,k}$ selected by minimizing $E_{b,k}$ converge (weakly) to solutions u_{eff} of the limiting problem in (2). The out-of-plane part $w_{b,k} \to 0$. These results follow by identifying the Γ-limit of $\frac{1}{2\sqrt{bk}} E_{b,k}$ [2].

There is a suggestive geometric interpretation of the limiting problem (2) for u_{eff}. It says that the shell should deform while on the bath so as to *maximize its projected area* in the plane, under the constraint that it is *tension-free*. We turn to discuss its consequences now.

A typical assumption in the literature is that the out-of-plane displacement satisfies

$$w(x) = l_{\text{wr}} a(x) \cos\left(\frac{x \cdot \hat{\eta}(x)}{l_{\text{wr}}}\right) \tag{3}$$

where $l_{\text{wr}} \approx 0$ and the rescaled amplitude $a(x)$ and wrinkle direction $\hat{\eta}(x)$ are slowly varying. But there is no fundamental reason to believe that (3) is always correct, especially given the possibility of a disordered response. A broader viewpoint is to keep track of the *defect measure*

$$\mu(dx) := \text{weak-}* \lim \nabla w_{b,k} \otimes \nabla w_{b,k}\, dx \tag{4}$$

defined in [2]. For instance, the ansatz (3) generates the defect measure

$$\mu = \frac{a^2}{2}\hat{\eta}\otimes\hat{\eta}\,dx \tag{5}$$

upon taking $l_{\mathrm{wr}} \to 0$. Conversely, proving that (5) holds for minimizing sequences is one way of justifying (3).

The paper [2] proves that the defect measure μ associated to any sequence of minimizers of (1) must solve the boundary value problem

$$\begin{cases} -\frac{1}{2}\mathrm{curlcurl}\,\mu = \kappa & \text{in } \Omega \\ \langle \mathrm{cof}\,\nabla\nabla\varphi, \mu\rangle = 0 & \text{in } \Omega \\ \hat{\nu}\cdot[\nabla\varphi]\,\langle \hat{\tau}\otimes\hat{\tau}, \mu\rangle = 0 & \text{at } \partial\Omega \end{cases}. \tag{6}$$

The converse holds for "almost minimizers" which, by definition, have minimal energy at leading order ($E_{b,k} = \min E_{b,k} + o(\sqrt{bk})$). The term $\kappa = \det\nabla\nabla p$ is the Gauss curvature in the shallow shell approximation, where p is the initial height of the shell. The vectors $\hat{\nu}$ and $\hat{\tau}$ are the outwards pointing unit normal and tangent vectors to the shell boundary, $\partial\Omega$. The coefficients in (6) depend on a new quantity $\varphi(x)$, which solves the maximization problem

$$\max_{\substack{\varphi(x)\text{ is convex} \\ \varphi(x)=\frac{1}{2}|x|^2\text{ outside }\Omega}} \int_\Omega \varphi\kappa. \tag{7}$$

This is the convex dual of the limiting problem (2) for u_{eff}.

These results give a method for predicting wrinkle patterns: find a function φ solving (7), and solve the boundary value problem (6) for the defect measure μ. This gives the layout of the patterns and their amplitude. Figure 2 shows various patterns recovered this way, overlaid against the experiments. Blue lines fill out the predicted ordered regions where $\nabla\nabla\varphi \neq 0$ so that a simple formula like (5) holds. The wrinkle direction $\hat{\eta}(x)$ points perpendicularly to the plotted lines. The amplitude a obeys an ODE along each line, obtained directly from the PDE (6).

There is one final ingredient needed to prove the simple rules for wrinkles, and it is what gives the diagrams from Panel (B) of Figure 1. Again, these diagrams link the patterns of positively ($\kappa > 0$) and negatively ($\kappa < 0$) curved shells. The surprise is that the corresponding solutions φ_+ and φ_- of the dual problem (7) are Legendre transforms of one another, for simply connected Ω. With this fact and some convex analysis (see the Supplementary Information of [1]), one finally deduces that

$$\mathrm{cof}\,\nabla\nabla\varphi_- = R^T\nabla P_{\partial\Omega}R \quad\text{and}\quad \mathrm{cof}\,\nabla\nabla\varphi_+ = R^T\nabla P_{\mathfrak{M}}R. \tag{8}$$

In these formulas R is a rotation by $\pi/2$, and $x \mapsto P_{\partial\Omega}(x)$ is the closest point projection taking $x \in \Omega\backslash\mathfrak{M}$ to $P_{\partial\Omega}(x) \in \partial\Omega$. The set $\mathfrak{M}$ where this map is not well-defined is the *medial axis* of Ω. The reciprocal projection $x \mapsto P_{\mathfrak{M}}(x)$ is depicted in Panel (B) of Figure 1: there, $P_{\mathfrak{M}}(x) = p$ if x belongs to the segment qr or to the polygon $qrst$ (see [1] for a formula). Plugging (8) into (6) proves the simple rules.

References

[1] I. Tobasco, Y. Timounay, D. Todorova, G.C. Leggat, J.D. Paulsen, and E. Katifori, Exact solutions for the wrinkle patterns of confined elastic shells, *Nat. Phys.*, **18(9)** (2022), 1099–1104.

[2] I. Tobasco, Curvature-driven wrinkling of thin elastic shells, *Arch. Ration. Mech. Anal.*, **239(3)** (2021), 1211–1325.

[3] P. Bella and R.V. Kohn, Coarsening of folds in hanging drapes, *Comm. Pure Appl. Math.*, **70(5)** (2017), 978–1021.

[4] B. Davidovitch, Y. Sun, and G.M. Grason, Geometrically incompatible confinement of solids, *Proc. Natl. Acad. Sci. USA*, **116(5)** (2019), 1483–1488.

[5] B. Audoly and Y. Pomeau, *Elasticity and Geometry: From Hair Curls to the Non-Linear Response of Shells*, Oxford University Press, Oxford, 2010.

Global Minimization of Polynomial Integral Functionals

Federico Fuentes
(joint work with Giovanni Fantuzzi)

The first part of this talk is based on recently published work [1]. We describe a discretize-then-relax strategy to globally minimize integral functionals over functions in a Sobolev space satisfying prescribed Dirichlet boundary conditions, i.e.

$$\mathcal{F}^* := \inf_{u \in W_0^{1,p}(\Omega;\mathbb{R}^m)} \int_\Omega f(x, u(x), \nabla u(x)) \, \mathrm{d}x \,.$$

The strategy applies whenever the integral functional, f, depends polynomially on the function and its derivatives, even if it is nonconvex. Naturally, suitable (and well established) growth, coercivity and quasiconvexity assumptions are imposed on f to ensure the existence of such a minimizer in the first place, without precluding the existence of multiple local (or even global) minimizers (see [2]).

The 'discretize' step uses a bounded finite element scheme to approximate the integral minimization problem with a convergent hierarchy of sparse polynomial optimization problems over a compact feasible set, indexed by the decreasing size h of the finite element mesh. The 'relax' step employs *sparse* moment-sum-of-squares relaxations to approximate each polynomial optimization problem (POP) with a hierarchy of convex semidefinite programs (SDPs), indexed by an increasing relaxation order ω (see [3]). At each value of h and ω an approximation for the infimum, $\lambda^*_{h,\omega}$, is computed explicitly, and a function $u^*_{h,\omega} \in W_0^{1,p}(\Omega)$, is constructed from the moments collected from the SDP solution.

We remark that solving these problems typically requires at least hundreds of "degrees of freedom" to accurately capture solutions that often contain localized features, so the resulting POPs involve hundreds of variables. Without sparsity, this is computationally infeasible due to the prohibitive cost and memory footprint of solving each SDP. Meanwhile, exploiting sparsity whenever possible dramatically lowers these costs, making this a necessity in practice. The finite element

discretization not only naturally produces such correlative sparsity, but also provides flexibility in terms of the domain shape and benefits from existing theoretical results in approximation theory.

We prove that, as $\omega \to \infty$ and $h \to 0$,

$$\lim_{h\to 0} \lim_{\omega\to\infty} \lambda^*_{h,\omega} = \mathcal{F}^* \,.$$

Moreover, if the global minimizer of the original problem, u^*, is unique, we show that $u^*_{h,\omega}$ converges weakly to u^* in the Sobolev space $W^{1,p}(\Omega)$. In particular,

$$\lim_{h\to 0} \lim_{\omega\to\infty} \|u^*_{h,\omega} - u^*\|_{L^p(\Omega)} = 0 \,.$$

For this theoretical result, technical conditions on the finite element spaces are required, which in practice are always possible to satisfy, but, more importantly, regarding the sparse moment-sum-of-squares hierarchy, the so-called running intersection property (RIP) is also needed. Lastly, with a further separability assumption on the form of f, we show that evaluating the functional on $u^*_{h,\omega}$ produces a convergent sequence of upper bounds to $\mathcal{F}^*$ (since, in principle, the $\lambda^*_{h,\omega}$ are not known a priori to be either upper nor lower bounds on $\mathcal{F}^*$).

We report computational experiments, including ones that show our numerical strategy works well even when technical conditions required by our theoretical analysis are not satisfied. In particular, in our examples, we obtained reasonable results without enforcing the RIP, which is convenient, since the necessary modifications to ensure the RIP often result in prohibitive increases of computational cost. Our results also suggest that knowing L^∞ bounds of the global optimizers a priori could be of great benefit to improve the quality of results, and that, in practice, using ω-converged solutions to jump-start Newton solvers is a viable strategy to more quickly obtain the desired solution.

Finally, as a preview of upcoming work, we demonstrate how to extend these techniques to the minimization of integral functionals constrained by well-posed semilinear partial differential equations, which appear ubiquitously in control theory. We show computational experiments in one and two spatial dimensions of relevant problems in control theory.

References

[1] G. Fantuzzi and F. Fuentes, Global minimization of polynomial integral functionals, *SIAM J. on Sci. Comp.*, **46(4)** (2024), A2123–A2149.

[2] B. Dacorogna, *Direct Methods in the Calculus of Variations*, vol. 78 of *Applied Mathematical Sciences*, Springer, New York, 2nd ed., 2008.

[3] J.B. Lasserre, Convergent SDP relaxations in polynomial optimization with sparsity, *SIAM J. on Opt.*, **46(4)** (2006), 822–843.

Shortest Paths in Graphs of Convex Sets

Pablo Parrilo

Given a graph, the shortest-path problem requires finding a sequence of edges with minimum cumulative length that connects a source vertex to a target vertex. We consider a variant of this classical problem in which the position of each vertex in the graph is a continuous decision variable constrained in a convex set, and the length of an edge is a convex function of the position of its endpoints. Problems of this form arise naturally in many areas, from motion planning of autonomous vehicles to optimal control of hybrid systems. The price for such a wide applicability is the complexity of this problem, which is easily seen to be NP-hard. Our main contribution is a strong and lightweight mixed-integer convex formulation based on perspective operators, that makes it possible to efficiently find globally optimal paths in large graphs and in high-dimensional spaces.

Data-Driven System Analysis Using Polynomial Optimization and the Koopman Operator

Jason J. Bramburger
(joint work with Giovanni Fantuzzi)

Many important statements about dynamical systems can be proved by finding scalar-valued auxiliary functions whose time evolution along trajectories obeys certain pointwise inequalities that imply the desired result. A familiar example from ordinary differential equations (ODEs) is a Lyapunov function: a positive definite function V that proves the global stability of a vector field F if the time derivative of V along trajectories is decreasing, i.e. $\mathrm{d}V/\mathrm{d}t = F \cdot \nabla V \leq 0$ everywhere in space. On paper, finding auxiliary functions can be a difficult task. However, for polynomial dynamical systems one can attempt to find auxiliary functions by expanding them in a monomial basis, thus resulting in a polynomial optimization problem. There are now well-developed computational methods to solving these polynomial optimization problems associated coming from finite-dimensional dynamical systems which involve strengthening the inequalities to sum-of-squares conditions and then translating them to an equivalent semidefinite program to be solved numerically.

What all auxiliary functions have in common is that they encode system dynamics through the time derivative along trajectories, the *Lie derivative*, which for ODEs took the form $F \cdot \nabla V$ above. The Lie derivative can further be recognized as the generator of the celebrated Koopman operator [3], for which there are now machine learning techniques that can estimate the Koopman operator (and its Lie derivative) directly from data. This work employs extended dynamic mode decomposition (EDMD) to approximate the action of the Koopman operator on the span of finitely many monomials [5], for which there are performance guarantees as the amount of data increases [2, 4, 5]. This means that one can identify auxiliary functions by estimating the action of the Lie derivative using EDMD on

spans of monomials to construct auxiliary functions in using only data gathered from the system. Thus, by estimating the Lie derivative using EDMD, we can then implement the well-developed polynomial optimization framework for identifying auxiliary function to provide statements about an underlying system using only data gathered from it.

This presentation will rigorously present the connection between auxiliary function methods and the Koopman operator, while also introducing a variant of EDMD for estimating the Lie derivative from data that is best suited to this application. Two major strengths of the method will be highlighted throughout the talk. First, it can be applied to a broad class of deterministic and stochastic processes with no adjustments needed to differentiate between these two settings. Second, it directly discovers approximations of the Lie derivative, meaning that no system identification from the data is necessary. With these strengths, we highlight applications into:

(1) Finding Lyapunov functions from data to prove global stability of a steady-state;
(2) Bounding infinite time averages of quantities of interest for both deterministic and stochastic systems;
(3) Introduce the dual formulation of these methods from [1] that can identify invariant measures of dynamical systems.

This presentation will show how one can bring together some of the most important topics in modern dynamical systems analysis to produce a simple, yet powerful, numerical procedure to analyze dynamic data.

References

[1] J.J. Bramburger and G. Fantuzzi, Data-driven discovery of invariant measures, *arXiv preprint*, arXiv:2308.15318, 2023.
[2] J.J. Bramburger and G. Fantuzzi, Auxiliary functions as Koopman observables: Data-driven analysis of dynamical systems via polynomial optimization, *J. Nonlinear Sci.*, **34** (2024), 8.
[3] B.O. Koopman, Hamiltonian systems and transformation in Hilbert space, *Proc. Nat. Acad. Sci.*, **17** (1931), 315–318.
[4] M. Korda and I. Mezić, On convergence of extended dynamic mode decomposition to the Koopman operator, *J. Nonlinear Sci.*, **28** (2018), 687–710.
[5] M.O. Williams, I.G. Kevrekidis, and C.W. Rowley, A data-driven approximation of the Koopman operator: extending dynamic mode decomposition, *J. Nonlinear Sci.*, **25** (2015), 1307–1346.

Constructive Methods for Robust Control of PDE Systems

Emilia Fridman

Many important plants (e.g. flexible manipulators or heat transfer processes) are governed by partial differential equations (PDEs) and are often described by models with a significant degree of uncertainty. Some PDEs may not be robust with respect to arbitrary small time-delays in the feedback. Robust finite-dimensional controller design for PDEs is a challenging problem.

In this talk two constructive methods for finite-dimensional control will be presented:

Spatial decomposition (or sampling in space) method, where the spatial domain is divided into N subdomains with N sensors and actuators located in each subdomain; Modal decomposition method, where the controller is designed on the basis of a finite-dimensional system that captures the dominant dynamics of the infinite-dimensional one. Sufficient conditions ensuring the stability and performance of the closed-loop system are established in terms of simple linear matrix inequalities that are always feasible for appropriate choice of controllers. We will discuss delayed and sampled-data implementations as well as application to network-based deployment of multi-agents.

On the Conditions of SDP Problems arising in Polynomial Optimization

Michael Stingl

We will first introduce a definition of a condition number for SDP problems and show (by experiments) its relation to trustworthiness of results of standard SDP solvers. The numerical experiments will be focused on SDP problems arising in polynomial optimization. We will further present a new generation of the code Loraine.jl, focused on high-accuracy solution of small but difficult SDP problems. The code uses extended-precision arithmetic provided by MultiFloats.jl. We will demonstrate that the solver can deliver reliable solutions to difficult (highly ill-conditioned or almost infeasible) problems arising in global polynomial optimization. In particular, we will give numerical confirmation of theoretical convergence rate of Lasserre hierarchy for certain polynomial optimization problems studied in the literature.

Strengthening Lasserre's Hierarchy in Real and Complex Polynomial Optimization

Jie Wang

Consider the real polynomial optimization problem:

$$\text{(RPOP)} \qquad \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathbf{K}\},$$

where $f \in \mathbb{R}[\mathbf{x}]$ and the feasible set $\mathbf{K}$ is given by

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

and the complex polynomial optimization problem:

$$\text{(CPOP)} \qquad \inf\{f(\mathbf{x}, \overline{\mathbf{x}}) : \mathbf{x} \in \mathbf{K}\},$$

where

$$\mathbf{K} \coloneqq \{\mathbf{x} \in \mathbb{C}^n \mid g_i(\mathbf{x}, \overline{\mathbf{x}}) \geq 0, i = 1, \ldots, m\},$$

and $f, g_1, \ldots, g_m \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ are self-conjugate polynomials.

Lasserre's hierarchy [1] is a well-established scheme for globally solving (RPOP) and attracts a lot of attentions of researchers from diverse fields due to its nice theoretical properties in recent years. There is also a complex variant of Lasserre's hierarchy [2] for globally solving (CPOP). However, a bottleneck of Lasserre's hierarchy is its limited scalability as the size of associated semidefinite relaxations grows rapidly with relaxation orders. To accelerate the convergence of Lasserre's hierarchy, we propose to strengthen moment relaxations by including additional positive semidefinite conditions. Concretely, let $\mathbf{M}_d^{\mathbb{R}}$ and $\mathbf{M}_d^{\mathbb{C}}$ denote the d-th order real and complex moment matrices, respectively. We can prove the following result.

(i) Suppose that $\mathbf{y}$ is a pseudo-moment sequence with $\mathbf{M}_r^{\mathbb{R}}(\mathbf{y}) \succeq 0$ for some $r \in \mathbb{N}$. Then for any $s \in \mathbb{N}$ with $s < r$, one has

$$\begin{bmatrix} \mathbf{M}_s^{\mathbb{R}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{R}}(x_i\mathbf{y}) \\ \mathbf{M}_s^{\mathbb{R}}(x_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{R}}(x_i^2\mathbf{y}) \end{bmatrix} \succeq 0, \quad i = 1, \ldots, n. \tag{1}$$

(ii) Suppose that $\mathbf{y}$ is a complex moment sequence admitting a Borel representing measure. Then for any $s \in \mathbb{N}$, one has

$$\begin{bmatrix} \mathbf{M}_s^{\mathbb{C}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(x_i\mathbf{y}) \\ \mathbf{M}_s^{\mathbb{C}}(\overline{x}_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(|x_i|^2\mathbf{y}) \end{bmatrix} \succeq 0, \quad i = 1, \ldots, n. \tag{2}$$

Note that the PSD conditions (2) arise from the characterization of normality of the so-called *shift operators*.

Now we can strengthen moment relaxations by including the PSD conditions (1) or (2). For the real case, it is

$$\begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & y_{\mathbf{0}} = 1, \\ & \mathbf{M}_{r-d_i}^{\mathbb{R}}(g_i\mathbf{y}) \succeq 0, \quad i = 1, \ldots, m, \\ & \begin{bmatrix} \mathbf{M}_r^{\mathbb{R}}(\mathbf{y}) & \mathbf{M}_r^{\mathbb{R}}(x_i\mathbf{y}) \\ \mathbf{M}_r^{\mathbb{R}}(x_i\mathbf{y}) & \mathbf{M}_r^{\mathbb{R}}(x_i^2\mathbf{y}) \end{bmatrix} \succeq 0, \quad i = 1, \ldots, n. \end{cases} \tag{3}$$

It turns out that (3) provides an intermediate relaxation between the r-th and $(r+1)$-th moment relaxations for (RPOP).

For the complex case, it is

$$\begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r^{\mathbb{C}}(\mathbf{y}) \succeq 0, \quad y_{\mathbf{0},\mathbf{0}} = 1, \\ & \mathbf{M}_{r-d_i}^{\mathbb{C}}(g_i\mathbf{y}) \succeq 0, \quad i = 1, \ldots, m, \\ & \begin{bmatrix} \mathbf{M}_s^{\mathbb{C}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(x_i\mathbf{y}) \\ \mathbf{M}_s^{\mathbb{C}}(\overline{x}_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(|x_i|^2\mathbf{y}) \end{bmatrix} \succeq 0, \quad i = 1, \ldots, n. \end{cases} \tag{4}$$

Here $s \in \mathbb{N}$ is a tunable parameter which we call the *normal order*. It turns out that (4) provides a bilevel hierarchy of relaxations for (CPOP) which is indexed by the relaxation order r and the normal order s.

To improve scalability, the strengthening technique can be further integrated into different sparse versions of Lasserre's hierarchy. Numerical experiments are

performed on a variety of real and complex polynomial optimization problems. It is shown that the strengthening technique can indeed improve the bound produced by the usual Lasserre's hierarchy, and most often allows to achieve global optimality at lower relaxation orders for complex polynomial optimization problems. Consequently, the strengthening technique provides substantial computational savings and yields considerable speedup (sometimes even by orders of magnitudes).

References

[1] J.B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Optim.*, **11** (2001), no. 3, 796–817.

[2] C. Josz and D.K. Molzahn, Lasserre hierarchy for large scale polynomial optimization in real and complex variables, *SIAM J. Optim.*, **28** (2018), no. 2, 1017–1048.

Sparse Moment-SOS Relaxations for Direct Optimal Control

Heng Yang

(joint work with S. Kang, X. Xu, J. Sarva, L. Liang)

Direct methods for optimal control, referred to as *trajectory optimization*, is the workhorse for computing numerical solutions of complex optimal control problems. Nonlinear programming with engineered initializations has been the de-facto approach for trajectory optimization, which however, can suffer from undesired local optimality.

In this talk, I show that many trajectory optimization problems that can be written as polynomial optimization problems (POPs) admit *tight* convex relaxations, and such relaxations can be solved at large scale by exploiting problem structures jointly with GPU-based parallelization in a customized ADMM solver. A full paper of this talk can be found in [1] with numerical experiments.

1. Problem Statement

Let N be the number of steps (with $[N] := \{1, \dots, N\}$), $\{x_k\}_{k=0}^N \subset \mathbb{R}^{d_x}$ be the state trajectory, and $\{u_k\}_{k=0}^{N-1} \subset \mathbb{R}^{d_u}$ be the control trajectory, we consider the following trajectory optimization problem:

$$\min_{\{u_k\}_{k=0}^{N-1}, \{x_k\}_{k=0}^N} \quad l_N(x_N) + \sum_{k=0}^{N-1} l_k(x_k, u_k) \tag{1a}$$

$$\text{subject to} \quad x_0 = x_{\text{init}} \tag{1b}$$

$$F_k(x_{k-1}, u_{k-1}, x_k) = 0, \ \forall k \in [N] \tag{1c}$$

$$(u_{k-1}, x_k) \in \mathcal{C}_k, \ \forall k \in [N] \tag{1d}$$

where $l_k, k = 0, \dots, N$ are the instantaneous and terminal losses; x_{init} is the initial state; F_k represents the discretized system dynamics as a differential algebraic equation (*e.g.* obtained from the continuous-time dynamics via multiple shooting); and $\mathcal{C}_k$ imposes constraints on u_{k-1} and x_k (*e.g.* control limits, obstacle

avoidance). We assume l_k and F_k are polynomial functions and $\mathcal{C}_k$ are basic semi-algebraic sets (*i.e.* described by polynomial constraints), in which case problem (1) is an instance of *polynomial optimization* (POP) that is nonconvex and NP-hard in general.

2. Sparse Moment-SOS Relaxations

We first introduce a special type of sparsity in polynomial optimization known as *chain-like* sparsity and show that problem (1) satisfies this pattern.

Definition 1 (POP with Chain-like Sparsity)**.** *Let $z \in \mathbb{R}^d$ and $I_1, \ldots, I_N \subset [d]$ be N index sets such that $\bigcup_{k=1}^N I_k = [d]$ and $\left(\bigcup_{j=1}^{k-1} I_j\right) \bigcap I_k = I_{k-1} \bigcap I_k,\ \forall k \in \{2, \ldots, N\}$, then the following POP is said to admit a chain-like sparsity pattern*

$$
\begin{aligned}
&p^\star = \min_{z \in \mathbb{R}^d} && \sum_{k=1}^N f_k(z(I_k)) && \text{(2a)}\\
&\text{subject to} && g_{k,i}(z(I_k)) \geq 0,\ \forall k \in [N], i \in \mathcal{G}_k && \text{(2b)}\\
& && h_{k,j}(z(I_k)) = 0,\ \forall k \in [N], j \in \mathcal{H}_k && \text{(2c)}
\end{aligned}
$$

where $\mathcal{G}_k$ (resp. $\mathcal{H}_k$) indexes the inequality (resp. equality) constraints for variables $z(I_k)$, and $f_k, g_{k,i}, h_{h,j} \in \mathbb{R}[z(I_k)]$ are polynomials in the variables $z(I_k)$ (called a clique).

For trajectory optimization problem (1), each clique contains variables $z(I_k) = \{x_{k-1}, u_{k-1}, x_k\} \in \mathbb{R}^{2d_x+d_u}$ and these cliques form a chain-like graph. Clearly, each dynamics constraint (1c) and set constraint (1d) only involves the k-th clique, and the objective function (1a) readily admits the decomposition in the form of (2a).

We now introduce the sparse moment-SOS hierarchy for relaxing the POP (2) as a sequence of semidefinite programs. Given a sequence of numbers $\varphi = (\varphi_\alpha) \in \mathbb{R}^{s(d,n)}$ indexed by the (exponents of) monomials in $[z]_n$, we define a one-to-one linear map from $\mathbb{R}[z]_n$ to $\mathbb{R}$,

$$
\ell_\varphi :\ f(z) = \sum_\alpha c_\alpha z^\alpha \ \mapsto\ \sum_\alpha c_\alpha \varphi_\alpha,\ \forall f(z) \in \mathbb{R}s[z]_n. \tag{3}
$$

The application of ℓ_φ to polynomial vectors and matrices is element-wise.

Proposition 2 (Sparse Moment Relaxation)**.** *In (2), denote $d_k^f := \deg(f_k)$, $d_{k,i}^g := \lceil \deg(g_{k,i})/2 \rceil$, and $d_{k,j}^h := \deg(h_{k,j})$. Given a positive integer κ such that*

$$
2\kappa \geq 2\kappa_0 = \max\{\{d_k^f\}_{k\in[N]},\ \{2d_{k,i}^g\}_{k\in[N],i\in\mathcal{G}_k},\ \{d_{k,j}^g\}_{k\in[N],j\in\mathcal{H}_k}\},
$$

the κ-th order sparse moment relaxation for problem (2) *reads*

$$p_\kappa^\star = \min_{\varphi \in \mathbb{R}s(d,2\kappa)} \sum_{k=1}^{N} \ell_\varphi \left(f_k(z(I_k)) \right) \tag{4a}$$

$$\text{subject to} \quad \varphi_0 = 1, \ \text{and } \forall k \in [N], i \in \mathcal{G}_k, j \in \mathcal{H}_k : \tag{4b}$$

$$M_k := \ell_\varphi \left([z(I_k)]_\kappa [z(I_k)]_\kappa^\mathsf{T} \right) \succeq 0, \tag{4c}$$

$$L_{k,i} := \ell_\varphi \left(g_{k,i}(z(I_k)) \cdot [z(I_k)]_{\kappa - d_{k,i}^g} [z(I_k)]_{\kappa - d_{k,i}^g}^\mathsf{T} \right) \succeq 0, \tag{4d}$$

$$\ell_\varphi \left(h_{k,j}(z(I_k)) \cdot [z(I_k)]_{2\kappa - d_{k,j}^h} \right) = 0. \tag{4e}$$

Problem (4) is a convex optimization whose variable is the sequence φ. The sequence φ is called the *truncated moment sequence* (TMS) because it can be interpreted as the moments of a probability measure supported on the feasible set of the POP (2). M_k and $L_{k,i}$, whose entries are linear in φ, are called the *moment matrix* and the *localizing matrix*, respectively.

The power of the sparse moment relaxation (4) is twofold. First, every relaxation generates a lower bound $p_\kappa^\star \leq p^\star$ and the lower bound increasingly converges to $p^\star$. Second, the convergence can be detected and certified.

Theorem 3 (Convergence of Sparse Moment Relaxations [2])**.** *Suppose $\forall k \in [N]$, $\exists R_k > 0$ such that $\|z(I_k)\|_\infty \leq R_k$. Then:*

(i) *Problem* (4) *is* solvable *for any $\kappa \geq \kappa_0$. Let $M_k^\star$ be an optimal solution.*
(ii) *$p_\kappa^\star \leq p^\star, \forall \kappa \geq \kappa_0$. Moreover, $p_\kappa^\star \to p^\star$, as $\kappa \to \infty$.*
(iii) *If under some relaxation order κ, it holds that rank$(M_k^\star) = 1, \forall k \in [N]$. Then $p_\kappa^\star = p^\star$, i.e. the relaxation is tight.*

Theorem 3 provides a way to check the tightness of the relaxation and extract globally optimal solutions $z^\star$ of the nonconvex POP (2). In fact, condition 3 guarantees the optimal TMS $\phi^\star$ admits a representing Dirac measure supported on the unique optimal solution $z^\star$. Therefore, one can extract $z^\star$ from the entries of $\varphi^\star$ corresponding to degree-one monomials. Even when the relaxation is not tight, one can extract a near-optimal solution $\hat{z}$ using a three-step heuristic described in [1]. Denote as $\hat{p}$ the POP objective (2a) evaluated at $\hat{z}$, we compute the *certificate of suboptimality*

$$\xi_\kappa = \frac{\hat{p} - p_\kappa^\star}{1 + |\hat{p}| + |p_\kappa^\star|} \geq 0. \tag{5}$$

An ξ_κ that is close to zero certifies the global optimality of $\hat{z}$.

3. Fast sGS-ADMM in GPUs

Focusing on the sparse moment relaxation (4), denote

$$X := (\{M_k\}_{k \in [N]}, \{L_{k,i}\}_{k \in [N], i \in \mathcal{G}_k}) \tag{6}$$

as the tuple of all moment and localizing matrices, which lives in a vector space Ω that is the Cartesian product of $\{\mathbb{S}^{s(|I_k|,\kappa)}\}_{k \in [N]}$ and $\{\mathbb{S}^{s(|I_k|,\kappa - d_{k,i}^g)}\}_{k \in [N], i \in \mathcal{G}_k}$.

Let $\Omega_+ \subset \Omega$ be the cone containing tuples X whose elements are PSD, we claim the moment relaxation (4) can be written as a standard multi-block SDP [3]

$$p_\kappa^\star = \min_X \{\langle C, X\rangle : \mathcal{A}(X) = b,\ X \in \Omega_+\} \tag{7}$$

for some $b \in \mathbb{R}^m$, $C \in \Omega$ and linear map $\mathcal{A}(X) := (\langle A_i, X\rangle)_{i\in[m]}$ with $A_i \in \Omega$. The inner product in the space Ω is defined element-wise. Due to space limitations, we provide a tutorial about how to convert the moment relaxation (4) as a standard SDP (7) in [1], where we also show our conversion is two orders of magnitude faster than SOSTOOLS and YALMIP .

Let $y \in \mathbb{R}^m$ and denote $\mathcal{A}^*$ as the adjoint of $\mathcal{A}$ defined as $\mathcal{A}^* y := \sum_{l\in[m]} y_l A_l$. The Lagrangian dual of (7) reads:

$$\max_{y\in\mathbb{R}^m, S\in\Omega} \{\langle b, y\rangle : \mathcal{A}^* y + S = C,\ S \in \Omega\}. \tag{8}$$

sGS-ADMM can be seen as a semi-proximal ADMM method applied to the Augmented Lagrangian of (8). A full description of the algorithm, its implementation details, and numerical experiments can be found in [1].

References

[1] S. Kang, X. Xu, J. Sarva, L. Liang, and H. Yang, Fast and certifiable trajectory optimization, *Int. Work. on the Alg. Found. of Rob. (WAFR)*, vol. 1, 2024, pp. 3.

[2] J.B. Lasserre, Convergent SDP-relaxations in polynomial optimization with sparsity, *SIAM J. Optim.*, **17** (2006), no. 3, 822–843.

[3] R.H. Tütüncü, K.C. Toh, and M.J. Todd, Solving semidefinite-quadratic-linear programs using sdpt3, *Math. Prog.*, **95** (2003), 189–217.

[4] X. Li, D. Sun, and K.C. Toh, A block symmetric Gauss-Seidel decomposition theorem for convex composite quadratic programming and its applications, *Math. Prog.*, **175** (2019), 395–418.

[5] L. Chen, D. Sun, and K.C. Toh, An efficient inexact symmetric Gauss-Seidel based majorized ADMM for high-dimensional convex composite conic programming, *Math. Prog.*, **161** (2017), 237–270.

A Mathematical Formalism for Safe Learning

Amir Ali Ahmadi

(joint work with Abraar Chaudhry, Vikas Sindhwani, Stephen Tu)

In many applications such as robotics, autonomous systems, and safety-critical control, one needs to learn a model of a dynamical system by observing a small set of its trajectories in a safe manner. This model can serve as a tool for making predictions about unobserved trajectories of the system. It can also be used for accomplishing downstream control objectives. Often, an important challenge during the initial stages of learning is that deploying even a conservative learning strategy on a real world system, such as a robot, is fraught with risk. How should the robot be "set loose" (i.e., initialized) in the real world so that our uncertainty about its dynamics is reduced, but with guarantees that the robot will remain safe (e.g., it does not exit a pre-specified region in state space)? How much more

aggressive can our learning strategy get "on the fly" as uncertainty is reduced? This interplay between *safety and uncertainty while learning dynamical systems* is the central theme of our work. We propose a mathematical formulation that captures the essence of this interplay and study the optimization problems that arise from the formulation in several settings.

The central object of our mathematical framework is a discrete-time dynamical system

$$x_{t+1} = f_\star(x_t), \tag{1}$$

where $f_\star : \mathbb{R}^n \to \mathbb{R}^n$ is an *unknown* map. This could be either a naturally arising autonomous system, or a closed-loop control system with a fixed feedback policy. Our interest is in the problem of safe data acquisition for estimating the unknown map $f_\star$ from a collection of length-T trajectories $\{\varphi_{f_\star,T}(x_j)\}_{j=1}^m$, where $\varphi_{f,T}(x) := (x, f(x), \ldots, f^{(T)}(x))$.

In our setting, we are given as input a set $S \subset \mathbb{R}^n$, called the *safety region*, in which the state should remain throughout the learning process. We say that an initial state x is *T-step safe* under a map $f : \mathbb{R}^n \to \mathbb{R}^n$ if $f^{(i)}(x) \in S$ for all $i = 0, \ldots, T$. We define $S^T(f) \subseteq S$ to be the set of states that are T-step safe under f. In order to safely learn $f_\star$, we require that measurements are made only at points in $S^T(f_\star)$. Obviously, if we make no assumptions about $f_\star$, this task is impossible. We assume, therefore, that the map $f_\star$ belongs to a set of dynamics U_0, which we call the *initial uncertainty set*. As experience is gathered, the uncertainty over $f_\star$ decreases. Let us denote the uncertainty set after we have observed k trajectories $\{\phi_{f_\star,T}(x_j)\}_{j=1}^k$ by,

$$U_k := \{f \in U_0 \mid \phi_{f,T}(x_j) = \phi_{f_\star,T}(x_j)\,, j = 1, \ldots, k\}.$$

Observe that $U_{k+1} \subseteq U_k$ for all k. For a nonnegative integer k, define

$$S_k^T := \bigcap_{f \in U_k} S^T(f)\,,$$

the set of points that are T-step safe under all dynamics consistent with the initial uncertainty set and the data after observing k trajectories. We refer to the set S_k^T as the *T-step safe set* (the dependence on k is implicit). Note that $S_k^T \subseteq S_{k+1}^T$ for all k. A primary goal of our work is to characterize the sets S_k^T as feasible regions of tractable optimization problems. In certain settings where an exact tractable characterization is not possible, our goal would be to find tractable inner approximations of these sets. For robustness reasons, we would like these inner approximations to be full-dimensional so that safe queries to the system can be made while tolerating perturbations which may arise during implementation.

A secondary goal of our work is to provide algorithms for what we define as the *T-step safe learning problem*. Fix a scalar $\bar{\varepsilon} > 0$ and a norm $||.||$ on $\mathbb{R}^n$. Given a safety region $S \subset \mathbb{R}^n$ and an initial uncertainty set U_0, the T-step safe learning problem (up to accuracy $\bar{\varepsilon}$ and with respect to norm $||.||$) is to sequentially choose vectors $x_1, \ldots, x_m$, for some nonnegative integer m, such that:

(1) **(Safety)** for each $k = 1, \dots, m$, $x_k \in S_{k-1}^T$,
(2) **(Learning)** $\sup_{f \in U_m, x \in S^T(f_\star)} \|f(x) - f_\star(x)\| \leq \bar{\varepsilon}$.

If for a given T, no such sequence of vectors $x_1, \dots, x_m$ exists (for any m), we say that T-step safe learning is impossible. Note that if T-step safe learning is possible, then T'-step safe learning is also possible for any $T' < T$. Moreover, since the highest rate of safe information assimilation is achieved when $T = 1$, to prove that safe learning is impossible for any T, it is necessary and sufficient to prove its impossibility for $T = 1$.

In many situations, the choice of the sequence of $\{x_1, \dots, x_m\}$ that achieves T-step safe learning may not be unique. We further suppose that for a function $c : \mathbb{R}^n \mapsto \mathbb{R}$ that takes nonnegative values over S, initializing the unknown system at a state $x \in S$ comes at a cost of $c(x)$. In such a setting, we are interested in safely learning the dynamical system at minimum total initialization cost. Ideally, we wish to minimize $\sum_{k=1}^m c(x_k)$ over sequences $\{x_1, \dots, x_m\}$ that satisfy the safe learning conditions 1 and 2 above. However, such an optimization problem cannot be solved without knowing the action of the true dynamics $f_\star$ on the initialization points $\{x_k\}$ ahead of time. Hence, a natural online algorithm is to sequentially solve the following greedy optimization problem

$$\min_{x \in S_{k-1}^T} \; c(x)\,, \tag{2}$$

whose optimal solution gives the next cheapest T-step safe initialization point x_k, given information gathered before time k. A byproduct of our primary goal of characterizing the sets S_k^T tractably is efficient algorithms for solving the optimization problem (2).

In our work, we derive tractable conic programs that exactly characterize or inner approximate T-step safe sets (for any k) for both linear systems and a general class of nonlinear systems in the extreme cases when $T = 1$ and $T = \infty$. For linear systems, we also address the case when $T = 2$, and provide algorithms for solving the exact (i.e., $\bar{\varepsilon} = 0$) T-step safe learning problem when $T = 1, 2, \infty$. The interested reader can find these results and more details in [1].

References

[1] A.A. Ahmadi, A. Chaudhry, V. Sindhwani, and S. Tu, Safely learning dynamical systems, *arXiv preprint* `arXiv:2305.12284`, 2024.

Verifiable Safe Data-Driven Control

Mario Sznaier

The goal of this presentation was twofold (i) to use the problem of safe data driven control to illustrate the difficulties entailed in using semi-algebraic optimization techniques in scenarios involving dynamical systems, due to poor scaling properties, and (ii) to propose ways to mitigate this issue exploiting duality and a novel

semi-definite optimization algorithm that requires solving only second-order cone or linear programs.

1. Verifiable safe control for known systems

The goal of safe data driven control is to synthesize safe control laws that guarantee that all trajectories of a dynamical system of the form

$$\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})u + \mathbf{w} \tag{1}$$

that start in a set $\mathcal{X}_o$ do not enter an unsafe set $\mathcal{X}_u$ for all possible polytopic disturbances $\mathbf{w} \in \mathcal{W} \doteq \{\mathbf{w} \colon \mathbf{W}\mathbf{w} \leq \mathbf{d}_w\}$

Safety verification and synthesis of safe control laws for the case where the dynamics are known have been the subject of intense research during the past decade. Level-set methods separate the initial and unsafe set by the 0-contour of a solved function. Barrier functions [1] are a level-set method to certify the safety of trajectories, given that the superlevel sets of the barrier function are invariant. This superlevel invariance can be relaxed through slack (class-$\mathcal{K}$) conditions, while ensuring that the 0-level set is invariant [2, 6]. The level-set certificate of stability may be solved jointly with a safety-guaranteeing control policy $u(\cdot)$ (Control Barrier Function). When a barrier function is given, the min-norm controller will ensure safety of trajectories, and can be found through quadratic programming [3]. Barrier functions and funnels [4, 5, 7] contain bilinearities when jointly synthesizing controllers and barriers. An alternative level-set certificate is Density functions [9], which are based on Dual Lyapunov methods for stability [8]. Controllers and density functions can be simultaneously solved in a convex manner. Specifically, it can be shown that the following modification of Rantzer's criteria [10] yields a safe control law that is robust to ℓ_∞ bounded disturbances:

Lemma 1 ([11])**.** *Assume that the set $\mathcal{X}_u$ has a description:*

$$\mathcal{X}_u \doteq \{\mathbf{x} \colon h_i(\mathbf{x}) \geq 0,\ i = 1 \ldots N_c\}\,.$$

If there exist scalar functions $\rho(\mathbf{x}), \psi(\mathbf{x}) \in C^1$ such that: (i) $u(\mathbf{x}) \doteq \frac{\psi(\mathbf{x})}{\rho(\mathbf{x})}$ is well defined over the safe region $\rho(\mathbf{x}) \geq 0$, (ii) for all $\mathbf{w}(\cdot) \in \mathcal{W}$ and initial condition $\mathbf{x}_0 \in \mathcal{X}_0$, the trajectories of (1) *are well defined, and (iii) the following conditions hold:*

$$\nabla \cdot [\rho(\mathbf{x})\,(f(\mathbf{x}) + \mathbf{w}) + \psi(\mathbf{x})g(\mathbf{x})] - \rho(\mathbf{x})h(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in \mathbf{R}^n \text{ and } \mathbf{w} \in \mathbf{W} \tag{2a}$$

$$\rho(\mathbf{x}) \geq 0,\ \forall \mathbf{x} \in \mathcal{X}_0,\ \rho(\mathbf{x}) < 0,\ \forall \mathbf{x} \in \mathcal{X}_u, \tag{2b}$$

where $h \doteq \min_i \{h_i(\mathbf{x})\}$, then the control law $u(\mathbf{x})$ renders the closed loop system robustly safe with respect to $\mathcal{X}_u$.

When the dynamics and the constraints defining the sets $\mathcal{X}_o$ and $\mathcal{X}_u$ are polynomial, ρ and ψ can be found by solving a semi-algebraic optimization problem in the 2n indeterminates $\mathbf{x}$ and $\mathbf{w}$.

2. The data driven case

Consider now the case where the only information available about the plant is that it can be expressed as linear combinations of functions $\phi : \mathbf{R}^n \to \mathbf{R}^{d_f}$, $\gamma : \mathbf{R}^n \to \mathbf{R}^{d_g}$ with

$$f(\mathbf{x}) = \mathbf{F}\phi(\mathbf{x});\ g(\mathbf{x}) = \mathbf{G}\gamma(\mathbf{x}) \tag{3}$$

for some unknown system parameter matrices $\mathbf{F} \in \mathbf{R}^{n\times d_f}$ and $\mathbf{G} \in \mathbf{R}^{n\times d_g}$, that must be inferred from experimental measurements of T derivative-state-input tuples, $\{\dot{\mathbf{x}}_s, \mathbf{x}_s, u_s\}_{s=t_1\ldots t_T}$, sampled from the trajectories of (1) under some bounded disturbance $\mathbf{w} \in W$, indexed by the observations times $t_1 \ldots t_T$. In this context, the safe data driven control problem can be formally stated as: *Problem 1*, Given a disturbance set description $(\mathbf{W}, \mathbf{d}_w)$, training data $\{(\dot{\mathbf{x}}_s, \mathbf{x}_s, u_s)\}_{s=t_1\ldots t_T}$, and basic semialgebraic sets $\mathcal{X}_0$, $\mathcal{X}_u$, find a state-feedback control law $u(\mathbf{x})$ that renders all closed-loop systems consistent with the observed data and priors $\mathcal{W}$-robustly safe with respect to $\mathcal{X}_0$ and $\mathcal{X}_u$.

Definition 1. *Given training data $\{(\dot{\mathbf{x}}_s, \mathbf{x}_s, u_s)\}_{s=t_1\ldots t_T}$ and the uncertainty description $(\mathbf{W}, \mathbf{d_w})$, the consistency set $\mathcal{C}$, which contains all systems that are consistent with the data is defined as:*

$$\mathcal{C} \doteq \{f, g\colon \mathbf{W}\left[\dot{\mathbf{x}}_s - f(\mathbf{x}_s) - g(\mathbf{x}_s)u_s\right] \leq \mathbf{d_w}, s = t_1 \ldots t_T\} \tag{4}$$

Exploiting the property of the Kronecker product

$$\mathrm{vec}(\mathbf{PXQ}) = (\mathbf{Q}^T \otimes \mathbf{P})\mathrm{vec}(\mathbf{X}),$$

with $\mathbf{f} = \mathrm{vec}(\mathbf{F}^T)$, $\mathbf{g} = \mathrm{vec}(\mathbf{G}^T)$ leads to an equivalent representation of (4)

$$\mathcal{C} = \left\{\mathbf{f}, \mathbf{g}\colon \begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}\begin{bmatrix}\mathbf{f}\\ \mathbf{g}\end{bmatrix} \leq \xi - \mathbf{1}\otimes\mathbf{d}_w\right\} \tag{5}$$

using the matrix blocks (with $f(\mathbf{x}_s) = \mathrm{vec}(\phi(\mathbf{x}_s)^T\mathbf{F}^T)$)

$$\mathbf{A} \doteq \begin{bmatrix}\mathbf{W}\otimes\phi^T(t_1)\\ \vdots \\ \mathbf{W}\otimes\phi^T(t_T)\end{bmatrix}, \mathbf{B} \doteq \begin{bmatrix}\mathbf{W}\otimes u_{t_1}\gamma^T(t_1)\\ \vdots \\ \mathbf{W}\otimes u_{t_T}\gamma^T(t_T)\end{bmatrix}, \xi \doteq \begin{bmatrix}\mathbf{W}\dot{\mathbf{x}}(t_1)\\ \vdots \\ \mathbf{W}\dot{\mathbf{x}}(t_T)\end{bmatrix} \tag{6}$$

Combining this description with the polytopic description of the disturbances leads to a polytopic augmented consistency set describing the set of all possible plants and disturbances:

$$\mathcal{P}_1 \doteq \left\{\mathbf{f}, \mathbf{g}, \mathbf{w}\colon \begin{bmatrix}\mathbf{A} & \mathbf{B} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{W}\end{bmatrix}\begin{bmatrix}\mathbf{f}\\ \mathbf{g}\\ \mathbf{w}\end{bmatrix} \leq \begin{bmatrix}\xi - \mathbf{1}\otimes\mathbf{d}_w\\ \mathbf{d_w}\end{bmatrix}\right\} \tag{7}$$

Thus, in principle the data driven safe control problem can be reduced to a semi-algebraic optimization problem by simply imposing that condition (2) holds for all $\mathbf{x}$ and all $\mathbf{w}, \mathbf{f}, \mathbf{g} \in \mathcal{P}_1$. Note however that this problem will involve polynomials in $\mathcal{O}(n^2)$ indeterminates (since $\mathbf{f}, \mathbf{g}$ are now variables). Thus, it quickly becomes intractable. In the next section we will show how to exploit duality to recover a semi-algebraic optimization problem in only $\mathcal{O}(n)$ indeterminates.

3. Reducing complexity via duality

For a given pair (ρ, ψ), consider the set of all systems of the form (1) that are rendered safe by the control action $u = \frac{\psi}{\rho}$, along with the admissible perturbations, that is, the set of all $(\mathbf{f}, \mathbf{g}, \mathbf{w})$ such that

$$\nabla \cdot [\rho(\mathbf{x})f(\mathbf{x}) + \psi(\mathbf{x})g(\mathbf{x}) + \rho(\mathbf{x})\mathbf{w}] - \rho(\mathbf{x})h(\mathbf{x}) > 0 \tag{8}$$

holds for all $\mathbf{x} \in \mathbb{R}^n$. For each $\mathbf{x}$, this set is a polytope of the form:

$$\mathcal{P}_2 \doteq \left\{ \mathbf{f}, \mathbf{g}, \mathbf{w} : - \begin{bmatrix} (\nabla \cdot [\rho(\mathbf{I}_n \otimes \phi^T)])^T \\ (\nabla \cdot [\psi(\mathbf{I}_n \otimes \gamma^T)])^T \\ (\nabla \rho)^T \end{bmatrix}^T \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \\ \mathbf{w} \end{bmatrix} < -\rho h \right\} \tag{9}$$

where the divergence operator is applied column-wise to the matrix. The term $\nabla \cdot [\rho(\mathbf{x})(\mathbf{I}_n \otimes \phi(\mathbf{x})^T)]\mathbf{f}$ may be interpreted as $\nabla \cdot \mathrm{vec}(\rho(\mathbf{x})\phi(\mathbf{x})^T\mathbf{F}^T) = \nabla \cdot [\rho(\mathbf{x})f(\mathbf{x})]$.

It follows that (8) holds for all admissible disturbances $\mathbf{w} \in W$ and all plants in the consistency $\mathcal{C}$ set if and only if $\mathcal{P}_1 \subseteq \mathcal{P}_2$. This inclusion can be enforced through duality as follows:

Lemma 2 ([11])**.** *Assume that the data and priors are consistent (e.g. $\mathcal{C} \neq \emptyset$). Then $\mathcal{P}_1 \subseteq \mathcal{P}_2$ if there exists a vector function $\mathbf{y}(\mathbf{x}) \geq 0, \mathbf{y}(\mathbf{x}) \in \mathbb{R}^{2nT+2n}$ such that the following functional set of affine constraints is feasible:*

$$\mathbf{y}^T(\mathbf{x})\mathbf{N} = \mathbf{r}(\mathbf{x}) \text{ and } \mathbf{y}^T(\mathbf{x})\mathbf{e} < -\rho(\mathbf{x})h(\mathbf{x}), \tag{10}$$

where

$$\begin{aligned} &\mathbf{N} \doteq \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W} \end{bmatrix}, \ \mathbf{e} \doteq \begin{bmatrix} \xi - \mathbf{1} \otimes \mathbf{d}_w \\ \mathbf{d}_\mathbf{w} \end{bmatrix}, \\ &\mathbf{r}(\mathbf{x}) \doteq - \begin{bmatrix} \nabla \cdot [\rho(\mathbf{I}_n \otimes \phi^T)] & \nabla \cdot [\psi(\mathbf{I}_n \otimes \gamma^T)] & \nabla \rho \end{bmatrix} \end{aligned} \tag{11}$$

Remark 1. *Proceeding as in Theorem 2 in* [12]*, it can be shown that if $\phi(\mathbf{x}), \gamma(\mathbf{x})$ are continuous functions, then $\mathbf{y}(\mathbf{x})$ can be chosen to be continuous.*

Exploiting this Lemma leads to the following result

Theorem 2 ([11])**.** *A sufficient condition for the existence of a state-feedback control law $u(\mathbf{x})$ such that all systems in the consistency set $\mathcal{C}$ are rendered robustly safe, is that there exists a continuous vector function $\mathbf{y}(\mathbf{x}) \geq 0$ and functions $\rho \in C^1$, $\psi \in C^1$ such that*

$$\mathbf{y}^T(\mathbf{x})\mathbf{N} = \mathbf{r}(\mathbf{x}), \ \forall \mathbf{x} \in \mathbb{R}^n \tag{12a}$$

$$\mathbf{y}^T(\mathbf{x})\mathbf{e} < -\rho(\mathbf{x})h(\mathbf{x}), \ \forall \mathbf{x} \in \mathbb{R}^n \tag{12b}$$

$$|\psi(\mathbf{x})| \leq -\rho(\mathbf{x})h(\mathbf{x}), \ \forall \mathbf{x} \in \mathbb{R}^n \tag{12c}$$

$$\rho(\mathbf{x}) \geq 0, \ \forall \mathbf{x} \in \mathcal{X}_0 \tag{12d}$$

$$\rho(\mathbf{x}) < 0, \ \forall \mathbf{x} \in \mathcal{X}_u. \tag{12e}$$

The corresponding control law is given by $u(\mathbf{x}) = \frac{\psi(\mathbf{x})}{\rho(\mathbf{x})}$.

Remark 2. *Constraint* (12c) *is a convex tightening of the condition that* $\psi = 0$ *when* $\rho = 0$ *in the safe zone* $\rho(\mathbf{x}) \geq 0$. *This ensures satisfaction of Assumption (i) in Lemma 1.*

Remark 3. *Note that since the entries of* $y(\mathbf{x})$ *are continuous, they can be approximated by polynomials. Thus, in the case of polynomial dynamics, finding polynomial* (ρ, ψ) *reduces to a semi-algebraic optimization problem involving polynomials only in the indeterminate* $\mathbf{x} \in \mathbb{R}^n$. *Hence, this dual problem has substantially better scaling properties than the original (primal) problem as one goes up the Lasserre hierarchy.*

4. An efficient algorithm for solving Semi-Definite Programs

It is well known that semi-algebraic optimization problems can be relaxed to a convex Semidefinite Program (SDP) of the form:

$$\begin{aligned} X^*_{PSD} = \underset{X \succeq 0}{\text{argmin}} \quad & \text{Tr}\left(C^T X\right) \\ \text{s.t.} \quad & \text{Tr}\left(A_i^T X\right) = b_i \quad i = 1, \ldots, M \end{aligned} \tag{13}$$

However, even when exploiting duality, the safe data driven control problem leads to SDPs involving matrices of size $\binom{2n+r}{r}$, where r is the order of the relaxation. For instance, for a system with 10 states, using polynomials of degree 3 leads to matrices of size 1140, and using 4^{th} order polynomials leads to matrices of size 5985. In turn, solving the resulting SDP using a conventional interior point method will involve at least $\mathcal{O}(3 \times 10^8)$ variables, which is beyond the capabilities of most solvers. In this section we review an algorithm that we recently proposed [13], based upon a combination of Cholesky factorizations and Linear Programs (LPs) or Second Order Cone Programs (SOCPs). This algorithm is guaranteed to converge, in polynomial time to an ϵ-optimal solution a generic SDP.

4.1. Interior Point Methods for SDPs and the Central Path.

First formulated by Karmakar [14], interior point methods (IPM) have become widely adopted due to their guaranteed polynomial runtime [15]. These methods handle conic constraints by adding to the cost a "barrier" that tends to infinity when approaching the boundary of the feasible set. To prevent numerical instability, IPMs solve a sequence of optimization problems in which the barrier is weighted by a factor $1/t$, where t is increased until ϵ-optimality is reached. In the case of SDPs, the most widely used barrier function is the negative log-determinant, which leads to problems of the form:

$$\begin{aligned} X_t = \underset{X}{\text{minimize}} \quad & \text{Tr}\left(C^T X\right) - \frac{1}{t} \log\left(|X|\right) \\ \text{s.t.} \quad & \text{Tr}\left(A_i^T X\right) = b_i \quad i = 1, \ldots, M \end{aligned} \tag{14}$$

The curve defined by the optimizers X_t of (14) as a function of $t > 0$ is called the Central Path of the problem. As $t \to \infty$, X_t converges to the optimizer of (13),

X^*_{PSD}. Moreover, due to duality theory, the elements of the central path satisfy the following inequality:

$$\text{Tr}\left(C^T X_t\right) \geq \text{Tr}\left(C^T X^*_{PSD}\right) \geq \text{Tr}\left(C^T X_t\right) - N/t \tag{15}$$

which provides an optimality bound at any point in the path.

4.2. **DD and SDD relaxations of Semidefinite Programming.** In [16], the authors proposed a relaxation for general SDPs based on replacing the positive semidefinite constraints by lower complexity ones involving diagonally-dominant and scaled diagonally-dominant matrices, defined below:

Definition 3. *A symmetric matrix X with elements $X(i,j)$ is diagonally-dominant (DD) if*

$$X(i,i) \;\geq\; \sum_{j \neq i} |X(i,j)| \quad \forall i$$

Definition 4. *A symmetric matrix X is scaled diagonally-dominant (SDD) if there exist a positive diagonal matrix D and a DD matrix Y such that $X = DYD$.*

From these definitions it follows that $DD_N \subset SDD_N \subset PSD_N$, where DD_N, SDD_N and PSD_N denote the cones of $N \times N$ DD, SDD and Positive Semi-Definite (PSD) matrices. Thus, relaxations of the SDP (13) can be obtained by simply replacing the constraint $X \succeq 0$ with the stronger ones $X \in DD_N$ or $X \in SDD_N$. The following results, adapted from [17, 18] provides an alternative characterization of DD and SDD matrices that was used in [17] to show that these relaxations indeed lead to LPs or SOCPs with lower complexity than the original SDP. Define the mapping $\Psi_{i,j}$ from 2×2 matrices to $N \times N$ matrices:

$$\Psi_{i,j}(\mathcal{M}) = \bar{\mathcal{M}} \quad \text{where} \begin{cases} \bar{\mathcal{M}}(\{i,j\},\{i,j\}) = \mathcal{M} \\ 0 \;\; \text{otherwise.} \end{cases}$$

i.e. the $\{i,j\}$ sub-matrix of $\bar{\mathcal{M}}$ is $\mathcal{M}$, and the other entries of $\bar{\mathcal{M}}$ are 0. $\Psi_{i,j}(.)$ allows for characterizing DD and SDD matrices in terms of "exploded" 2×2 matrices as follows:

Lemma 3 ([17, 18])**.**

$$Y \in DD_N \iff Y = \sum_{i,j}^{N} \Psi_{i,j}(\mathcal{M}_{i,j}), \quad \mathcal{M}_{i,j} \in DD_2$$

Similarly,

$$Y \in SDD_N \iff Y = \sum_{i,j}^{N} \Psi_{i,j}(\mathcal{M}_{i,j}), \quad \mathcal{M}_{i,j} \succeq 0$$

Thus, enforcing the constraint $X \in DD_N (SDD)_N$ indeed reduces to a set of linear (second order cone) constraints.

4.3. **Iterative Basis Update.** While replacing the PSD constraint in (13) with the stronger one $X \in DD_N$ or $X \in SDD_N$ leads to a computationally cheaper optimization, the solution to these relaxed problems can be far from the true optimum. To alleviate this [17] proposed an iterative algorithm with improve performance, based on alternating between solving a sequence of DD/SDD problems and performing Cholesky factorizations. Briefly, the idea is to solve at step k a problem of the form

$$\begin{aligned} X_k^*\left(U_{k-1}\right) = \underset{X,Y}{\operatorname{argmin}} \ & \operatorname{Tr}\left(C^T X\right) \\ \text{s.t. } & \operatorname{Tr}\left(A_i^T X\right) = b_i \quad i = 1, \ldots, M \\ & U_{k-1}^T Y U_{k-1} = X, \quad Y \in DD_N \ / \ SDD_N \end{aligned} \tag{16}$$

where U_{k-1} is a Cholesky factor of the previous solution, e.g. $X_{k-1} = U_{k-1}^T U_{k-1}$. Note that since $I \in DD_N(SDD_N)$, the previous iterate X_{k-1}^*, is always a feasible solution of (16). Hence the algorithm generates a sequence of solutions with non-increasing cost. This sequence, however, is not guaranteed to converge to the optimizer of the original SDP (13). Indeed, in numerical tests the algorithm tends to converge to strictly suboptimal values for all medium to large size problems ($N >> 10$). This is precisely what motivated our research showing that indeed, it is possible to obtain an LP/SOCP based globally convergent algorithm.

4.4. **Solving SDPs by solving a sequence of LPs/SOCPs.** In this section we present a globally convergent algorithm for solving the SDP (13) based on DD and SDD programs. The algorithm is split in two phases. The first phase, the *decrease phase*, consists of solving a sequence of DD/SDD programs, exactly as in [17]. As noted above, this sequence tends to stagnate on a suboptimal objective cost as the iterates approach the boundary of the PSD cone and their conditioning worsens. To prevent this, a second phase of the algorithm that consists of a series of steps designed to improve the iterates' conditioning starts after the decrease phase. We call these steps *centering* steps, as they guide the iterates towards the center path of the SDP by solving a sequence of analytic centerings on the DD/SDD set. These centering steps constitute the *centering phase* of the algorithm.

Figure 1 shows a graphical description of the proposed algorithm. In the *decrease phase*, a sequence of problems of the form (16) are solved, decreasing the cost. After a fixed number s_d of decrease steps, the *centering phase* starts and a sequence of problems of the form (17) are solved :

$$\begin{aligned} X_l\left(U_{l-1}\right) = \underset{X,Y}{\operatorname{argmin}} \ & -\phi\left(Y\right) \text{ s.t.} \\ & \operatorname{Tr}\left(A_i^T X\right) = b_i \quad i = 1, \ldots, M \\ & \operatorname{Tr}\left(C^T X\right) = \operatorname{Tr}\left(C^T X_{l-1}\right) \\ & U_{l-1}^T Y U_{l-1} = X, \quad Y \in DD_N \ / \ SDD_N \end{aligned} \tag{17}$$

Here the function $-\phi(Y)$ is the logarithmic barrier of the DD/SDD sets and U_{l-1} is the Cholesky factor of X_{l-1}. The sequence of centering steps converges to a

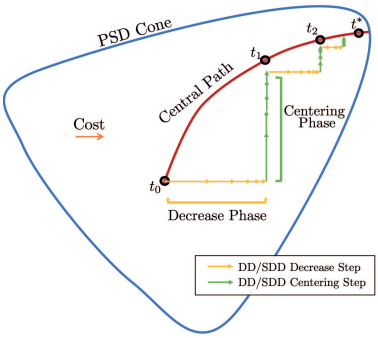


FIGURE 1. The algorithm alternates between a *cost decreasing phase* and a *centering phase* that brings the iterate to a point ϵ_c-close to the central path. After a finite number of phases, the iterate is brought to a point ϵ_c-close to the central path with parameter $t_\kappa \geq t^*$, guaranteeing ϵ_g-convergence to the solution of the SDP.

point ϵ_c close to the central path of the SDP, i.e. the optimizer of the analytic centering defined on the PSD cone. At the end of this sequence the conditioning of the current iterate has improved and due to the properties of the SDP central path, the optimality gap is given by (15). At this point, a new *decrease phase* starts and the algorithm keeps iterating between *decrease* and *centering* phases.

The main result in [13] shows that, under the following assumptions:

A1: The data matrices C and A_i all satisfy that:

$$\begin{aligned} \mathrm{Tr}\left(A_i^T A_j\right) = \mathrm{Tr}\left(A_i^T C\right) = 0 \quad & \forall 1 \leq i, j \leq M \\ ||C||_F = ||A_i||_F = 1 \quad & \forall i = 1, \ldots, M \end{aligned}$$

A2: There exists at least one feasible X for (13) such that $X \succ 0$ (Slater's condition).

A3: The cost function evaluated at the optimizer of (13) satisfies $\mathrm{Tr}(C^T X^*_{PSD}) > -\infty$.

the algorithm converges to a global optimum of the problem. Note that Assumption A1 can be made to hold trivially for any SDP by orthogonalization and projection; A2 guarantees strong duality; and A3 guarantees that the optimal cost is finite. Specifically, we have the following result:

Theorem 5. *Algorithm converges to ϵ_g-optimality in at most κ iterations of Decrease and Centering Phases, where κ is given by:*

$$\kappa = \left\lceil \frac{\log\left(N/\epsilon^*\right) - \log\left(t_0\right)}{\log\left(\chi\right)} \right\rceil \quad \chi = \frac{N\sqrt{1+\Theta}}{N\sqrt{1+\Theta} - \sqrt{\Phi}} > 1 \tag{18}$$

where Φ is a positive constant that satisfies $\Phi \leq \{2/(N+1), 1\}$ for the DD and SDD case, respectively, and Θ is a positive constant that depends only on the problem data.

In principle Sum-of-Squares optimization offers a powerful tool for establishing properties of the trajectories of a dynamical system. However, the poor scaling properties of SoS render this approach impractical beyond relatively low order systems. As we illustrated in this presentation using as an example the problem of data-driven verifiably safe control, computational complexity can be mitigated by the use of duality to obtain equivalent problems with a substantially reduced number of variables. Computational complexity can be further reduce through the use of the algorithm that we presented in the second portion of the talk. This algorithm can solve a generic SDP to optimality by solving a sequence of lower complexity Second Order Cone Programs or Linear Programs.

References

[1] S. Prajna and A. Jadbabaie, Safety verification of hybrid systems using barrier certificates, *Proceedings of the 7th International Workshop on Hybrid Systems: Computation and Control (HSCC)*, vol. 2993, Springer, 2004, pp. 477–492.

[2] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, Control barrier functions: Theory and applications, *Proceedings of the 18th European Control Conference (ECC)*, 2019, pp. 3420–3431.

[3] A. D. Ames, J. W. Grizzle, and P. Tabuada, Control barrier function based quadratic programs with application to adaptive cruise control, *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014, pp. 6271–6278.

[4] A. Majumdar, A. A. Ahmadi, and R. Tedrake, Control design along trajectories with sums of squares programming, *Proceedings of the 2013 IEEE International Conference on Robotics and Automation*, 2013, pp. 4054–4061.

[5] A. Clark, Verification and synthesis of control barrier functions, *Proceedings of the 2021 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 6105–6112.

[6] W. Xiao and C. Belta, Control barrier functions for systems with high relative degree, *Proceedings of the 2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 474–479.

[7] P. Jagtap, A. Swikir, and M. Zamani, Compositional construction of control barrier functions for interconnected control systems, *Proceedings of the 23rd International Conference on Hybrid Systems: Computation and Control*, 2020, pp. 1–11.

[8] A. Rantzer, A dual to Lyapunov's stability theorem, *Systems Control Lett.*, **42** (2001), no. 3, 161–168.

[9] A. Rantzer and S. Prajna, On analysis and synthesis of safe control laws, *42nd Allerton Conf. Commun., Control, Comput.*, 2004, pp. 1468–1476.

[10] S. Prajna and A. Rantzer, Convex programs for temporal verification of nonlinear dynamical systems, *SIAM J. Control Optim.*, **46**(3), 999–1021, 2007.

[11] J. Zheng, T. Dai, J. Miller, and M. Sznaier, Robust data-driven safe control using density functions, *IEEE Control Systems Lett.*, **7**, 2611–2616, 2023.

[12] T. Dai and M. Sznaier, A semi-algebraic optimization approach to data-driven control of continuous-time nonlinear systems, *IEEE Control Systems Lett.*, **5**(2), 487–492, 2020.

[13] B. Roig-Solvas and M. Sznaier, A globally convergent LP and SOCP-based algorithm for semidefinite programming, *arXiv preprint*, arXiv:2202.12374, 2022. Available at `https://arxiv.org/abs/2202.12374`.

[14] N. Karmarkar, A new polynomial-time algorithm for linear programming, in *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*, 302–311, 1984.

[15] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, 1994.

[16] A.A. Ahmadi and A. Majumdar, *DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization*, *SIAM J. Appl. Algebra Geom.*, **3**(2), 193–230, 2019.
[17] A.A. Ahmadi and G. Hall, *Sum of squares basis pursuit with linear and second order cone programming*, *Algebraic and Geometric Methods in Discrete Mathematics, Contemp. Math.*, **685**, 27–53, 2015.
[18] E.G. Boman, D. Chen, O. Parekh, and S. Toledo, *On factor width and symmetric H-matrices*, *Linear Algebra Appl.*, **405**, 239–248, 2005.

K Positivity Preservers and their Generators

Philipp Di Dio

In this talk we present very recent results on K-positivity preservers and their generators. These studies are motivated by a strange observation, namely that several non-negative polynomials which are not sums of squares become sums of squares under the heat equation.

The heat equation acts as a positivity preserver, i.e., a linear map from the polynomials to the polynomials such that non-negativity is preserved. It was an open problem for 14 years how all K-positivity preservers for general $K \subseteq \mathbb{R}$ look like. This question is fully answered in this talk.

Since the heat equation is generated by the Laplace operator we ask which other operators generate positivity preserver. We use the theory of Fréchet Lie groups and the Lévy-Khinchin formula to answer the question for the constant coefficient case. For the non-constant coefficient case we reduce it to the constant coefficient case.

We give an outline how these and future results shall be used in optimization, i.e., making non-negative polynomials continuously into sums of squares.

Theoretical and Practical Applications of Signomial Rings to Polynomial Optimization

Mareike Dressler
(joint work with Riley Murray)

Signomials generalize polynomials by allowing arbitrary real exponents at the expense of restricting the resulting function to the positive orthant. In [1], we introduce the concept of signomial rings and show how this and a newly presented signomial Positivstellensatz lead to a novel convex relaxation hierarchy of lower bounds of signomial and polynomial optimization. We provide numerical examples to illustrate the hierarchy and its performance on problems in chemical engineering, reaction networks, and optimal control (peak estimation.)

Signomial rings. 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 $e^{\boldsymbol{\alpha}}(\mathbf{x}) = \langle \boldsymbol{\alpha}, \mathbf{x} \rangle$. A *signomial* supported on a finite set $A \subseteq \mathbb{R}^n$ is a real-linear combination $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in A} c_{\boldsymbol{\alpha}} e^{\boldsymbol{\alpha}}(\mathbf{x})$. Its *support*, denoted by $\operatorname{supp}(f)$, is formally defined as the smallest set $A \subset \mathbb{R}^n$ for which $f \in \operatorname{span}\{e^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in A}$. A

posynomial is a signomial with only nonnegative terms. The *signomial ring* $\mathbb{R}[\mathcal{A}]$ is the $\mathbb{R}$-algebra generated by the basis functions $\{e^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathcal{A}}$.

Conditional SAGE as nonnegativity certificates. 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 is called *X-SAGE*. Detecting whether a signomial is X-SAGE boils down to a *relative entropy program*, a convex optimization problem. An important property of conditional SAGE is that they preserve sparsity, i.e., if f is X-SAGE supported on a set A, then f is a sum of X-AGE signomials, each supported on A, see [2, 3].

Positivstellensatz for X-SAGE. We prove the following Positivstellensatz, which is the first signomial Positivstellensatz to leverage conditional SAGE in the presence of nonconvex constraints and the first to permit irrational exponents.

Theorem 1 (Positivstellensatz for X-SAGE)**.** *Consider a compact convex set X, signomials $g_1,\dots,g_m\in\mathbb{R}[\mathcal{A}]$, and $K=\{\mathbf{x}\in X\,:\,g_i(\mathbf{x})\geq 0\ \text{ for all } i=1,\dots,m\}$. 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. We emphasize that neither X nor K (nor their images under exponential maps) need be semialgebraic.

Complete hierarchy of lower bounds. Given signomials $f,g_1,\dots,g_m$ and a closed convex set X, we want to compute

$$f_K^\star=\inf_{\mathbf{x}\in K} f(\mathbf{x}),\ \text{where } K=\{\mathbf{x}\in X\,:\,g_i(\mathbf{x})\geq 0\ \text{ for all } i=1,\dots,m\}.$$

As usual, we want to use our Positivstellensatz to compute lower bounds on $f_K^\star$. To search for an identity given in the Positivstellensatz, we have to decide on $r\in\mathbb{N}$ and permissible supports for the posynomials λ_i. Since signomials have no concept of "degree" that is central to polynomial optimization theory we reclaim it via artificially imposing an $\mathcal{A}$-degree on signomials. Hence, let $\mathcal{A}_d$ be the set of sums of at most d vectors from $\mathcal{A}$, then we define the *$\mathcal{A}$-degree* of a signomial f as $\deg_{\mathcal{A}}(f)=\inf\{d:\operatorname{supp}(f)\subseteq\mathcal{A}_d\}$.

With this, we can grade the certificates from the Positivstellensatz according to the largest $\mathcal{A}$-degree of the constituent signomials which leads us to bounds for the signomial optimization problem.

Definition 2. *If $r:=d-\deg_{\mathcal{A}}(f)\geq 0$, the* $\mathcal{A}$-degree d SAGE bound *is defined as*

$$f_K^{(d)}:=\sup\gamma\ \ \textit{such that}\ \ \left(\textstyle\sum_{\boldsymbol{\alpha}\in\mathcal{A}} e^{\boldsymbol{\alpha}}\right)^r(f-\gamma)=\lambda_0+\textstyle\sum_{i=1}^m\lambda_i\cdot g_i,$$

whith $\gamma\in\mathbb{R}$ and λ_0 and λ_i being X-SAGE signomials supported on $\mathcal{A}_d$ and $\operatorname{invsupp}_d(g_i)$*, respectively. If otherwise $d<\deg_{\mathcal{A}}(f)$, we set $f_K^{(d)}=-\infty$.*

Here, $\operatorname{invsupp}_d(g):=\{\boldsymbol{\alpha}\in\mathcal{A}_d\,:\,\boldsymbol{\alpha}+\operatorname{supp}(f)\subseteq\mathcal{A}_d\}$. Each bound $f_K^{(d)}$ can be computed via relative entropy programming.

Corollary 3. *The sequence* $f_K^{(1)}, f_K^{(2)}, \ldots$ *is nondecreasing and bounded above by* $f_K^\star$. *If the signomials* $f, g_1, \ldots, g_m$ *belong to* $\mathbb{R}[\mathcal{A}]$ *and* X *is compact, then*

$$\lim_{d\to\infty} f_K^{(d)} = f_K^\star.$$

This is the first completeness result for minimizing an arbitrary signomial subject to constraints given by a compact convex set and a conjunction of arbitrary (but finitely many) signomial inequalities. It is also the first completeness result for a hierarchy that uses conditional SAGE in the presence of nonconvex constraints.

For practical purposes it is an important open question how to choose the "best" signomial ring $\mathbb{R}[\mathcal{A}]$ resp. ground set $\mathcal{A}$.

References

[1] M. Dressler and R. Murray, Algebraic perspectives on signomial optimization, *SIAM J. Appl. Algebra Geom.*, **6** (2022), no. 4, 650–684.

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

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

Determining Multistationarity in n-site Phosphorilation using Sums of Nonnegative Circuit Polynomials

Timo de Wolff

Multisite phosphorylation is a signaling mechanism known to give rise to multiple steady states, a property termed multistationarity. When phosphorylation occurs in a sequential and distributive manner, one obtains a family of networks indexed by the number of phosphorylation sites n. We address the problem of understanding the parameter region where this family of networks displays multistationarity, by focusing on the projection of this region onto the set of kinetic parameters. The problem is reduced to studying whether a specific polynomial attains negative values over the positive orthant. The coefficients of the polynomial are symbolic, given as polynomials in terms of the kinetic parameters. We provide sufficient conditions for the polynomial to be positive and hence, preclude multistationarity, and also sufficient conditions for it to attain negative values and hence, enable multistationarity. We derive these conditions by exploiting the structure of the polynomial, its Newton polytope, and employing a certificate of nonnegativity called sums of nonnegative circuit polynomials (SONC).

Reporter: Ali Arslan

Participants

Prof. Dr. Amir Ali Ahmadi
Department of Operations Research
and Financial Engineering
Princeton University
Princeton, NJ 08540
UNITED STATES

Dr. Ali Arslan
ETH Zürich
Institute of Geophysics
Sonneggstrasse 5
8092 Zürich
SWITZERLAND

Dr. Nicolas Augier
LAAS-CNRS
7, av. du colonel Roche
31031 Toulouse Cedex 4
FRANCE

Dr. Benoit Bonnet-Weill
CNRS – L2S, CentraleSupélec
3, rue Joliot Curie
91190 Gif-sur-Yvette
FRANCE

Dr. Jason Bramburger
Dept. of Mathematics and Statistics
Concordia University
Montréal H3J1E1
CANADA

Dr. Elizabeth Carlson
Department of Computing and
Mathematical Sciences
California Institute of Technology
Pasadena, CA 91125
UNITED STATES

Prof. Dr. Sergei Chernyshenko
Imperial College London
Department of Aeronautics
City and Guilds Building
Exhibition Road
London SW7 2AZ
UNITED KINGDOM

Saroj Prasad Chhatoi
LAAS-CNRS
7, av. du colonel Roche
31400 Toulouse Cedex 4
FRANCE

Prof. Dr. Timo de Wolff
Institut für Analysis und Algebra
AG Algebra
TU Braunschweig
Universitätsplatz 2
38106 Braunschweig
GERMANY

Dr. Philipp di Dio
Fachbereich Mathematik u. Statistik
Universität Konstanz
Postfach 5560
78464 Konstanz
GERMANY

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

Dr. Giovanni Fantuzzi
Department Mathematik
Friedrich-Alexander-Universität
Erlangen-Nürnberg
Cauerstr. 11
91058 Erlangen
GERMANY

Prof. Dr. Emilia Fridman
School of Electrical Engineering
Tel Aviv University
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Dr. Federico Fuentes
Institute for Mathematical and
Computational Engineering (IMC)
Pontificia Universidad Católica de Chile
Vicuña Mackenna 4860
Macul, Santiago 7820436
CHILE

David Goluskin
Dept. of Mathematics & Statistics
University of Victoria
Victoria V8P 5C2
CANADA

Soodeh Habibi
Department of Electrical Engineering
and Electronics
University of Liverpool
Liverpool L69 3GJ
UNITED KINGDOM

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

Dr. Maria Infusino
Dipartimento di Matematica e
Informatica
Università degli Studi di Cagliari
Via Ospedale 72
09124 Cagliari
ITALY

Shucheng Kang
John A. Paulson School of Engineering
and Applied Science
Harvard University
Science Center
Akron 10 Street
Cambridge, MA 02138-2901
UNITED STATES

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

Lars-Luca Langer
Zukunftskolleg
Universität Konstanz
Postfach 216
78457 Konstanz
GERMANY

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

Shenyuan Ma
Artificial Intelligence Center
Faculty of Electrical Engineering
Czech Technical University in Prague
186 75 Praha 8
CZECH REPUBLIC

Dr. Swann Marx
LS2N
1 rue de la Noë
44300 Nantes
FRANCE

Jared Miller
Automatic Control Laboratory
ETH-Zentrum
ETL I 29
Physikstr. 3
8092 Zürich
SWITZERLAND

Prof. Dr. Antonis Papachristodoulou
Dept. Engineering Science
University of Oxford
Parks Road
Oxford OX1 3PJ
UNITED KINGDOM

Dr. Jeremy Parker
Department of Mathematics
University of Dundee
Dundee DD1 4HN
UNITED KINGDOM

Prof. Dr. Pablo A. Parrilo
Department of Electrical Engineering and Computer Science
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139-4307
UNITED STATES

David Rackl
Institut für Mathematik
Universität Klagenfurt
Universitätsstr. 65-67
9020 Klagenfurt
AUSTRIA

Dr. Rodolfo Rios-Zertuche
Department of Mathematics and Statistics
UiT The Arctic University of Norway
Postboks 953
9000 Tromsø
NORWAY

Corbinian Schlosser
INRIA Paris-Centre
2 Rue Simone IFF
75012 Paris
FRANCE

Prof. Dr. Markus Schweighofer
Fachbereich für Mathematik und Statistik
Universität Konstanz
78457 Konstanz
GERMANY

Prof. Dr. Michael Stingl
Lehrstuhl für Angewandte Mathematik
Department Mathematik
Friedrich-Alexander-Universität Erlangen-Nürnberg
Cauertstr. 11
91058 Erlangen
GERMANY

Prof. Dr. Mario Sznaier
Department of Electrical and Computer Engineering
Northeastern University
360 Huntington Avenue
Boston MA 02115-5000
UNITED STATES

Dr. Matteo Tacchi-Bénard
GIPSA-lab Grenoble Images Parole Signal Automatique
UMR 5216 CNRS – Grenoble INP
Université Grenoble Alpes
11 rue des Mathématiques
38402 Saint-Martin-d'Hères
FRANCE

Prof. Dr. Ian Tobasco
Rutgers University
Department of Mathematics
110 Frelinghuysen Road
Piscataway, NJ 08854
UNITED STATES

Prof. Dr. Giorgio Valmorbida
L2S, CentraleSupélec
3 rue Joliot Curie
91190 Gif-sur-Yvette
FRANCE

Dr. Luis Felipe Vargas
IDSIA
Via Santa 1
6962 Lugano
SWITZERLAND

Dr. Jie Wang
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Haidan District
No. 55 Zhongguancun East Road
Beijing 100 190
CHINA

Dr. Stephan Weis
Czech Technical University in Prague
Karlovo náměstí 13
12000 Praha 2
CZECH REPUBLIC

Dr. Andrew Wynn
Department of Aeronautics
Imperial College London
South Kensington Campus
London SW7 2AZ
UNITED KINGDOM

Dr. Heng Yang
John A. Paulson School of Engineering and Applied Sciences
Harvard University
150 Western Avenue
Cambridge, MA 02134
UNITED STATES

Christian Mugisho Zagabe
Department of Mathematics
The University of Namur
FUNDP
61, rue de Bruxelles
5000 Namur
BELGIUM