Mathematisches Forschungsinstitut Oberwolfach

Report No. 35/2023

DOI: 10.4171/OWR/2023/35

Mixed-integer Nonlinear Optimization: A Hatchery for Modern Mathematics

Organized by
Leo Liberti, Palaiseau
Sebastian Sager, Magdeburg
Angelika Wiegele, Klagenfurt

13 August – 18 August 2023

Abstract. Mixed-integer nonlinear programming (MINLP) is concerned with finding optimal solutions to mathematical formulations of optimization problems combining discrete and nonlinear phenomena. The scientific program was organized around three areas: convex envelopes and relaxation hierarchies, mixed-integer optimal control, and current trends. These topics were addressed with a variety of tutorials, talks, and short research announcements.

Mathematics Subject Classification (2020): 90-06, 90C11, 90C22, 90C26, 90C30.

Introduction by the Organizers

This report refers to the third edition of the workshop *Mixed-Integer Nonlinear Optimization: A Hatchery for Modern Mathematics* organized at Oberwolfach by Leo Liberti (Palaiseau), Sebastian Sager (Magdeburg), and Angelika Wiegele (Klagenfurt). There were 46 participants at the MFO and additional 5 participants taking the opportunity to attend online.

The workshop was organized in 4 tutorial talks (one each day from Monday to Thursday, one hour long, including 15 minutes for questions), 19 "normal" talks (45 minutes long, including 15 minutes for questions), and 15 short research announcements (SRA – 15 minutes long, including 5 minutes for questions). The discussion after practically all talks was lively and filled with questions, remarks and suggestions from many attendees. As Oberwolfach tradition warrants, we spent wednesday afternoon hiking to St. Roman, a little more than 7 km away from the Institute.

1. Participants

The 51 participants are from 11 countries, distributed as follows:

DE	US	AT	FR	NL	IT	SE	BR	CA	UK	CH
21	12	4	3	3	2	2	1	1	1	1

The participants have affiliations not only with mathematics departments, but also with computer science and engineering departments. This exhibits a notable attraction towards a diverse array of cultures.

Gender balance. Genderwise, this workshop attracted 15 women (29% of the total number of participants), 9 of which gave talks (in particular, two tutorials and seven normal talks were given by women). We received a particularly large number of cancellations from female colleagues at this workshop, which is probably due to the fact that it was during the vacation season.

2. Scientific Areas

Many relevant practical decision problems in energy, engineering, economics, medicine, and systems biology can be formulated as a MINLP. But it is also a research area that by its very nature touches many areas of mathematics. The links are bidirectional: many of the advances in MINLP stem from neighboring mathematical areas. And the research in MINLP theory and algorithms, as well as the application of MINLP software to mathematical problems, may yield deep mathematical insights.

The scientific organization of this workshop was divided into three main areas. I) Convex envelopes and relaxation hierarchies; II) Mixed-integer optimal control; and III) Current trends.

2.1. **Convex envelopes and relaxation hierarchies.** Convex envelopes of ever more complicated mixed-integer nonlinear sets provide the foundation of most solution algorithms for MINLP. Relaxation hierarchies were introduced to study finite dimensional nonlinearity by means of infinitely-dimensional but linear entities: they are now mature enough to be be computationally viable. Convex envelopes and relaxation hierarchies are related insofar as both provide relaxations of the original MINLP, giving rise to a plethora of algorithmic tools for solving MINLP (from starting points to provable bounds).

2.2. **Mixed-integer optimal control.** MINLP has historically many links to systems engineering and optimal control. Finding optimal control strategies for nonlinear processes is often difficult due to high dimensions resulting from fine control discretizations. Deriving efficient algorithms from theoretical arguments in weak topologies has resulted in a research boost that we want to continue supporting, with many open questions and opportunities for the future.

2.3. **Current trends.** We identified new, exciting, and promising MINLP research areas, most of which are still too young to be promoted to "topic", yet too important to be ignored: (a) The growing link between MINLP and Machine Learning (ML); (b) The long-awaited attack on bi-level MINLP (problems where one constraint requires some first-level variables to belong to the set of optima of a second-level MINLP), a problem class of formidable difficulty; (c) Computational complexity and (in)approximability; (d) Challenges arising from applications in power flow.

3. Short Research Announcements

There were three sessions of Short Research Announcements, one on Tuesday and two on Thursday.

- Amir Ali Ahmadi presented results on the complexity of finding local minima in polynomial optimization, including a representation theorem about the set of local minima of any cubic polynomial is a semidefinite set. He also presented work on higher-order Newton methods with polynomials.
- Dominik Cebulla demonstrated that applying mixed-integer optimal control for chromatography can reduce the batch-cycle time significantly.
- Julius Martensen talked about automated model discovery and presented a MINLP formulation for symbolic regression.
- Second-order partial outer convexification for switched dynamical systems was presented by Christoph Plate.
- A lot of questions and discussion arose after Ruth Misener's presentation about optimizing over trained graph neural networks.
- Sven Leyffer's short research announcement was titled "Beyond Mixed-Integer Nonlinear Optimization" and was about applications with massive nonlinear MIPs and data analytics for large gamma-ray spectrometers.
- Volker Kaibel proved a connection between the diameter of the polyhedron of feasible solutions and the complexity of a linear program.
- Diane Puges introduced strong SDP based bounds on the cutwidth of a graph as well as a cutting-plane algorithm to compute these bounds.
- Using Copositive Duality for the Discrete Energy Market has been addressed by Merve Bodur.
- Daniel Brosch asked the question "Is the Set of Trees Convex?" where he extends Razborov's flag algebras, allowing the application of sums-of-squares and moment techniques to tackle this question.
- Nick Sahinidis reported about the status of the BARON project.
- Jan Schwiddessen presented his work on machine learning, namely how to use semidefinite programming for semi-supervised support vector machines.
- Driven by an application from chromatography, Jan Rolfes exhibits a safe approximation of distributionally robust optimization depending on univariate indicator functions.

- Marc Pfetsch talked about solving Mixed-Integer Semidefinite Programs and presented numerical results on using SCIP-SDP, an open-source solver for mixed-integer semidefinite programs.
- Finally, Florian Rösel talked dabout Optimality Certificates for Convex MINLPs.

4. The Future

This is the third edition of MINLP workshops at Oberwolfach. Again, the participants were enthusiastic and expressed curiosity regarding potential future editions. While there exist various MINLP workshops organized by different members of our community, the Oberwolfach workshops stand out for their strong mathematical relevance. Our intention is to continue applying for MINLP workshops at Oberwolfach and hope that this will be acknowledged by the MFO Scientific Committee.

Leo Liberti volunteered to step down from the organizing committee. We are grateful to Jon Lee for agreeing to take his place, as well as for helping to organize the 2023 edition of the workshop.

Workshop: Mixed-integer Nonlinear Optimization: A Hatchery for Modern Mathematics

Table of Contents

Abstracts

A Walk Through the ACOPF World

Merve Bodur

In this survey-type talk, we overview the developments around the alternating current optimal power flow problem (ACOPF). Starting with the basic concepts about the power flow, we go over the main components of ACOPF formulations, then detail the commonly studied convex relaxations as well as approximations, briefly mention techniques to generate feasible solutions along with available software, and lastly mention some extensions of the problem.

1. Basic concepts

ACOPF models rely on adequately representing the actual power flow physics. In that regard, the power law, Ohm's law, Kirchhoff's circuit laws can be deemed as the most important aspects. Power is the product of current and voltage, which are respectively defined as the rate at which charge (i.e., the basic entity of electricity) is flowing and the difference in charge (potential energy) between two points in a circuit. Ohm's law defines voltage as the product of current and resistance, whereas Kirchhoff provides two circuit laws: (i) total current entering a junction is exactly equal to the current leaving the node, and (ii) in any closed-loop network, the sum of voltage drops around the loop is equal to zero.

A power system consists of generators, transmission system, and loads. A transmission system is modelled as a directed graph with a node for each bus of the system, which may be a load as well as host generators, and with links for transmission lines (a pair of anti-parallel arcs for each line).

2. Formulations

ACOPF aims to generate sufficient power to satisfy demand at minimum cost, deciding on power generation amounts and power flow on the network that abide by the laws of physics.

Due to the use of alternating current, stemming from the magnetic field generated by a magnetic wheel rotation, the voltage, current, and accordingly power over time are represented by sinusoidal curves. Given the frequency, amplitude and rotation angle of such a curve, the instantaneous value at any time can be simply calculated using the curve defining formula. However, solving ACOPF models using such time-dependent decision variables would be computationally infeasible for practical cases. As an acceptable compromise, the ACOPF literature considers time-independent but *complex* decision variables. The use of complex variable form for all the current, voltage, and power decisions yields a very good representation of the actual power flow. More specifically, all the important laws of physics still apply, the steady-state average power can be readily calculated as the real part of the complex power variables (called real power, which is the actual quantity used by the load), and the reactive power (which is essential for the power system to

run, and need to be managed to maximize the throughput of active power of the system) is captured by the imaginary part of the complex power variables.

ACOPF formulations are based on the two fundamental characteristics: (i) power law that links power to voltage and current variables, and (ii) relationship between voltage and current variables. For the latter, a linear relationship is employed, where the coefficients in the associated equations are obtained through the admittance matrix, a complex-valued matrix constructed using the electrical parameters and network topology (where a Π-circuit model is commonly used for each transmission line).

The main components of ACOPF formulations are the objective function usually minimizing power generation cost, bounds on most of the decision variables (e.g., voltage magnitudes, difference of voltage angles between adjacent buses, maximum power magnitudes injected on lines, and generated power), explicit constraints on power balance at each bus and the two aforementioned fundamental characteristics (i) and (ii), and other technical constraints (including the voltage angle fixing for a reference bus). The feasible space of ACOPF models is usually highly non-convex, and the practical cases mostly have multiple local optima, as such ACOPF is a very challenging problem to solve.

There are various ACOPF formulations in the literature. They are usually classified with respect to the use of bus or line power injection, the choice of the coordinate system for the decision variables (namely, rectangular or polar), and the existance of the current variables in addition to the voltage variables. (The literature presents some special treatment to the radial, i.e., tree-based, network cases.) Based on the class of the formulation, different complexities are involved in the models, e.g., different sources and forms of nonlinearity and existence of trigonometric functions. Many of the distinctions among ACOPF relaxations and approximations are related to various mechanisms for exploiting these different mathematical features.

3. Convex relaxations

ACOPF relaxations are typically chosen to be convex to leverage the existing convex optimization tools. By design, relaxations provide valid bounds on the optimal objective value of the problem. Moreover, the literature provides some sufficient conditions for certifying global optimality of relaxation solutions.

Commonly studied relaxations fall into the following three classes:

(1) Semidefinite programming (SDP) relaxations: Shor relaxation, moment/sum-of-square relaxation hierarchies, some other hierarchies
(2) Second-order cone programming (SOCP) relaxations: Jabr's relaxation, strong SOCP relaxation, quadratic convex (QC) relaxation
(3) Linear programming (LP) relaxations

Dominance (or its lack of) among different relaxations has been established in the literature. For computational efficiency, relaxations are often further strengthened by applying bound tightening and adding valid inequalities (or cuts in a

cutting-plane framework). Moreover, for further tractability, sparsity of power networks has also been leveraged.

4. Approximations

Approximations assume specific quantities for certain decisions to simplify the ACOPF model. However, those quantities are usually chosen inspired by typical operating conditions for power systems, as such practically useful solutions can be obtained based on approximation models. Approximations based on SOCP and LP are most commonly proposed in the literature.

5. Feasible Solutions

In practice, locally optimal solutions are often acceptable. There exist many approaches to generate feasible solutions, including SDP-based, SCOP-based, convex restriction based, and practically most efficient ones as local solution techniques, most notably interior point methods. Local solution algorithms can benefit from the outputs resulting from problem relaxations and approximations (e.g., using the decision variable values and the set of binding constraints for initialization).

6. Software and Extensions

For interior point methods, Knitro and IPOPT have been some excellent choices. There also exist some specialized tools such as `MATPOWER` and `PowerModels.jl`. Recently, Gurobi released a new module called OptiMods including ACOPF related content.

There is a vast (in particular) recent literature considering extensions related to ACOPF, some of which adding integer decision variables such as the reactive OPF (which considers switchable shunts and adjustable tap ratios) and transmission switching (which considers turning off some lines). Security constraints constitute an important aspect to ensure solution robustness. Integration of ACOPF with other decision-making problems such as unit commitment yields further challenges to be addressed in the literature. Furthermore, incorporation of uncertainty and efficient implementations (e.g., by means of parallelization or with the help of GPUs) have been considered as valuable directions. Lastly, ARPA-E Grid Optimization Competitions have been pushing researchers to incorporate many interesting practical features to optimization models.

References

[1] Molzahn, D. K., & Hiskens, I. A. (2019). A survey of relaxations and approximations of the power flow equations. *Foundations and Trends in Electric Energy Systems*, 4(1-2), 1–221.

[2] Bienstock, D., Escobar, M., Gentile, C., & Liberti, L. (2022). Mathematical programming formulations for the alternating current optimal power flow problem. *Annals of Operations Research*, 314(1), 277–315.

Perfect graphs and algebraic certificates of nonnegativity

Amir Ali Ahmadi
(joint work with Cemil Dibek)

In this talk, we present a precise connection between the notion of *"perfectness"* in graph theory, and the notion of a *"sum of squares"* decomposition in the study of nonnegative polynomials. We also discuss some consequences of this connection.

A graph is *perfect* if for each of its induced subgraphs, the chromatic number equals the cardinality of a largest clique. Perfect graphs, introduced by Berge in 1960, have elegant theoretical properties and curious connections with linear, integer, and semidefinite programming. For instance, perfect graphs appear in the study of exactness of linear programming relaxations of integer programs. As an example, for a matrix $A \in \{0,1\}^{m\times n}$, all vertices of the polytope $\{x \in \mathbb{R}^n : Ax \leq 1, x \geq 0\}$ are integral if and only if the undominated rows of A are the incidence vectors of the maximal cliques of a perfect graph. Moreover, several combinatorial problems that are NP-hard on general graphs can be solved efficiently on perfect graphs using semidefinite programming. Examples include the maximum independent set and the minimum clique cover problems. More generally, perfect graphs have been the subject of much research in recent decades due to the fact that they are at the crossroad of several mathematical disciplines, including graph theory, information theory, combinatorial optimization, polyhedral and convex geometry, and semidefinite programming.

The second notion of interest to this work is that of sum of squares polynomials. A polynomial is a *sum of squares* (sos) if it can be written as a sum of squares of some other polynomials. There has been a growing interest in sos polynomials recently due to the fact that they provide semidefinite programming-based sufficient conditions for problems involving nonnegative polynomials. It is well known that several important problems in applied and computational mathematics can be formulated as optimization problems over the set of nonnegative polynomials. Although these problems are generally intractable to solve exactly, they can be efficiently approximated by replacing nonnegativity constraints with sum of squares requirements. By connecting ideas from real algebraic geometry and semidefinite programming, sum of squares polynomials have significantly impacted both discrete and continuous optimization over the last two decades.

In this work, we introduce and study the notion of sos-perfectness, a notion that brings together perfect graphs and sos polynomials. For a graph $G = (V, E)$ with clique number $\omega(G)$, we define the following quartic (homogeneous) polynomial in the variables $x = (x_1, \ldots, x_{|V(G)|})^T$:

$$p_G(x) = -2\omega(G) \sum_{ij\in E(G)} x_i^2 x_j^2 + (\omega(G) - 1) \left(\sum\nolimits_{i=1}^{|V(G)|} x_i^2 \right)^2 . \tag{1}$$

It turns out that for every graph G, the polynomial $p_G(x)$ is nonnegative by construction. We say that a graph G is *sos-perfect* if $p_H(x)$ is sos for every induced subgraph H of G. We prove the following theorem.

Theorem 1. *A graph is perfect if and only if it is sos-perfect.*

We remark that our proof of this theorem does not rely on the celebrated strong perfect graph theorem (which states that a graph is perfect if and only if it does not contain an odd hole or an odd antihole).

In 1888, Hilbert proved the existence of nonnegative polynomials that are not sums of squares. The first examples of such polynomials were constructed by Motzkin and Robinson many decades later. Understanding the distinction between nonnegative polynomials and sos polynomials continues to be an active area of research today. In relatively low degrees and dimensions, constructing examples of nonnegative polynomials that are not sos seems to be a nontrivial task.

In this work, we use Theorem 1 to explicitly construct several infinite families of nonnegative polynomials that are not sos. These polynomials come from special families of non-perfect graphs such as powers of cycles, Paley graphs, and Mycielski graphs. We also show that certain graph operations—adding edges or verticies without increasing the clique number, strong graph products, and graph joins—can be used to generate even more nonnegative polynomials that are not sos.

We end by noting that by Theorem 1, the following is an algebraic reformulation of the strong perfect graph theorem:

If a graph G does not contain an odd hole or an odd antihole, then the (nonnegative) form $p_G(x)$ defined in (1) is a sum of squares.

It would be interesting to find an algebraic proof of the strong perfect graph theorem based on this reformulation. The interested reader can find more details in [1].

References

[1] A.A. Ahmadi and C. Dibek, *A sum of squares characterization of perfect graphs*, SIAM Journal on Applied Algebra and Geometry (2023).

On the convex hull of convex quadratic optimization problems with indicators

Simge Küçükyavuz

(joint work with Linchuan Wei, Alper Atamtürk, Andrés Gómez)

Given a symmetric positive semidefinite matrix $Q \in \mathbb{R}^{n\times n}$, and set $Z \subset \{0,1\}^n$, we consider the mixed-integer nonconvex set

$$X = \left\{(x,z,t) \in \mathbb{R}^n \times Z \times \mathbb{R} : t \geq x^\top Q x,\ x_i(1-z_i) = 0,\ \forall i = 1,\dots,n\right\}.$$

Set X is associated with the mixed-integer convex quadratic optimization problem with indicator variables (MIQO) where the objective function is a convex quadratic function, and there is an indicator variable $z_i, i = 1,\dots,n$ that turn "on/off" the associated continuous variable x_i. MIQO finds various applications in statistical learning with sparsity, portfolio optimization, electricity power production, and so on. A critical step toward solving MIQO is to convexify the set X. Indeed, MIQO is equivalent to minimizing a linear objective function over the closure of the

convex hull of X (clconv(X)). However, MIQO is $\mathcal{NP}$-hard even if $Z = \{0,1\}^n$ [1]. Thus, a simple description of clconv(X) is, in general, not possible unless $\mathcal{NP} = \text{Co} - \mathcal{NP}$.

An important class of convex relaxation of X that has received attention in the literature is obtained by decomposing matrix $Q = \sum_{i=1}^{\ell} \Gamma_i + R$, where Γ_i, $i = 1, \ldots, \ell$ are "simple" and R and Γ_i's are positive semidefinite (psd) matrices. Then

$$(1) \qquad t \geq x^\top Q x \iff t \geq \sum_{i=1}^{\ell} \tau_i + x^\top R x, \text{ and } \tau_i \geq x^\top \Gamma_i x, \ \forall i \in \{1, \ldots, \ell\},$$

and each constraint $\tau_i \geq x^\top \Gamma_i x$ is replaced with a system of inequalities describing the convex hull of the associated mixed-integer set. This idea was originally used in [3] where $\ell = n$, Γ_i's are diagonal matrices, and constraints $\tau_i \geq d_i x_i^2$ are strengthened using the perspective relaxation [2], i.e., reformulated as $\tau_i \geq d_i \frac{x_i^2}{z_i}$. A generalization of the above approach is rank-one decomposition, which lets $\Gamma_i = h_i h_i^\top$ be a rank-one matrix [7, 8, 9]. Alternative generalizations of perspective relaxation that have been considered in the literature include exploiting substructures based on 2×2 matrices [4, 5] or tridiagonal [6]. In this work, we show that clconv(X) can be described in a compact extended formulation with $\mathcal{O}(n^2)$ additional variables with linear constraints and a single positive semidefiniteness constraint. We also characterize clconv(X) in the original space of variables.

Given a matrix $W \in \mathbb{R}^{p \times q}$, its Moore-Penrose inverse is denoted as $W^\dagger$, and W_S is the submatrix of W induced by S for some $S \subset \{1, \ldots, n\}$. Let $\hat{W}_S \in \mathbb{R}^{n \times n}$ be the $n \times n$ matrix obtained from W_S by filling the missing entries with zeros. We first provide a representation of clconv(X) under the assumption that Q is positive definite. Given Q positive definite, define the polytope $P \subset \mathbb{R}^{n+n^2}$ as

$$P = \left(\left\{ (\hat{e}_S, \hat{Q}_S^\dagger) \right\}_{\hat{e}_S \in Z} \right),$$

where e denotes the vector of all ones.

Theorem 1. *Let Q be a positive definite matrix. Then*

$$\textit{clconv}(X) \ = \ \{(z, x, t) \in [0,1]^n \times \mathbb{R}^{n+1} \mid \exists W \in \mathbb{R}^{n \times n} s.t. \begin{pmatrix} W & x \\ x^\top & t \end{pmatrix} \succeq 0, (z, W) \in P\}.$$

When Q is positive semidefinite, there exists some $F \in \mathbb{R}^{n \times k}$ such that $Q = FF^\top$. Define $F_S \in \mathbb{R}^{S \times k}$ as the submatrix of F corresponding to the rows indexed by S, and let $\hat{F}_S \in \mathbb{R}^{n \times k}$ be the matrix obtained by filling the missing entries with zeros. Define the polytope $P_F \subset \mathbb{R}^{n+k^2}$ as

$$P_F = \text{conv} \left(\left\{ (\hat{e}_S, \hat{F}_S^\dagger \hat{F}_S) \right\}_{\hat{e}_S \in Z} \right).$$

Let $\pi_S : \mathbb{R}^n \to \mathbb{R}^S$ be the projection onto the subspace indexed by S and col(W) denote the column space of W.

Theorem 2. *Let $Q = FF^\top$, where $F \in \mathbb{R}^{n\times k}$ is a full-column rank matrix satisfying col$(F) = \bigcap_{\hat{e}_S \in Z} \pi_S^{-1}(col(F_S))$. Then*

$$clconv(X) = \{(z,x,t) \in [0,1]^n \times \mathbb{R}^{n+1} \mid \exists W \in \mathbb{R}^{k\times k},\ \begin{pmatrix} W & F^\top x \\ x^\top F & t \end{pmatrix} \succeq 0,\ (z,W) \in P_F\}.$$

An immediate consequence of Theorem 2 is that if Q is rank-deficient, i.e., $k < n$, then the extended formulation describing clconv(X) is simpler than the full rank case, i.e., it has fewer additional variables and a lower-dimensional semidefiniteness constraint.

We now turn our attention to describing clconv(X) in the original space of variables. The main task is projecting out the matrix variable W in the extended formulation of clconv(X) given in Theorem 1 for Q positive definite. Identical arguments hold for the representation in Theorem 2 for low-rank matrices.

Suppose that a minimal description of polyhedron P is given by the facet-defining inequalities

$$\langle \Gamma_i, W\rangle - \gamma_i^\top z \le \beta_i, \quad i = 1,\dots,m_1,$$

and equalities

$$\langle \Gamma_i, W\rangle - \gamma_i^\top z = \beta_i, \quad i = m_1+1,\dots,m,$$

where $\Gamma_i \in \mathbb{R}^{n\times n}$, $\beta_i \in \mathbb{R}$ and $\gamma_i \in \mathbb{R}^n$. We define a set $\mathcal{Y}$ as

$$\mathcal{Y} = \left\{ y \in \mathbb{R}_+^{m_1} \times \mathbb{R}^{m-m_1} : \sum_{i=1}^m \Gamma_i y_i \succeq 0,\ \sum_{i=1}^m \mathrm{Tr}(\Gamma_i) y_i \le 1 \right\}.$$

Theorem 3. *If Q is positive definite, point $(x,z,t) \in clconv(X)$ if and only if $z \in conv(Z)$, $t \ge 0$ and*

$$t \ge \frac{x^\top \left(\sum_{i=1}^m \Gamma_i y_i\right) x}{y^\top \beta + \left(\sum_{i=1}^m y_i \gamma_i\right)^\top z}, \qquad \forall y \in \mathcal{Y}, \tag{2}$$

or equivalently,

$$t \ge \max_{y\in\mathcal{Y}} \frac{x^\top \left(\sum_{i=1}^m \Gamma_i y_i\right) x}{y^\top \beta + \left(\sum_{i=1}^m y_i \gamma_i\right)^\top z}. \tag{3}$$

An analogous result holds for low-rank matrices, where $(\Gamma_i, \gamma_i, \beta_i)$, $i \in \{1,...,m\}$ defines P_F.

From Theorem 3, we see that clconv(X) can be described by an infinite number of fractional quadratic/affine inequalities. More importantly, the convex hull is finitely generated: the infinite number of quadratic and affine functions are obtained from conic combinations of a *finite* number of base matrices Γ_i and vectors (γ_i, β_i), which correspond precisely to the minimal description of P.

References

[1] X. Chen, D. Ge, Z. Wang, and Y. Ye. *Complexity of unconstrained $L_2 - L_p$ minimization.* Mathematical Programming, 143(1):371–383, 2014.

[2] A. Frangioni and C. Gentile. *Perspective cuts for a class of convex 0–1 mixed integer programs*, Mathematical Programming, 106:225–236, 2006.

[3] A. Frangioni and C. Gentile. *SDP diagonalization and perspective cuts for a class of nonseparable MIQP*, Operations Research Letters, 35:181–185, 2007.
[4] A. Frangioni, C. Gentile and J. Hungerford. *Decompositions of Semidefinite Matrices and the Perspective Reformulation of Nonseparable Quadratic Programs*, Mathematics of Operations Research, 45(1):15–33, 2020.
[5] S. Han, A. Gómez and A. Atamtürk. *2x2 convexifications for convex quadratic optimization with indicator variables*, Mathematical Programming, 2023. doi: 10.1007/s10107-023-01924-w. Online First Article.
[6] P. Liu, S. Fattahi, A. Gómez and S. Küçükyavuz. *A Graph-based Decomposition Method for Convex Quadratic Optimization with Indicators*, Mathematical Programming, 2022. doi: 10.1007/s10107-022-01845-0. Article in Advance.
[7] A. Atamtürk and A. Gómez. *Rank-one Convexification for Sparse Regression*, arXiv preprint arXiv:1901.10334, 2019.
[8] L. Wei, A. Gómez and S. Küçükyavuz. *On the Convexification of Constrained Quadratic Optimization Problems with Indicator Variables*, Lecture Notes in Computer Science, vol 12125, 433–447, Springer, Cham, 2020.
[9] L. Wei, A. Gómez and S. Küçükyavuz. *Ideal formulations for constrained convex optimization problems with indicator variable*, Mathematical Programming, 192(1-2):57–88, 2022.

Utilising Deep Neural Networks in Mixed Integer Programming Problems

Shudian Zhao
(joint work with Jan Kronqvist, Boda Li, Calvin Tsay, Jan Rolfes)

This work presents a common way of encoding a Deep Neural Network (DNN) by mixed-integer programming techniques and two applications solved by algorithms based on this technique. This talk starts with a brief introduction to the application of DNNs in image classification and the architecture of a common example. Then, the technique of encoding DNNs with mixed-integer programming with the big-M formulation is illustrated.

The first part of the talk presents the results from the paper "Model-based feature selection for neural networks: A mixed-integer programming approach", which is co-authored with Jan Kronqvist and Calvin Tsay. In this paper, we develop a novel input feature selection framework for ReLU-based DNNs, which builds upon a mixed-integer optimization approach. While the method is generally applicable to various classification tasks, we focus on finding input features for image classification for clarity of presentation. The idea is to use a trained DNN, or an ensemble of trained DNNs, to identify the salient input features. The input feature selection is formulated as a sequence of mixed-integer linear programming (MILP) problems that find sets of sparse inputs that maximize the classification confidence of each category. These "inverse" problems are regularized by the number of inputs selected for each category and by distribution constraints. Numerical results on the well-known MNIST and FashionMNIST datasets show that the proposed input feature selection allows us to drastically reduce the input size to $\sim$15% while maintaining a good classification accuracy. This allows us to design DNNs with significantly fewer connections, reducing computational effort and producing DNNs that are more robust toward adversarial attacks.

In the next part, this talk presents the method that is introduced in "Alternating mixed-integer programming and neural network training for approximating stochastic two-stage problems", which is co-authored with Jan Kronqvist, Boda Li, and Jan Rolfes. In this paper, the presented work addresses two-stage stochastic programs (2SPs), a broadly applicable model to capture optimization problems subject to uncertain parameters with adjustable decision variables. In case the adjustable or second-stage variables contain discrete decisions, the corresponding 2SPs are known to be NP-complete. The standard approach of forming a single-stage deterministic equivalent problem can be computationally challenging even for small instances, as the number of variables and constraints scales with the number of scenarios. To avoid forming a potentially huge MILP problem, we build upon an approach of approximating the expected value of the second-stage problem by a neural network (NN) and encoding the resulting NN into the first-stage problem. The proposed algorithm alternates between optimizing the first-stage variables and retraining the NN. We demonstrate the value of our approach with the example of computing operating points in power systems by showing that the alternating approach provides improved first-stage decisions and a tighter approximation between the expected objective and its neural network approximation.

This talk concludes with a discussion about advanced formulations for encoding DNNs with mixed-integer programming and generalizing the techniques in solving other optimization problems such as Capacitated Facility location, Investment, Stochastic server location, and pooling problems.

Branch-and-bound for D-Optimality with fast local search and variable-bound tightening

Marcia Fampa

(joint work with Jon Lee, Gabriel Ponte)

We develop a branch-and-bound algorithm for the D-optimality problem, a central problem in statistical design theory, based on two convex relaxations, employing variable-bound tightening and fast local-search procedures, testing our ideas on randomly-generated test problems.

The D-Optimality problem is formulated as

$$\text{(D-Opt)} \qquad z := \max\left\{\operatorname{ldet}\textstyle\sum_{\ell\in N} x_\ell v_\ell v_\ell^\top \,:\, \mathbf{e}^\top x = s,\ l \le x \le u,\ x \in \mathbb{Z}^n\right\}$$
$$= \max\left\{\operatorname{ldet}\left(\textstyle\sum_{\ell\in N} l_\ell v_\ell v_\ell^\top + \sum_{\ell\in N} x_\ell v_\ell v_\ell^\top\right) : \mathbf{e}^\top x = s - \mathbf{e}^\top l\,,\ 0 \le x \le u - l,\ x \in \mathbb{Z}^n\right\},$$

where $v_\ell \in \mathbb{R}^m$, for $\ell \in N := \{1,\dots,n\}$, with $m \le s < n$ natural numbers, and $0 \le l < u \in \mathbb{Z}^n$, with $\mathbf{e}^\top l \le s \le \mathbf{e}^\top u$. It will be very useful to define $A := (v_1, v_2, \dots, v_n)^\top$ (which we always assume has full column rank), and so we have $\sum_{\ell\in N} x_\ell v_\ell v_\ell^\top = A^\top \operatorname{Diag}(x) A$.

D-Opt is a fundamental problem in statistics, in the area of "experimental designs" (see [23], for example). Ideally we would be considering the least-squares

regression problem $\min_{\theta\in\mathbb{R}^m} \|A_u\theta - b_u\|_2$, where $A_u \in \mathbb{R}^{\mathbf{e}^\intercal u\times m}$ has $v_\ell^\intercal$ repeated u_ℓ times, and $b_u \in \mathbb{R}^{\mathbf{e}^\intercal u}$ is a corresponding response vector. But we consider a situation where each v_ℓ corresponds to a costly experiment, which should be carried out between l_ℓ and u_ℓ times. Overall, we have a budget to carry out a total of $s(\geq m)$ experiments, and so we model the choices by x (in D-Opt). For a given feasible solution $\tilde{x}$, we define $A_{\tilde{x}}$ to be a matrix that has $v_\ell^\intercal$ repeated $\tilde{x}_\ell$ times as its rows, with $b_{\tilde{x}}$ as the associated response vector. This leads to the (reduced) least-squares problem $\min_{\theta\in\mathbb{R}^m} \|A_{\tilde{x}}\theta - b_{\tilde{x}}\|_2$, with solution $\hat{\theta} := (A_{\tilde{x}}^\intercal A_{\tilde{x}})^{-1} A_{\tilde{x}}^\intercal b_{\tilde{x}}$. The squared volume of a standard ellipsoidal confidence region (centered at $\hat{\theta}$) for the true θ is inversely proportional to $\det \sum_{\ell\in N} \tilde{x}_\ell v_\ell v_\ell^\intercal$. So D-Opt corresponds to picking the set of allowable experiments to minimize the volume of the confidence region for θ.

The design criterion of D-optimality was first suggested by A. Wald (see [27]). The term "D-optimality" was coined by J. Kiefer (see [11]). Also see [12, 26, 8], for example.

There is a large literature on heuristic algorithms for D-Opt and its variations. See [25] and [20] for approximation algorithms with guarantees. [28] was the first to approach D-Opt with an exact branch-and-bound (B&B) algorithm, employing a bound based on Hadamard's inequality and another based on continuous relaxation (apparently without using state-of-the art NLP solvers of that time). [14, 13] proposed a spectral bound and analytically compared it with the Hadamard bound; also see [17]. [18] applied a local-search procedure and an exact algorithm to the Data-Fusion problem, a particular case of the D-optimality problem where the positive definite $\sum_{\ell\in N} l_\ell v_\ell v_\ell^\intercal$ is known as the "existing Fisher Information Matrix (FIM)". Moreover, the Data-Fusion problem consider only the case where the variables are binary, i.e., $l = 0$ and $u = \mathbf{e}$. Although the Data-Fusion and the D-optimality problems have similarities, most techniques used in [18] rely on the positive definiteness of the existing FIM and cannot be applied to our problem.

Next, we highlight our contributions:

- a new upper bound for the binary D-optimality problem, which we call the Γ-bound. We prove that the Γ-bound is a generalization for the binary D-optimality problem of the "M-DDF-complementary bound" for the Data-Fusion problem presented in [18]. We note that we can reformulate D-Opt as a binary D-optimality problem by repeating row ℓ of A $u_\ell - l_\ell$ times;
- numerical and theoretical results showing some relations between different bounds for binary D-optimality problem and the Data-Fusion problem, including the Γ-bound and other bounds from the literature, specifically the so-called spectral bound, Hadamard bound, "M-DDF bound", "M-DDF-complementary bound", and natural bound;
- three local-search heuristics for D-Opt considering different neighborhoods of the current point to visit at each iteration;
- five algorithms to construct initial solutions for the local-search procedures;

- how we compute the determinant of a rank-one update of a given matrix, knowing the determinant of the matrix. This procedure is essential to the successful application of the local-search procedures;
- different strategies to compute the direction to move at each iteration of the local-search procedures;
- variable-bound tightening (VBT) inequalities, which are constructed based on a lower bound for D-Opt and on the knowledge of a feasible solution for the Lagrangian dual of its continuous relaxation;
- a B&B algorithm based on a convex mixed-integer nonlinear programming formulation of D-Opt. We investigate possible methodologies to accelerate the convergence of the B&B algorithm, by combining the use of the VBT inequalities, local-search procedures, and the use of different upper bounds;

We note that although [28], in 1982, already considered the application of a B&B algorithm for D-optimality, that work did not have access to our new Γ-bound, and did not use variable tightening based on convex optimization, nor discussed the linear algebra of doing a fast local search.

A similar solution approach has been successfully applied to the related maximum-entropy sampling problem (MESP) (see [3, 1, 2, 6]), where given the covariance matrix C of a Gaussian random n-vector, one searches for a subset of s random variables that maximizes the "information" (measured by Shannon's "differential entropy") (see [24, 4, 16, 6], for example).

References

[1] Kurt M. Anstreicher. Maximum-entropy sampling and the Boolean quadric polytope. *Journal of Global Optimization*, 72(4):603–618, 2018.

[2] Kurt M. Anstreicher. Efficient solution of maximum-entropy sampling problems. *Operations Research*, 68(6):1826–1835, 2020.

[3] Kurt M. Anstreicher, Marcia Fampa, Jon Lee, and Joy Williams. Using continuous nonlinear relaxations to solve constrained maximum-entropy sampling problems. *Mathematical Programming*, 85:221–240, 1999.

[4] William F. Caselton and James V. Zidek. Optimal monitoring network design. *Statistics and Probability Letters*, 2:223–227, 1984.

[5] Zhongzhu Chen, Marcia Fampa, and Jon Lee. On computing with some convex relaxations for the maximum-entropy sampling problem. *INFORMS Journal on Computing*, 35(2):368–385, 2023.

[6] Marcia Fampa and Jon Lee. *Maximum-Entropy Sampling: Algorithms and Application*. Springer, 2022.

[7] Marcia Fampa, Jon Lee, Gabriel Ponte, and Luze Xu. Experimental analysis of local searches for sparse reflexive generalized inverses. *Journal of Global Optimization*, 81:1057–1093, 2021.

[8] Valerii V. Fedorov. *Theory of Optimal Experiments*. Academic Press, New York-London, 1972. Translated from the Russian and edited by W. J. Studden and E. M. Klimko.

[9] Gene H. Golub and Charles F. Van Loan. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[10] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, First edition, 1985.

[11] Jack Kiefer. On the nonrandomized optimality and randomized nonoptimality of symmetrical designs. *The Annals of Mathematical Statistics*, 29(3):675–699, 1958.

[12] Jack Kiefer and Jacob Wolfowitz. Optimum designs in regression problems. *The Annals of Mathematical Statistics*, 30(2):271–294, 1959.
[13] Chun-Wa Ko, Jon Lee, and Kevin Wayne. A spectral bound for D-optimality, 1994. Unpublished.
[14] Chun-Wa Ko, Jon Lee, and Kevin Wayne. Comparison of spectral and Hadamard bounds for D-optimality. In *MODA 5*, pages 21–29. Physica, Heidelberg, 1998.
[15] Ole Kröger, Carleton Coffrin, Hassan Hijazi, and Harsha Nagarajan. Juniper: An open-source nonlinear branch-and-bound solver in Julia. In *Proceedings of CPAIOR 2018*, pages 377–386. Springer, 2018.
[16] Jon Lee. Maximum entropy sampling. In A.H. El-Shaarawi and W.W. Piegorsch, editors, *Encyclopedia of Environmetrics, 2nd ed.*, pages 1570–1574. Wiley, Boston, 2012.
[17] Jon Lee and Joy Lind. Generalized maximum-entropy sampling. *INFOR: Information Systems and Operational Research*, 58(2):168–181, 2020.
[18] Yongchun Li, Marcia Fampa, Jon Lee, Feng Qiu, Weijun Xie, and Rui Yao. D-optimal data fusion: Exact and approximation algorithms, 2022. Preprint arXiv:2208.03589.
[19] Yongchun Li and Weijun Xie. Best principal submatrix selection for the maximum entropy sampling problem: Scalable algorithms and performance guarantees. *Operations Research*, 2023. `https://doi.org/10.1287/opre.2023.2488`.
[20] M Madan, M. Singh, U Tantipongpipat, and W. Xie. Combinatorial algorithms for optimal design. *Proceedings of Machine Learning Research (Conference on Learning Theory, COLT 2019)*, 99:2210–2258, 2019.
[21] Aleksandar Nikolov. Randomized rounding for the largest simplex problem. In *Proceedings of STOC 2015*, pages 861–870. ACM, 2015.
[22] Gabriel Ponte, Marcia Fampa, and Jon Lee. Exact and heuristic solution approaches for the D-optimality problem, 2022. In: Proceedings of the LIV Brazilian Symposium on Operations Research (SOBRAPO 2022), `https://proceedings.science/proceedings/100311/_papers/157447/download/fulltext_file2`.
[23] Friedrich Pukelsheim. *Optimal Design of Experiments*, volume 50 of *Classics in Applied Mathematics*. SIAM, 2006. Reprint of the 1993 original.
[24] Michael C. Shewry and Henry P. Wynn. Maximum entropy sampling. *Journal of Applied Statistics*, 46:165–170, 1987.
[25] Mohit Singh and Weijun Xie. Approximation algorithms for D-optimal design. *Mathematics of Operations Research*, 45(4):1512–1534, 2020.
[26] Ralph C. St. John and Norman R. Draper. D-optimality for regression designs: A review. *Technometrics*, 17(1):15–23, 1975.
[27] Abraham Wald. On the efficient design of statistical investigations. *Annals of Mathematical Statistics*, 14:134–140, 1943.
[28] William J. Welch. Branch-and-bound search for experimental designs based on D-optimality and other criteria. *Technometrics*, 24(1):41–48, 1982.
[29] Fuzhen Zhang, editor. *The Schur Complement and its Applications*, volume 4 of *Numerical Methods and Algorithms*. Springer-Verlag, New York, 2005.

Lasserre hierarchy for mixed-integer nonlinear optimization

Matteo Tacchi

Mixed-Integer NonLinear Programming (MINLP) involves optimization problems with nonconvex integer constraints. A common strategy to handle such problems consists in relaxing them into convex problems, that are as close as possible to the original one. Among such convex relaxation techniques, one can cite the class of lift-and-project methods, which are based on two steps: (1) a lifting operation that recasts the nonconvex problem at hand under the form of a convex cone

programming problem in a higher dimensional space and (2) a projection of the resulting optimal solutions back onto the decision space of the original problem.

The integer programming literature contains various examples of lift-and-project methods such as the Lovasz-Schrijver and Sherali-Adams hierarchies, as well as Lasserre's moment-SoS hierarchy [4, 6]. The idea of leveraging convex or linear structures by adding degrees of freedom to a problem is common to various techniques such as support vector machines, kernel methods or Koopman operators. In this talk, we intend to give a general overview on Lasserre's hierarchy for polynomial optimization, and to interpret it as a two-levels lift-and-project framework.

Consider the polynomial mixed 0–1 program

$$f^\star := \min\{f(x) \mid x = (\hat{x}, \bar{x}) \in \mathbb{R}^{\hat{n}} \times \{0,1\}^{\bar{n}} \wedge\ g(x) \geq 0\} \tag{1}$$

where f, g are polynomials (g can be vector-valued, then inequality constraints are interpreted component-wise), and such that $K := \{x \in \mathbb{R}^{\hat{n}} \times \{0,1\}^{\bar{n}} \mid g(x) \geq 0\}$ is compact. Noticing that $\bar{x} \in \{0,1\}^{\bar{n}}$ is equivalent to the set of polynomial inequality constraints $\forall i \in [\bar{n}]$, $\bar{x}_i^2(1-\bar{x}_i)^2 \leq 0$, K can be described as a semialgebraic set $K = \{x \in \mathbb{R}^n \mid h(x) \geq 0\}$ (with appropriate polynomial h and $n = \hat{n} + \bar{n}$) and (1) can be recast under the general form of a Polynomial Optimization Problem (POP):

$$f^\star = \min\{f(x) \mid x \in K\} = \min\{f(x) \mid x \in \mathbb{R}^n \wedge\ h(x) \geq 0\}. \tag{2}$$

Then, the moment-SoS hierarchy consists of three steps. The **first step** is an *infinite dimensional* lifting: the finite dimensional variable x is recast into a random variable X of law $\mu \in \mathcal{M}(K)_+$ ($\mathcal{M}(K)_+$ denotes the cone of Radon measures supported on K, dual to the cone $C(K)_+$ of nonnegative continuous functions):

$$f_\infty^\star := \min\left\{\mathbb{E}_\mu[f(X)] = \int f\, d\mu \;\middle|\; \mu \in \mathcal{M}(K)_+ \wedge \int 1\, d\mu = 1\right\}. \tag{3}$$

For any random variable X, $\mathbb{E}_\mu[f(X)] \geq f^\star$, hence $f_\infty^\star \geq f^\star$. Then, if $f^\star = f(x^\star)$, setting $\mathbb{P}_\mu(X = x^\star) = 1$ yields $f_\infty^\star \leq \mathbb{E}_\mu[f(X)] = f(x^\star) = f^\star$, hence $f_\infty^\star = f^\star$.

Then, problem (3) is equivalent to problem (2); however, while (2) is a nonconvex finite dimensional optimization problem, (3) is an infinite dimensional linear program. Another linear programming formulation can be derived from (2), using the definition of the minimum as the largest possible lower bound:

$$f'_\infty := \max\left\{\gamma \in \mathbb{R} \;\middle|\; \forall x \in K,\ \gamma \leq f(x)\right\} = \max\left\{\gamma \in \mathbb{R} \;\middle|\; f - \gamma \in C(K)_+\right\}, \tag{4}$$

which is in lagrangian duality with (3). Then, weak duality holds: $f'_\infty \leq f_\infty^\star$; moreover, here $\gamma = f^\star$ is feasible for (4), so that strong duality holds: $f'_\infty = f_\infty^\star$.

Problems (3) and (4) are actually instances of the Generalized Moment Problem (GMP) and its dual:

$$v^\star := \inf\left\{\int f\, d\mu \;\middle|\; \mu \in \mathcal{M}(K)_+ \wedge\ A\,\mu = b \in \Gamma'\right\} \tag{5a}$$

$$v' := \sup\left\{\langle \gamma, b\rangle_\Gamma \;\middle|\; f - A'\,\gamma \in C(K)_+ \wedge\ \gamma \in \Gamma\right\} \tag{5b}$$

where Γ is an appropriate normed vector space of polynomials (recall that $\mathbb{R}$ is the space of degree 0 polynomials) with topological dual Γ' and duality $\langle\cdot,\cdot\rangle_\Gamma$, $A : \mathcal{M}(K) \longrightarrow \Gamma'$ is linear and bounded for the operator norms of $\mathcal{M}(K)$ and Γ' and A' is the adjoint of A such that $\forall \mu \in \mathcal{M}(K)_+$, $\gamma \in \Gamma$, $\langle \gamma, A\,\mu\rangle_\Gamma = \int A'\,\gamma \; d\mu$. Notice that problems (5a) and (5b) are in duality through the Lagrange functional

$$L(\mu,\gamma) = \int f \; d\mu + \langle \gamma, b - A\mu\rangle = \langle \gamma, b\rangle_\Gamma + \int (f - A'\,\gamma) \; d\mu. \tag{6}$$

This GMP natuarlly extends to multiple Radon measures $\mu_1, \ldots, \mu_M$ on various euclidean spaces. In general, only weak duality is guaranteed: $v' \leq v^\star$, the question of strong duality often being nontrivial.

The **second step** of the moment-sum-of-squares (moment-SoS) hierarchy consists in solving *finite dimensional projections* of (3)–(4), using Putinar's theorem.

Theorem 1 ([9])**.** *Consider the space $\mathbb{R}[x]$ of polynomials in n variables and the convex cone $\Sigma := \{p_1^2 + \ldots + p_N^2 \mid N \in \mathbb{N},\ p_1, \ldots p_N \in \mathbb{R}[x]\}$ of SoS polynomials. Let $h = (h_1, \ldots, h_m) \in \mathbb{R}[x]^m$ and consider the quadratic module $Q(h) := \{\sigma_0 + \sigma_1\, h_1 + \ldots + \sigma_m\, h_m \mid \sigma_0, \ldots, \sigma_m \in \Sigma\}$ and the semialgebraic set $K = \{x \in \mathbb{R}^n \mid h(x) \geq 0\}$.*

Suppose that $Q(h)$ is archimedean, that is $\exists R > 0$; $R^2 - x^\top x \in Q(h)$. Let $y \in \mathbb{R}[x]'$ (continuous linear form). The following statements are equivalent:

(1) *y is a moment operator* i.e. $\exists \mu \in \mathcal{M}(K)_+$; $\forall p \in \mathbb{R}[x]$, $\langle p, y\rangle_{\mathbb{R}[x]} = \int p \; d\mu$

(2) *y is in the dual cone $Q(h)'$ of $Q(h)$* i.e. $\forall q \in Q(h)$, $\langle q, y\rangle_{\mathbb{R}[x]} \geq 0$

Notice that any $q \in Q(h)$ is nonnegative on K, so that $Q(h)$ being archimedean implies compactness of K. This theorem has a dual formulation:

Theorem 2 (Positivstellensatz [9])**.** *Using the same notation as in* **Theorem 1***:*

if $p \in \mathbb{R}[x]$ is positive on K, then $p \in Q(h)$.

Proof. If $p \notin Q(h)$ then one can use the separation theorem to build a $y \in \mathbb{R}[x]'$ s.t. $\langle p, y\rangle_{\mathbb{R}[x]} < 0$ and $\forall q \in Q(h)$, $\langle q, y\rangle_{\mathbb{R}[x]} \geq 0$. From **Theorem 1**, $\exists \mu \in \mathcal{M}(K)_+$; $0 > \langle p, y\rangle_{\mathbb{R}[x]} = \int p \; d\mu \geq 0$ since p is positive on K, which is a contradiction. □

Notice that **Theorem 2** is very similar to Farkas' lemma[1]. Next, **Theorems 1** and **2** can be used to replace $C(K)_+$ with $Q(h)$ and $\mathcal{M}(K)_+$ with $Q(h)'$ in problems (3)–(5), without changing their oprimal values. The announced projection then consists in restricting the degrees of the σ_i in the description of $Q(h)$, defining the bounded degree quadratic module, for $k \in \mathbb{N}$:

$$Q(h)_k := \{\sigma_0 + \sigma_1\, h_1 + \ldots + \sigma_m\, h_m \in Q(h) \mid \forall i \in [m], \sigma_0, \sigma_i\, h_i \in \mathbb{R}[x]_{2k}\} \tag{7}$$

Where $\mathbb{R}[x]_d$ denotes the space of degree at most d polynomials. This yields

$$f_k^\star := \min \left\{ \langle f, y\rangle_{\mathbb{R}[x]_{2k}'} \;\middle|\; y \in Q(h)_k' \;\wedge\; \langle 1, y\rangle_{\mathbb{R}[x]} = 1 \right\} \tag{8a}$$

[1]case where p and h are linear; then the σ_i are nonnegative scalars and the result still holds when p vanishes on K.

$$f'_k := \max\left\{\gamma \in \mathbb{R} \;\middle|\; f - \gamma \in Q(h)_k\right\}. \tag{8b}$$

Problems (8a) and (8b) are equivalent to convex, finite dimensional SDP that are in duality (hence $\forall k, f'_k \leq f^\star_k$), and can be solved numerically using softwares such as SeDuMi or Mosek [5], generating sequences of optimal solutions $(y^\star_k)_k \in \mathbb{R}[x]'$. The power of **Theorem 1** (resp. **2**) is that it yields $\lim f^\star_k = f^\star_\infty$ (resp. $\lim f'_k = f'_\infty$).

In the POP setting, a **third step** complements this hierarchy: generically (but not always, see [8] for finite convergence conditions), there is a $k \in \mathbb{N}$ such that $f^\star_k = f'_k = f^\star$ (in 0–1 programming such k always exists with $k \leq n$, see e.g. [4, 6]). It is then possible to extract the actual minimizers of (2) from the optimal solution $y^\star_k$ of (8a) (see [2, 5]): this is the final *projection onto the original decision space.*

Lasserre's hierarchy usually comes with three questions: (1) does strong duality hold for the GMP and its relaxations? (2) Does $(y^\star_k)$ converge to an optimal solution $\mu^\star$ of (5a)? (3) What is the convergence rate of the relaxations? In the POP setting, strong duality was proved in [3], [5, Section 4.10.1] contains a proof of convergence of $(y^\star_k)_k$ and the latest convergence rates for $(f^\star_k)_k$ can be found in [1]. In the general case, sufficient conditions for positively answering (1) and (2) have been given in [10], and a general methodology is under study for computing convergence rates based on [1] and function approximation theorems.

References

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

[2] R.E. Curto and L.A. Fialkow. *The truncated complex K-moment problem.* Transactions of the American Mathematical Society **352** (352), 2825–2855.

[3] C. Josz and D. Henrion, *Strong duality in lasserre's hierarchy for polynomial optimization*, Optimization Letters **10** (2016), 3–10.

[4] J.B. Lasserre, An explicit exact SDP relaxation for nonlinear 0-1 programs, in *Proceedings of the 8thIPCO Conference*, Utrecht NL (2001), 293–303.

[5] J.B. Lasserre, *Moments, Polynomials and Their Applications*, Imperial College Press, 2010.

[6] M. Laurent, *A Comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre Relaxations for 0–1 Programming*, Mathematics of Operations Research **28**, 3 (2003), 470–496.

[7] M. Laurent, Sums of squares, moment matrices and optimization over polynomials, in *Emerging Applications of Algebraic Geometry*, Springer Verlag (2009), 155–270.

[8] J. Nie, *Optimality conditions and finite convergence of Lasserre's hierarchy*, Mathematical Programming Series A. (2014) **146**, 97–121.

[9] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana University Mathematics Journal **42**, 3 (1993), 969–984.

[10] M. Tacchi, *Convergence of Lasserre's hierarchy: the general case*, Optimization Letters **16**, 3 (2022), 1015–1033.

Recent advances in the relaxation approach to mixed-integer optimal control

Christian Kirches
(joint work with Felix Bestehorn, Paul Manns)

Mixed-integer nonlinear optimization problems constrained by ordinary differential equations (*mixed-integer optimal control*, short MIOCP) or by partial differential equations (*mixed-integer PDE-constrained optimization*, short MIPDECO) have attracted significant uptake over the past decade. This is true for research in solution and approximation theory as well as for applications in diverse areas such as automotive control, biochemistry, energy, process control, pharmaceutics, or security. In the most simple form, we consider the problem class

$$\text{(MIOCP)} \qquad \begin{aligned} \min_{y,v}\ & J(y(T)) + C(v) \\ \text{s.t}\ & E(y) = f(y(t), v(t)) && \text{a.e. } t \in [0,T] \\ & 0 \leq c(y(t), v(t)) && \text{a.e. } t \in [0,T] \\ & v(t) \in \{v^1, \ldots, v^M\} && \text{a.e. } t \in [0,T], \end{aligned}$$

wherein a state y depends on the control v via a differential equation specified by the operator E and the right hand side f. The goal is to minimize a cost functional J and a general switch cost term C subject to mixed state-control constraints c. The point of interest here is that the free variable v is a *distributed* and *discrete-valued* control function. When discretizing (MIOCP) using a direct or indirect approach, v becomes a *mesh-dependent* degree of freedom.

We review a variety of solution approaches to variants of (MIOCP). In a direct approach to optimal control, *direct transscription* of (MIOCP) to a finite-dimensional problem results in a non-convex mixed-integer nonlinear program (MINLP). Tailored strategies for node selection and branching [1] that exploit the "arrow of time" help global MINLP solvers to perform better on such instances than general-purpose MINLP strategies. Under additional convexity assumptions, *state elimination* techniques sometimes permit to entirely remove the dependent variable y, which results in a smaller convex MINLP to which outer approximation can be applied with great efficiency [2, 3]. If rigorous global solutions are not important, *mode insertion gradients* can be computed to gradually improve a reference control up to near stationarity. Once the sequence of mode choices for v has been fixed, *switching time optimization*, a continuous optimization technique, may be applied to refine the switching time grid [4]. Besides MINLP techniques, relaxation approaches to (MIOCP) have proven powerful to approximately solve real-world instances. [5] describes a *variable time transformation* approach, and [4, 6, 7] develop a *convexification and relaxation* approach. The latter decomposes

(MIOCP) into a continuous relaxation

$$
\begin{aligned}
&\min_{y,\alpha} J(y(T)) \\
&\text{s.t } E(y) = \textstyle\sum_{i=1}^{M} \alpha_i(t) f(y(t), v^i) && \text{a.e. } t \in [0,T] \\
&\quad 0 \le \alpha_i(t) c(y(t), v^i) && \text{a.e. } t \in [0,T] \\
&\quad \alpha(t) \in [0,1]^M,\ \textstyle\sum_{i=1}^{M} \alpha_i(t) = 1 && \text{a.e. } t \in [0,T],
\end{aligned} \tag{RC}
$$

and a *combinatorial integral approximation* (CIA) problem, cf. [9], to reconstruct $v(t) = \sum_{i=1}^{M} \omega_{j,i} v^i$ on $t \in (t_j, t_{j+1})$ from a discretized relaxed solution α^*. Given α^*, *sum-up rounding* algorithms [4, 7] produce a sequence (ω^N) of binary feasible controls on ever finer grids as $N \to \infty$. For these, tight approximation bounds are known, cf. [7], and induce weak-*-convergence in the space of controls v as well as norm convergence in the space of states y under suitable assumptions on the spaces, the differential operator E, and the function f, c.f. [4, 10] for MIOCPs. For semilinear hyperbolic MIPDECOS, this has been developed in [8]. Space-filling curves admit an extension of sum-up rounding procedures also to higher-dimensional domains, e.g. [11] for elliptic MIPDECOs.

Sum-up rounding controls provide converging but sub-optimal integer control reconstructions, and do not take the switch cost term $C(w)$ into account. Hence, optimal reconstruction procedures are reviewed next. One example is the mixed-integer linear program (MILP)

$$
\begin{aligned}
&\min_{\omega} C(\omega) \\
&\text{s.t } \max_{j \in [N-1]} \left| \textstyle\sum_{i \in [M]} (t_{j+1} - t_j)(\alpha^*_{j,i} - \omega_{j,i}) \right| \le \theta \ \forall\ i \in [M] \\
&\quad \omega \in \{0,1\}^{(N-1)\times M},\ \textstyle\sum_{i \in [M]} \omega_{j,i} = 1 && \forall\ j \in [N-1],
\end{aligned} \tag{MILP}
$$

which minimizes the switch cost subject to a maximum integral control deviation approximation bound θ. We review further reconstruction problems, such as optimizing the sum-up rounding gap and dwell-time restrictions [12], the penalty-ADM method [13] for coupling the effect of reconstruction decisions to the state space, rounding via matching [14] for sequence independent switch cost terms, and rounding via shortest paths [15] on exponential graphs for sequence dependent ones.

References

[1] M. Hahn, S. Leyffer, and V. Zavala. Mixed-Integer PDE-Constrained Optimal Control of Gas Networks. Mathematics and Computer Science Division, Argonne National Laboratory. Preprint ANL/MCS-P7095-0817, 2017.

[2] C. Buchheim, R. Kuhlmann, and C. Meyer. Combinatorial optimal control of semilinear elliptic PDEs, COAP, 70(3), pp. 641–675, 2018.

[3] A. Thünen, S. Leyffer, and S. Sager. State elimination for mixed-integer optimal control of partial differential equations by semigroup theory. Optimal Control Applications and Methods, 43(3), pp. 867–883, 2022.

[4] S. Sager. Numerical methods for mixed-integer optimal control. Dissertation, Heidelberg University, 2005.

[5] M. Gerdts. A variable time transformation method for mixed-integer optimal control problems. Optimal Control Applications and Methods, 27(3), pp. 169–182, 2006.
[6] S. Sager, H.G. Bock, and M. Diehl. The Integer Approximation Error in Mixed-Integer Optimal Control. Mathematical Programming, 133, pp. 1–23, 2012.
[7] C. Kirches, F. Lenders, and P. Manns. Approximation Properties and Tight Bounds for Constrained Mixed-Integer Optimal Control. SIAM Journal on Control and Optimization, 58 (3), pp. 1371–1402, February, 2020.
[8] F.M. Hante. Relaxation methods for hyperbolic PDE mixed-integer optimal control problems. Optimal Control Applications and Methods, 38(6), pp. 1103–1110, 2017.
[9] S. Sager, M. Jung, and C. Kirches. Combinatorial Integral Approximation. Mathematical Methods of Operations Research, 73, pp. 363–380, 2011.
[10] P. Manns and C. Kirches. Improved Regularity Assumptions for Partial Outer Convexification of Mixed-Integer PDE-Constrained Optimization Problems. ESAIM: Control, Optimisation and Calculus of Variations, 26, 2020.
[11] P. Manns and C. Kirches. Multi-dimensional Sum-Up Rounding for elliptic control systems. SIAM Journal on Numerical Analysis, 58 (6), pp. 3427–3447, 2020.
[12] S. Sager and C. Zeile. On Mixed-Integer Optimal Control with Constrained Total Variation of the Integer Control. Computational Optimization and Applications, 78, pp. 575–623, 2020.
[13] S. Göttlich, F.M. Hante, A. Potschka, and L. Schewe. Penalty alternating direction methods for mixed-integer optimal control with combinatorial constraints. Mathematical Programming, 188, pp. 599–619, 2021.
[14] F. Bestehorn and C. Kirches. Matching Algorithms and Complexity Results for Constrained Mixed-Integer Optimal Control with Switching Costs. Preprint, Optimization Online Preprint 8059, October 2020.
[15] F. Bestehorn, C. Hansknecht, C. Kirches, and P. Manns. Mixed-Integer Optimal Control Problems with switching costs: A shortest path approach. Mathematical Programming, 188, pp. 621–652, 2021.

Computing lower bounds for the maximum perimeter of small polygons

Andreas Potschka
(joint work with Bernd Mulansky)

A small polygon is defined as a convex polygon in the plane with diameter not greater than one. Remarkably, for over a century the problem of finding small n-gons with maximum perimeter has not been solved completely. In 1922, Karl Reinhardt [12] published a geometric construction based on Reuleaux polygons (polygons whose facets are bent outwards to form circular arcs of constant radius). The construction allowed Reinhardt to derive an upper bound of $2n \sin\left(\frac{\pi}{2n}\right)$ on the perimeter and to solve the problem for all $n \geq 3$ that possess an odd divisor $d \neq 1$. The regular n-gon belongs to the solutions if and only if n is odd. There are (often) multiple solutions, which take the form of equilateral n-gons whose vertices lie on the boundary of some regular Reuleaux polygon with an odd number of sides.

The case $n = 2^s$ is solved for $n = 4$ [14] and $n = 8$ [1] but remains an open problem for larger n. We present a novel mixed-integer nonlinear programming (MINLP) formulation of the problem based on zonogons (point-symmetric polygons) with $2n$ vertices. We argue that excellent local solutions to the MINLP

can be found by solving a subset-sum problem (SSP) and a numerical homotopy approach using arbitrary precision arithmetic.

We confirm a known example for $n = 16$ and present new examples for the cases $n = 32$ and $n = 64$, which improve existing numerical lower bounds.

1. Related work

A good point of entry is [10]. The solution for the case $n = 8$ can be found in [1] and a recent survey in [6]. The case $n \neq 2^s$ has been further investigated in [8, 9] leading to additional solutions. Asymptotic lower bounds on the maximal perimeter have been provided and continuously improved in [11, 5, 3, 4] to $\frac{\pi^9}{8n^8} + O(n^{-10})$. These bounds, however, are not sharp (e.g, for $n = 64$, the difference to the upper bound amounts to $\pi^9/(8 \cdot 64^8) \approx 1.3 \cdot 10^{-11}$, while we know of local solutions with a distance of $2.8 \cdot 10^{-23}$). In [4], Couenne [2] was used to solve the problem for $n \leq 32$ and a candidate for $n = 64$ with a difference of $3.56 \cdot 10^{-23}$ to the upper bound was provided, which is 26% worse than the best solution we could compute. Ideas for symmetrization and constructions with zonogons have already been considered in the Russian literature [7].

2. Zonogon Mixed-Integer Nonlinear Program

We apply a construction based on zonogons, which leads to the MINLP

$$
\begin{aligned}
&\max \sum_{k=1}^{n} 2 \sin \frac{\varphi_k - \varphi_{k-1}}{2} \\
&\text{s.t.} \sum_{k=1}^{n} c_k(\cos\varphi_k - \cos\varphi_{k-1}) = 0, \quad \sum_{k=1}^{n} c_k(\sin\varphi_k - \sin\varphi_{k-1}) = 0, \\
&\qquad 0 = \varphi_0 \leq \varphi_1 \leq \ldots \leq \varphi_n = \pi, \quad c_k \in \{\pm 1\}, k = 1, \ldots, n
\end{aligned} \tag{1}
$$

with $n - 1$ real variables, n "binary" variables, and two nonlinear inequality constraints. The integer vector c is called a *code*. It is possible to enumerate all codes for $n \leq 32$ using fast algorithms such as [13]. Computational evidence suggests that there is only one local maximum of φ for each fixed code c.

We compute excellent feasible points of (1) with a two-step strategy, which is based on the observation that the code c only enters in the nonlinear equality constraints and that Reinhardt's upper bound would be attained by equidistantly spaced φ_k: Assuming Mossinghoff's conjecture [10] that optimal polygons are axially symmetric, we can restrict our search in step 1 to symmetric codes via a (often difficult to solve) SSP, while keeping φ fixed at the supersymmetric configuration. In step 2, corresponding angles φ can be found using an arbitrary precision (inexact) Newton method in a homotopy approach (1).

3. Numerical results

Using 100 digits of accuracy and after symmetry reduction, we obtain the results presented in Fig. 1 and Tab. 1.

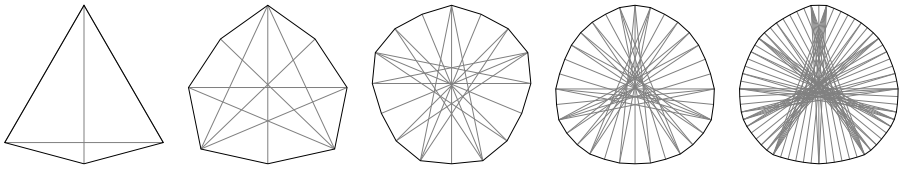

Figure 1. Computed polygons with large perimeter for $n = 2^s$, $s = 2, \ldots, 6$.

n	distance to u.b.	perimeter
4	$2.619 \cdot 10^{-2}$	3.035276180410083049395595350496193313396275605279722055256012829260227898995207987689471897 8
8	$2.980 \cdot 10^{-4}$	3.121147134059831353864659503638086530909542166469760122452478912381640349042889495925235033 6
16	$7.741 \cdot 10^{-7}$	3.136547716486607386085967031941228227298136765809232692789218203577745755473817628905857361 5
32	$1.335 \cdot 10^{-13}$	3.140331156954619365825401380577458672312053098339521869910414855946883777463454396416438368 5
64	$2.836 \cdot 10^{-23}$	3.141277250932772868061991415502468297956262096308096411175077343971836218350978865731726760 3

Table 1. Perimeters and their corresponding distance to the upper bound $2n \sin \frac{\pi}{2n}$ of the polygons in Fig. 1.

References

[1] C. Audet, P. Hansen, and F. Messine. The small octagon with longest perimeter. *J. Combin. Theory Ser. A*, 114(1):135–150, 2007.

[2] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wächter. Branching and bounds tightening techniques for non-convex MINLP. *Optim. Methods Softw.*, 24(4-5):597–634, 2009.

[3] C. Bingane. Tight bounds on the maximal perimeter and the maximal width of convex small polygons. *J. Global Optim.*, 84(4):1033–1051, 2022.

[4] C. Bingane. Maximal perimeter and maximal width of a convex small polygon, arXiv:2106.11831, 2023.

[5] C. Bingane and C. Audet. Tight bounds on the maximal perimeter of convex equilateral small polygons. *Arch. Math. (Basel)*, 119(3):325–336, 2022.

[6] G. Fejes Tóth. Finite variations on the isoperimetric problem. In L. Fejes Tóth, G. Fejes Tóth, and W. Kuperberg, editors, *Lagerungen: Arrangements in the Plane, on the Sphere, and in Space*, Grundlehren der mathematischen Wissenschaften, pages 243–247. Springer International Publishing, 2023.

[7] S. Gashkov. Inequalities for the area and perimeter of a convex polygon (russian). *Kvant*, 10:15–19, 1985.
[8] K.G. Hare and M.J. Mossinghoff. Sporadic Reinhardt polygons. 49(3):540–557, 2013.
[9] K.G. Hare and M.J. Mossinghoff. Most Reinhardt polygons are sporadic. *Geom. Dedicata*, 198:1–18, 2019.
[10] M.J. Mossinghoff. A $1 problem. *Amer. Math. Monthly*, 113(5):385–402, 2006.
[11] M.J. Mossinghoff. Isodiametric problems for polygons. *Discrete Comput. Geom.*, 36(2):363–379, 2006.
[12] K. Reinhardt. Extremale Polygone gegebenen Durchmessers. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 31:251–270, 1922.
[13] J. Sawada. Generating bracelets in constant amortized time. *SIAM Journal on Computing*, 31(1):259–268, 2001.
[14] N.K. Tamvakis. On the perimeter and the area of the convex polygon of a given diameter. *Bull. Greek Math. Soc*, 28:115–132, 1987.

The Robust Bilevel Selection Problem

Dorothee Henke

In bilevel optimization problems, two players, the leader and the follower, make their decisions in a hierarchy, and both decisions influence each other; see, e.g, [5] for an introduction to bilevel optimization. Usually one assumes that both players have full knowledge also of the other player's data. In a more realistic model, uncertainty can be quantified, e.g., using the robust optimization approach: Assume that the leader does not know the follower's objective function precisely, but only knows an uncertainty set of potential follower's objectives, and her aim is to optimize the worst case of the corresponding scenarios. We refer to the survey [1] for an overview of different uncertainty concepts in bilevel optimization. While bilevel optimization problems without uncertainty are often already NP-hard or even Σ_2^p-hard [6], now the question arises how the computational complexity of these problems changes under the additional complications of uncertainty.

We make a step towards answering this question by examining an easy bilevel problem. In the Bilevel Selection Problem, we are given finite leader's and follower's item sets $\mathcal{E}_l$ and $\mathcal{E}_f$, respectively, a number $b \in \{0, \dots, |\mathcal{E}_l \cup \mathcal{E}_f|\}$, and leader's and follower's item values $c\colon \mathcal{E}_l \cup \mathcal{E}_f \to \mathbb{Q}$ and $d\colon \mathcal{E}_f \to \mathbb{Q}$, respectively. The problem can then be written as follows:

$$\begin{aligned} \max_{X} \; & c(X \cup Y) \\ \text{s.t. } & X \subseteq \mathcal{E}_l \\ & Y \in \arg\max_{Y'} \; d(Y') \\ & \qquad \text{s.t. } Y' \subseteq \mathcal{E}_f \setminus X \\ & \qquad\quad |X \cup Y'| = b \end{aligned}$$

For the sake of simplicity, we assume that all follower's item values are distinct such that the optimum follower's solution Y is always unique. Observe that the follower's problem, for a fixed feasible leader's solution X, is a single-level Selection Problem that can be solved by a simple greedy approach, i.e., an optimum

solution is given by the best $b - |X|$ items from $\mathcal{E}_f \setminus X$ according to the follower's item values d. The leader's problem can be solved as follows: Enumerate all possible numbers b_l of items that the leader can select. For each b_l, greedily select b_l items from $\mathcal{E}_l$ according to the leader's item values c, and solve the corresponding follower's problem greedily as well. Return the solution that has the best overall leader's value. In case the sets $\mathcal{E}_l$ and $\mathcal{E}_f$ are disjoint, it is easy to see that this algorithm is correct. But also in the general case, this can be proven to be true. Hence, the Bilevel Selection Problem is a bilevel problem that is solvable in polynomial time.

We now investigate the complexity of its robust version. Here, we are given an uncertainty set $\mathcal{U}$ of possible follower's objectives $d\colon \mathcal{E}_f \to \mathbb{Q}$ and write the problem as follows:

$$\begin{aligned} \max_{X \subseteq \mathcal{E}_l} \ \min_{d \in \mathcal{U}} \ & c(X \cup Y) \\ & \text{s.t. } Y \in \arg\max_{Y'} d(Y') \\ & \qquad \text{s.t. } Y' \subseteq \mathcal{E}_f \setminus X \\ & \qquad\quad |X \cup Y'| = b \end{aligned}$$

The uncertainty can be seen as a third player, an adversary of the leader, who chooses a follower's objective $d \in \mathcal{U}$ after the leader has selected an item set X, but before the follower finally selects the additional items Y according to the values d chosen by the adversary. From the follower's perspective, the problem is still a single-level Selection Problem that can be solved greedily. We now study the complexity of the adversary's problem, which is the evaluation of the leader's objective function, and of the leader's problem for different types of uncertainty sets that are often used in robust optimization; see, e.g., [4]. Our focus here lies on discrete and interval uncertainty.

In case of a discrete uncertainty set $\mathcal{U}$, i.e., a finite set of scenarios that are explicitly given in the input, the adversary's problem can simply be solved by enumerating all scenarios and solving the follower's problem for each of them. This implies a polynomial-time algorithm also for the leader's problem in case of disjoint sets $\mathcal{E}_l$ and $\mathcal{E}_f$, similarly to the problem without uncertainty. For general sets $\mathcal{E}_l$ and $\mathcal{E}_f$ however, the robust problem with discrete uncertainty can be shown to be NP-hard. Hence, the uncertainty indeed increases the complexity of the underlying bilevel problem here.

If the uncertainty is given by an interval of possible values for each of the items, i.e., an interval uncertainty set $\mathcal{U} = \prod_{e \in \mathcal{E}_f} [d^-(e), d^+(e)]$ with $d^-, d^+\colon \mathcal{E}_f \to \mathbb{Q}$, then the adversary's problem can again be solved in polynomial time. However, more effort is required than in classical robust optimization with interval uncertainty where often each of the intervals can trivially be replaced by one of its endpoints [4]. Our algorithm makes use of the fact that the optimum follower's solutions in the different scenarios can be described as prefixes of an interval order derived from the uncertainty set and and is based on an idea from [7]. As above, this algorithm for the adversary's problem leads to a polynomial-time algorithm

also for the leader's problem in case of disjoint sets $\mathcal{E}_l$ and $\mathcal{E}_f$. In the general case, the complexity of the robust version of the Bilevel Selection Problem under interval uncertainty is still open.

Similar results were obtained for a robust bilevel continuous knapsack problem in [2]. The complexity of more general bilevel optimization problems under robust uncertainty regarding the follower's objective from the leader's point of view was studied in [3]. In particular, also interval uncertainty has been shown to increase the complexity of a bilevel optimization problem significantly in general.

References

[1] Y. Beck, I. Ljubić, and M. Schmidt, *A survey on bilevel optimization under uncertainty*, European Journal of Operational Research **311** (2023), 401–426.

[2] C. Buchheim and D. Henke, *The robust bilevel continuous knapsack problem with uncertain coefficients in the follower's objective*, Journal of Global Optimization **83** (2022), 803–824.

[3] C. Buchheim, D. Henke, and F. Hommelsheim, *On the complexity of robust bilevel optimization with uncertain follower's objective*, Operations Research Letters **49** (2021), 703–707.

[4] C. Buchheim and J. Kurtz, *Robust combinatorial optimization under convex and discrete cost uncertainty*, EURO Journal on Computational Optimization **6** (2018), 211–238.

[5] S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnykova, *Bilevel Programming Problems. Theory, Algorithms and Applications to Energy Networks*, Springer Berlin, Heidelberg (2015).

[6] R. G. Jeroslow, *The polynomial hierarchy and a simple model for competitive analysis*, Mathematical Programming **32** (1985), 146–164.

[7] G. J. Woeginger, *On the approximability of average completion time scheduling under precedence constraints*, Discrete Applied Mathematics **131** (2003), 237–252.

An exact projection-based algorithm for a class of bilevel MINLPs

Maximilian Merkert

(joint work with Galina Orlinskaya, Dieter Weninger)

Bilevel optimization problems model hierarchical decision-making of two agents, called *leader* and *follower*. Solving bilevel mixed-integer problems with lower-level integer variables to global optimality is extremely challenging as a single-level reformulation that is suitable for MINLP solvers is usually not available. Relying on a (hypothetical) exact single-level solver, we present an algorithmic framework that is capable of solving certain bilevel MINLPs to exact global optimality. Our method is an enhancement of an approximative projection-based algorithm by Yue, Gao, Zeng and You [1], which was designed for linear mixed-integer bilevel problems. We extend it to a problem class involving nonlinearities and show that one can get rid of an explicit ϵ approximation in the algorithm from [1] under one additional assumption (see Assumption (4) below).

Problem class and assumptions. We consider bilevel MINLPs with discrete and continuous variables on both levels, using the optimistic assumption and the

value-function reformulation

$$\begin{aligned}
\text{(BL-MINLP)} \qquad \max_{x^u,y^u,x^l,y^l} \quad & F(x^u,y^u,x^l,y^l) \\
\text{s.t.} \quad & G(x^u,y^u,x^l,y^l) \;\le\; 0 \\
& g(y^u,x^l,y^l) \;\le\; 0 \\
& f(x^u,y^u,x^l,y^l) \;\ge\; \theta\,(x^u,y^u) \\
& x^u \in \mathbb{R}_+^{m_R}, y^u \in \mathbb{Z}_+^{m_Z}, x^l \in \mathbb{R}_+^{n_R}, y^l \in \mathbb{Z}_+^{n_Z}.
\end{aligned}$$

x^u, y^u denote upper-level continuous and discrete variables, respectively, x^l, y^l similarly denote lower-level variables. Objective and constraint functions of the upper level are denoted by F and G, respectively, while the corresponding lower-case letters f and g are used for the lower-level problem. Moreover,

$$\theta\,(x^u,y^u) = \max_{x^l\in\mathbb{R}_+^{n_R}, y^l\in\mathbb{Z}^{n_Z}} \left\{ f(x^u,y^u,x^l,y^l) \colon g(y^u,x^l,y^l) \le 0 \right\}$$

denotes the optimal-value function. Dropping the value-function constraint from (BL-MINLP) yields the so-called *high point relaxation* (HPR).

We work with the following set of assumptions:

(1) **Boundedness, continuity** All variables have finite bounds in the HPR and follower problem, respectively. F, G, f and g are continuous.
(2) **Follower optimality conditions** For any fixed upper-level decisions $\bar{x}^u$ and $\bar{y}^u$, and lower-level integer decisions $\bar{y}^l$, the follower problem is convex and in case of feasibility satisfies Slater's condition. Functions f and g are continuously differentiable.
(3) **Solvability of master problems** We can solve the HPR of (BL-MINLP) together with the necessary and sufficient optimality conditions for the lower level as assumed in Assumption (2) to global optimality.
(4) **Restriction on continuous leader variables** Lower-level constraints do not contain any continuous upper-level variables.

Algorithmic concept. We follow the general iterative framework from [1].

- Start with the High Point Relaxation as the initial master problem.
- In each iteration:
  (1) Solve the master problem. Since it is a relaxation of (BL-MINLP), this yields a dual bound (upper bound for maximization problems).
  (2) For the obtained upper-level decisions $(\bar{x}^u, \bar{y}^u)$, solve two subproblems for first computing $\theta\,(\bar{x}^u, \bar{y}^u)$ and then checking whether there is $(\bar{x}^l, \bar{y}^l)$ that extends $(\bar{x}^u, \bar{y}^u)$ to a bilevel-feasible solution. If successful, this gives a new primal bound (lower bound).
  (3) Add conditions to the master problem that essentially make the leader anticipate any follower response involving the optimal $\bar{y}^l$ from the previous step.
- Terminate as soon as lower and upper bounds agree.

In step (3) of the loop we iteratively add constraints to the master problem that can be formally written as

$$\left[y^u \in \mathrm{Proj}_{(y^u)} P\left(\bar{y}^l\right)\right] \implies \left[f(x^u, y^u, x^l, y^l) \geq \theta\left(x^u, y^u, \bar{y}^l\right)\right], \tag{1}$$

where

$$P\left(\bar{y}^l\right) = \left\{\left(y^u, x^l\right) : g(y^u, x^l, \bar{y}^l) \leq 0\right\}$$

and thus

$$\mathrm{Proj}_{(y^u)} P\left(\bar{y}^l\right) = \left\{y^u : \exists x^l \text{ with } \left(y^u, x^l\right) \in P\left(\bar{y}^l\right)\right\}$$

denotes the set obtained by projecting out x^l from the lower-level feasible set. These constraints state that any master problem solution must leave the follower with an objective value at least as good as the best-possible had they played $\bar{y}^l$, if $\bar{y}^l$ is playable. Note that we extend the notation for the optimal-value function in a natural way to include fixing y^l to $\bar{y}^l$.

It can be shown that the procedure eventually terminates with an optimal solution under the above assumptions. Details can be found in [2]. However, constraints (1), which we call *projections implications*, are difficult to model.

An exact realization of the framework. Let Y^L denote the set of all $\bar{y}^l$ for which a projection implication shall exist. First, we use a binary variable ψ^k for every $y^{l,k} \in Y^L$ to split the projection implication into two:

$$\left[y^u \in \mathrm{Proj}_{(y^u)} P\left(y^{l,k}\right)\right] \implies \left[\psi^k = 1\right] \tag{2a}$$

$$\left[\psi^k = 1\right] \implies \left[f(x^u, y^u, x^l, y^l) \geq \theta\left(x^u, y^u, y^{l,k}\right)\right]. \tag{2b}$$

The first implication (2a) is equivalent to the disjunction

$$\left[y^u \notin \mathrm{Proj}_{(y^u)} P\left(y^{l,k}\right)\right] \vee \left[\psi^k = 1\right].$$

However, for linearly relaxed y^u the set on which this disjunction is true is not necessarily closed, causing modeling problems. In [1] this issue is dealt with by effectively considering a slightly larger set in place of $\mathrm{Proj}_{(y^u)} P\left(y^{l,k}\right)$. However, this comes at the cost of introducing an ϵ constant and that the result is only an approximation.

Our approach instead considers a slightly smaller set $U \subset \mathrm{Proj}_{(y^u)} P\left(y^{l,k}\right)$ for which the implication $\left[y^u \in U\right] \implies \left[\psi^k = 1\right]$ can be modeled in an exact way using duality theory due to Assumption 2. The remaining requirement

$$\left[y^u \in \left(\mathrm{Proj}_{(y^u)}\left(P\left(y^{l,k}\right)\right) \setminus U\right)\right] \implies \left[\psi^k = 1\right]$$

is modeled via no-good cuts, so U should be chosen as large as possible. For details on how to choose U and model the resulting implication in a MINLP, the reader is referred to [2].

Discussion and further considerations. The result shows that solvers for bilevel MINLPs of the considered type do not need to explicitly introduce sources of inexactness ("epsilons") besides the ones implicitly rooted in the single-level solver. It also proves that a bilevel optimum is attained for the considered class of problems. The main limiting assumptions are that the follower problem is convex in its continuous variables and that g does not depend on x^u. Although an exact MINLP solver assumed in Assumption (3) is purely hypothetical, the result is not only of theoretical relevance: Computational experiments (see [2]) on modified library instances show that the number of no-good cuts needed is usually moderate, and suggest that the method could form the basis for a practical solver implementation. It is worth noting that the framework implicitly makes use of all advanced features of MINLP solvers and will benefit from any future improvements. Our prototype implementation may of course also be improved by features tailored towards the specific framework. Besides specialized cutting planes, a candidate is *projection branching*: Instead of modeling the implications (2a) and (2b), which involves big-M constraints, we can branch on whether $y^u \in \mathrm{Proj}_{(y^u)} P\left(y^{l,k}\right)$ and add the corresponding constraints for either case. This avoids big-M constraints altogether. Details and possible further advantages, both theoretical and computational, are the topic of ongoing research.

This is an extended abstract on results published in [2].

References

[1] D. Yue, J. Gao, B. Zeng, F. You, *A projection-based reformulation and decomposition algorithm for global optimization of a class of mixed integer bilevel linear programs*, Journal of Global Optimization **73** (2019), 27–57. doi:10.1007/s10898-018-0679-1

[2] M. Merkert, G. Orlinskaya, D. Weninger, *An Exact Projection-Based Algorithm for Bilevel Mixed-Integer Problems with Nonlinearities*, Journal of Global Optimization **84** (2022), 607–650. doi:10.1007/s10898-022-01172-w

Extended Formulations for Binary Optimal Control Problems

Christoph Buchheim

(joint work with Alexandra Grütering, Maja Hügging, Christian Meyer)

Extended formulations are an important tool in polyhedral combinatorics. Many combinatorial optimization problems require an exponential number of inequalities when modelled as a linear program in the natural space of variables. However, by adding artificial variables, one can often find a compact linear formulation, i.e., one containing a polynomial number of variables and constraints, such that the projection to the original space of variables yields the exact linear formulation.

We propose to use the same approach for binary optimal control problems where the controls may be chosen from a set $D \subseteq BV(0, T, \{0, 1\})$. We assume that the

latter satisfies the following conditions:

(D1) D is a bounded set in $BV(0,T,\mathbb{R})$,

(D2) D is closed in $L^1(0,T,\mathbb{R})$.

Condition (D1) implies that there exists a common bound $\sigma_{\max} \in \mathbb{N}_0$ such that all controls $u \in D$ have at most $\sigma_{\max}$ switching points. In the following, we assume for sake of simplicity that the switch starts being switched off. More precisely, when the switch is on at $t = 0$, we already count this as one switching.

We now define an *extended formulation* of D as a set $D^{\text{ext}} \subseteq BV(0,T,\mathbb{R})^{d+1}$, for some $d \in \mathbb{N}_0$, such that the projection of D^{ext} to the first coordinate agrees with $\overline{\text{conv}}(D)$, where the closure is taken in $L^1(0,T,\mathbb{R})$. Moreover, we require that the formulation is linear and compact, i.e., contains a polynomial number of controls and a polynomial number of linear constraints. Finally, we require that the extended formulation is compatible with discretization.

On the positive side, motivated by a result in [2], we devise an extended formulation for the case where the only condition in D is that the switch changes at most $\sigma_{\max}$ times. For even $\sigma_{\max}$, the extended model is

$$\left\{u \in BV(0,T,[0,1]),\ z \in BV(0,T,[0,\tfrac{\sigma_{\max}}{2}])\colon Dz \geq Du,\ Dz \geq 0\right\} .$$

For odd $\sigma_{\max}$, a similar model can be derived. We also propose an extended formulation for the so-called dwell time constraints that is motivated by the extended formulation of Rajan and Takriti in the discrete case [3].

On the negative side, we show that such extended formulations cannot exist for general linear switching constraints, i.e., for sets $D(A,b)$ consisting of all controls having $\sigma_{\max}$ switching points $(t_1,\dots,t_{\sigma_{\max}})$ satisfying $At \leq b$; the latter condition generalizes the dwell-time constraints. We show this by proving that, in general, it is NP-complete to decide whether a discretization of $D(A,b)$ contains a feasible control or not. This is even true when the number of grid cells is not part of the input, meaning that the hardness of the problem does not disappear when the discretization grid is refined.

References

[1] C. Buchheim, A. Grütering, and C. Meyer, *Parabolic optimal control problems with combinatorial switching constraints – Part I: Convex relaxations*, arXiv preprint arXiv:2203.07121, (2022).

[2] C. Buchheim and M. Hügging, *The polytope of binary sequences with bounded variation*, Discrete Optimization. To appear.

[3] D. Rajan and S. Takriti, *Minimum up/down polytopes of the unit commitment problem with start-up costs*, Tech. Rep. RC23628, IBM Research Report, 2005.

Modelling Guide for Mixed-Integer Semidefinite Programming

Renata Sotirov
(joint work with Frank de Meijer)

Mixed-integer semidefinite programming can be viewed as a generalization of mixed-integer programming where the vector of variables is replaced by mixed-integer positive semidefinite matrix variables.

In this talk we show that many optimization problems may be modeled as mixed-integer semidefinite programs (MISDPs), either by a generic approach for certain large problem classes, or by a more problem-specific approach. Based on a comprehensive study on discrete positive semidefinite matrices, we introduce a generic approach to derive MISDP for quadratically constrained quadratic programs and quadratic matrix programs. We also provide a combinatorial, polyhedral, set-completely positive and integer hull description of the set of positive semidefinite binary matrices bounded by a certain rank. By applying a problem-specific approach, we derive a compact MISDP of the quadratic assignment problem (QAP). To the best of our knowledge, our MISDP-QAP formulation provides the most compact convex mixed-integer formulation of the problem in the literature. Complementary to the recent advances on algorithmic aspects related to MISDP, our work opens new perspectives on solution approaches for the here considered problems.

In the second part of the talk we present a formulation of the elementary closure of spectrahedra that relies on the data matrices of the integer semidefinte program (ISDP) and positive semidefinite matrices. Our formulation provides a constructive description of the elementary closure of spectrahedra rather than the implicit description that is known for general convex sets. Equivalent to the case of polyhedra, the elementary closure operation can be repeated, leading to a hierarchy of stronger approximations of the integer hull of the spectrahedron. Our explicit formulation of the elementary closure of spectrahedra enables us to introduce Chvátal-Gomory (CG) cuts for integer semidefinite programs. The CG cuts are introduced by Chvátal [1] and Gomory [2] and it is considered to be among the most celebrated results in integer programming.

In the third part of the talk, we first show how to derive ISDP for the the quadratic traveling salesman problem (QTSP) by exploiting the algebraic connectivity of the directed Hamiltonian cycle. The QTSP is the problem of finding a Hamiltonian cycle in a graph that minimizes the total interaction costs among consecutive arcs. Then, we show that the CG cuts resulting from these formulations contain several well-known families of cutting planes. Numerical results verify the practical strength of the CG cuts in our branch-and-cut algorithm, which outperforms alternative ISDP solvers and is able to solve large QTSP instances to optimality.

For the details on the first part of the talk see [3], and for the details on the remaining part of the talk see [4].

References

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

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

[3] F. de Meijer and R. Sotirov, *On integrality in semidefinite programming for discrete optimization*, Preprint 2023, https://arxiv.org/abs/2306.09865.

[4] F. de Meijer and R. Sotirov, *The Chvátal-Gomory procedure for integer SDPs with applications in combinatorial optimization*, Preprint 2022, https://arxiv.org/abs/2201.10224.

Strong cutting planes for convex mixed-integer programming

Jan Kronqvist

Outer approximation-based algorithms remain computationally more efficient than pure nonlinear branch and bound [1]. The main reasons behind this is the computational efficiency and robustness of LP solvers, i.e., we can explore nodes efficiently with LP relaxations, but also due to the mature technology of cutting planes in MILP. The combinatorial challenges are then to some extent passed on to the MIP subsolvers which employ a variety of cutting planes. However, the MIP subsolver will not be aware of the actual nonlinear constraints which can prevent the subsolver from utilizing this information to derive stronger cuts.

Deriving valid inequalities from convex nonlinear inequalities is trivial as any gradient cut (first-order Taylor Series expansion) results in a valid inequality. The simplest approach for generating an outer approximation of the feasible set is to solve a mixed-integer linear relaxation to obtain a trial solution $\mathbf{x}^k$, and generate the cuts

$$g_i(\mathbf{x}^k) + \nabla g_i(\mathbf{x}^k)^\top (\mathbf{x} - \mathbf{x}^k) \leq 0 \quad \forall i : g_i(\mathbf{x}^k) \geq 0,$$

to improve the mixed-integer linear relaxation. These cuts are often referred to as gradient cuts. Unfortunately, the gradient cuts can be quite weak. Unless $\mathbf{x}^k$ is on the boundary of the feasible set, the gradient cut may not even form a supporting hyperplane to the continuously relaxed feasible set. Several strategies for for obtaining stronger gradient cuts have been proposed, for example by projecting the trial solution $\mathbf{x}^k$ onto the continuously relaxed feasible set [2]. Keep in mind, that a supporting hyperplane to the continuous relaxation may still be a weak inequality in relation to the convex hull of the feasible set intersected with the integer lattice. The main shortcoming of classic outer approximation based methods is, thus, that nonlinearly and integrality are dealt with separately, leading to more iterations and unnecessarily big branch-and-bound trees.

The lift-and-project framework can also be applied to convex MINLP [3], but solving the resulting nonlinear cut generation problems is not computationally trivial. Furthermore, we illustrate by a simple example that one may want to consider multi-branch split disjuncts in order to get a stronger cut, which results in larger and computationally more expensive cut generation problems. In the talk, we discuss a simple, yet effective, cut-strengthening procedure that utilizes

disjunctive structures to derive cuts that consider both nonlinearity and integrality to generate a stronger cut. This disjunctive cut strengthening technique was presented by Kronqvist and Misener [4], and assumes the problem contains at least one constraint of the form

$$\sum_{i \in \mathcal{I}_D} x_i = 1,$$

where $\mathcal{I}_D$ contains indices of binary variables. Then starting from a valid cut $\alpha^\top \mathbf{x} \leq \beta$, the technique produces a strengthened cut of the form

$$\alpha^\top \mathbf{x} \leq \sum_{i \in \mathcal{I}_D} b_i x_i,$$

where each $b_i \leq \beta$.

Compared to lift-and-project cuts, the cut-strengthening technique has the advantage that the cut is obtained by solving $|\mathcal{I}_D|$ independent subproblems instead of a single large cut-generating subproblem. The effectiveness of the cut-strengthening is clearly demonstrated by a numerical study.

References

[1] Kronqvist, J., Bernal, D.E., Lundell, A. and Grossmann, I.E., *A review and comparison of solvers for convex MINLP*, Optimization and Engineering, **20** (2019), 397–455.

[2] Kronqvist, J., Lundell, A. and Westerlund, T., *The extended supporting hyperplane algorithm for convex mixed-integer nonlinear programming*, Journal of Global Optimization, **64**, (2016), 249–272.

[3] Bonami, P., *Lift-and-project cuts for mixed integer convex programs* In Integer Programming and Combinatoral Optimization: 15th International Conference, IPCO 2011, New York, NY, USA, Proceedings **15** (2011) 52–64.

[4] Kronqvist, J. and Misener, R., *A disjunctive cut strengthening technique for convex MINLP*, Optimization and Engineering, **22(3)** (2021), 1315–1345.

Time-Domain Decomposition for Mixed-Integer Optimal Control Problems

Falk M. Hante

(joint work with Richard Krug, Martin Schmidt)

We consider optimal control problems of the form

$$\begin{aligned}
&\min && \varphi_0(x(t_0)) + \varphi_{\mathrm{f}}(x(t_{\mathrm{f}})) && \text{(1a)}\\
&\text{s.t.} && \dot{x}(t) = f(x(t), u(t)) \quad \text{a. e. in } [t_0, t_{\mathrm{f}}], && \text{(1b)}\\
& && \chi_j(x(t_0)) = 0, \quad j = 1, \dots, p, && \text{(1c)}\\
& && \psi_j(x(t_{\mathrm{f}})) = 0, \quad j = 1, \dots, q, && \text{(1d)}\\
& && u(t) \in U \quad \text{a. e. in } [t_0, t_{\mathrm{f}}], && \text{(1e)}
\end{aligned}$$

where $[t_0, t_{\mathrm{f}}]$ is a fixed and finite time interval, $x\colon [t_0, t_{\mathrm{f}}] \to \mathbb{R}^n$ is a state function, and $u\colon [t_0, t_{\mathrm{f}}] \to \mathbb{R}^m$ is a control function. In Problem (1), $\chi_j, \psi_j : \mathbb{R}^n \to \mathbb{R}$ model constraints on the initial and terminal state, and $\varphi_0, \varphi_{\mathrm{f}} : \mathbb{R}^n \to \mathbb{R}$ model initial and terminal costs; $U \subset \mathbb{R}^m$ models control constraints, e.g. integrality

of certain components; and the minimum is taken over all absolutely continuous functions $x(\cdot)$ and over all measurable and essentially bounded functions $u(\cdot)$. Such problems appear for example in context of optimizing physical network flow with applications such as gas transportation [4].

In the case of (partially) discrete control sets U, direct transcription methods such as collocation or Runge–Kutta discretizations lead to mixed-integer non-linear programs (MINLPs). The limit behavior of such approximations for MIOCPs are discussed in [2], but typical step-sizing in such discretizations make these problems often large-scale and computationally intracktable.

We therefore consider computing Pontryagin-minima of (1), i.e., admissible state-control points (x, u), such that for any constant N there exists an $\varepsilon = \varepsilon(N) > 0$ with the properties that for any admissible point (x', u') satisfying

$$\|x - x'\|_{\mathcal{C}^0([t_0,t_\mathrm{f}];\mathbb{R}^n)} < \varepsilon, \quad \|u - u'\|_{\mathcal{L}^1(t_0,t_\mathrm{f};\mathbb{R}^m)} < \varepsilon, \quad \|u - u'\|_{\mathcal{L}^\infty(t_0,t_\mathrm{f};\mathbb{R}^m)} \leq N,$$

it holds $\varphi_0(x(t_0)) + \varphi_\mathrm{f}(x(t_\mathrm{f})) \leq \varphi_0(x'(t_0)) + \varphi_\mathrm{f}(x'(t_\mathrm{f}))$.

By Pontryagin's principle, and under certain regularity conditions but for arbitrary U, a Pontryagin-minimum satisfies

$$\begin{aligned}
\dot{x} &= f(x,u) \quad \text{a. e. in } (t_0,t_\mathrm{f}),\\
\chi_j(x(t_0)) &= 0, \quad j = 1,\dots,p,\\
\psi_j(x(t_\mathrm{f})) &= 0, \quad j = 1,\dots,q,\\
u(t) &\in U \quad \text{a. e. in } [t_0,t_\mathrm{f}],\\
\dot{\lambda} &= -\lambda^\top f_x(x,u) \quad \text{a. e. in } (t_0,t_\mathrm{f}),\\
\lambda(t_0) &= \varphi_0'(x(t_0)) + \textstyle\sum_{j=1}^{p} \beta_j \chi_j'(x(t_0)),\\
\lambda(t_\mathrm{f}) &= -\varphi_\mathrm{f}'(x(t_\mathrm{f})) - \textstyle\sum_{j=1}^{q} \beta_{p+j} \psi_j'(x(t_\mathrm{f})),\\
\max_{u\in U} H(\lambda(t),x(t),u) &= H(\lambda(t),x(t),u(t)) \quad \text{a. e. in } (t_0,t_\mathrm{f}).
\end{aligned} \tag{2}$$

The remarkable fact that the pointwise maximization is global motivates a MINLP approach for solving (2), but direct transcription alone would not bring any advantage over applying this to the original problem (1).

We therefore consider a non-overlapping and iteratively decoupled decomposition of (2) given by a sequence of so-called *virtual control problems* with fixed parameters $\gamma > 0$ and $\varepsilon \in (0,1)$, and subintervals (t_k, t_{k+1}), $k = 0,\dots,K$. For $k = 0$, the problem reads

$$\begin{aligned}
\min_{x_0,u_0} \quad & \varphi_0(x_0(t_0)) + \frac{1}{2\gamma}\|x_0(t_1) - \phi_{0,1}^{\ell-1}\|^2\\
\text{s.t.} \quad & \dot{x}_0 = f(x_0,u_0) \quad \text{a. e. in } (t_0,t_1),\\
& \chi_j(x_0(t_0)) = 0,\\
& u_0(t) \in U \quad \text{a. e. in } [t_0,t_1],
\end{aligned}$$

For $k = K$, the problem reads

$$\begin{aligned}\min_{x_K, u_K, h_K} \quad & \frac{1}{2\gamma}\|h_K\|^2 + \varphi_{\mathrm{f}}(x_K(t_{\mathrm{f}})) \\ \text{s.t.} \quad & \dot{x}_K = f(x_K, u_K) \quad \text{a. e. in } (t_K, t_{K+1}), \\ & x_K(t_K) = \phi^{\ell-1}_{K,K-1} + h_K, \\ & \psi_j(x_K(t_{\mathrm{f}})) = 0, \\ & h_K \in \mathbb{R}^n, \quad u_K(t) \in U \quad \text{a. e. in } [t_K, t_{K+1}].\end{aligned}$$

And for the inner time sub-intervals $k = 1, \dots, K-1$, the problems read

$$\begin{aligned}\min_{x_k, u_k, h_k} \quad & \frac{1}{2\gamma}\|h_k\|^2 + \frac{1}{2\gamma}\|x_k(t_{k+1}) - \phi^{\ell-1}_{k,k+1}\|^2 \\ \text{s.t.} \quad & \dot{x}_k = f(x_k, u_k) \quad \text{a. e. in } (t_k, t_{k+1}), \\ & x_k(t_k) = \phi^{\ell-1}_{k,k-1} + h_k, \\ & h_k \in \mathbb{R}^n, \quad u_k(t) \in U \quad \text{a. e. in } [t_k, t_{k+1}].\end{aligned}$$

The values $\phi^{\ell}_{k,k+1}$ and $\phi^{\ell}_{k,k-1}$ are updated using

$$\begin{aligned}\phi^{\ell-1}_{k,k+1} &= (1-\varepsilon)\left(x^{\ell-1}_{k+1}(t_{k+1}) + \gamma\lambda^{\ell-1}_{k+1}(t_{k+1})\right) + \varepsilon\left(x^{\ell-1}_{k}(t_{k+1}) + \gamma\lambda^{\ell-1}_{k}(t_{k+1})\right), \\ \phi^{\ell-1}_{k,k-1} &= (1-\varepsilon)\left(x^{\ell-1}_{k-1}(t_{k}) - \gamma\lambda^{\ell-1}_{k-1}(t_{k})\right) + \varepsilon\left(x^{\ell-1}_{k}(t_{k}) - \gamma\lambda^{\ell-1}_{k}(t_{k})\right).\end{aligned}$$

The Pontryagin optimality conditions of these subproblems can be seen to match exactly the time-decomposition of (2) if the iterates converge.

A direct transcription of these subproblems then yields a sequence of K but in dimension typically smaller MINLPs to be solved together with the update step above. For linear dynamics with quadratic costs and under certain technical conditions, we can prove convergence in error, i.e., the iterates $(x^{\ell}_k, \lambda^{\ell}_k)$ satisfy

$$\begin{aligned}x^{\ell}_k(t_{k+1}) - x^{\ell}_{k+1}(t_{k+1}) &\to 0, \\ \lambda^{\ell}_k(t_{k+1}) - \lambda^{\ell}_{k+1}(t_{k+1}) &\to 0\end{aligned}$$

for all $k = 0, \dots, K-1$ as $\ell \to \infty$.

As in [1], we demonstrate on numerical examples that this procedure can provide significant computational advantages compared to solving the direct discretization of the original problem on the entire time horizon. We also discuss the choice of additional parameters of the proposed algorithm. Typical convergence rates are shown in Figure 1.

Our computational experiments show that this advantage also pays off for the case of nonlinear problems if the procedure converges. For other than the linear-quadratic case however, a convergence theory or convergence enhancing mechanisms are still to be investigated.

These findings motivate to explore special mixed-integer nonlinear/quadratic programming technics to efficiently solve a sequence of similarly structured problems of low or moderate dimension. Moreover, they motivate to explore the convergence properties if additional constraints or more challenging dynamics are to

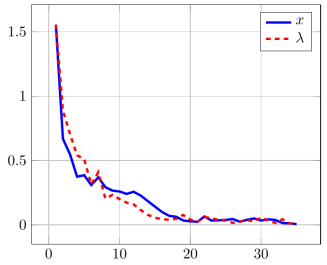

FIGURE 1. Convergence in error for a linear quadratic test problem from [1] over the iterates k.

be considered. Finally, it would also be desireable to replace technical assumptions such as the existence of optimal solutions to a less restictive setting based on convex relaxations as in [3] and the references given therein.

Acknowledgements. The research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under TRR154 (project ID: 239904186) and the Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

REFERENCES

[1] F.M. Hante, R. Krug, M. Schmidt: Time-Domain Decomposition for Mixed-Integer Optimal Control Problems. Applied Mathematics and Optimization **87**, Nr. 36 (2023). http://dx.doi.org/10.1007/s00245-022-09949-x.

[2] F.M. Hante, M. Schmidt: Convergence of Finite-Dimensional Approximations for Mixed-Integer Optimization with Differential Equations. Control and Cybernetics **48**, Nr. 2, 209–230 (2019).

[3] F.M. Hante: Relaxation methods for optimal switching control of PDE-dynamical systems. Handbook of Numerical Analysis, Numerical Control: Part B, Enrique Zuazua and E. Trelat (Eds.) **24** (2023), 61–76. http://dx.doi.org/10.1016/bs.hna.2022.10.004.

[4] F.M. Hante et al.: Challenges in optimal control problems for gas and fluid flow in networks of pipes and canals: From modeling to industrial applications. In: Industrial Mathematics and Complex Systems, P. Manchanda, R. Lozi, A.H. Siddiqi (Eds.), Springer Singapore, 77–122, (2017). http://dx.doi.org/10.1007/978-981-10-3758-0_5.

A Sequential Mixed-Integer Quadratic Programming Algorithm for Solving MINLP Arising in Optimal Control

Moritz Diehl

(joint work with Andrea Ghezzi, Sebastian Sager, Wim Van Roy)

The aim of the presented Sequential Mixed-Integer Quadratic Programming (S-MIQP) algorithm is to address nonconvex mixed-integer nonlinear programs (MINLP)

$$\begin{aligned} \min_{x \in X, y \in Y} \quad & f(x,y) \\ \text{s.t.} \quad & g(x,y) \leq 0, \\ & h(x,y) = 0, \end{aligned} \tag{1}$$

where $Y = \mathbb{Z}^{n_y} \cap \bar{Y}$, both $X \subset \mathbb{R}^{n_x}$ and $\bar{Y} \subset \mathbb{R}^{n_y}$ are convex polyhedral sets, and functions $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}$, $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_g}$, $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_h}$ are once continuously differentiable. When an MINLP arises from the discretization of a nonlinear mixed integer optimal control problem, the equality constraints $h(x,y) = 0$ are nonlinear and thus render the integer relaxation of the MINLP nonconvex. However, we formulate the following **optional convexity assumption**: *Function h is affine, functions f and g are convex on $X \times \bar{Y}$, and the integer set Y is finite.* While the S-MIQP algorithm is designed to work independently of this assumption, it is be guaranteed to find the global solution if accidentally applied to a MINLP that satisfies this assumption.

For notational compactness, we denote the first-order Taylor series of any nonlinear differentiable function $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_h}$ at a linearization point $(\bar{x}, \bar{y})$ by $h_{\mathrm{L}}(x, y; \bar{x}, \bar{y}) := h(\bar{x}, \bar{y}) + \frac{\partial h}{\partial x}(\bar{x}, \bar{y})(x - \bar{x}) + \frac{\partial h}{\partial y}(\bar{x}, \bar{y})(y - \bar{y})$. For any function $F : \mathbb{R}^{n_x} \to \mathbb{R}^{n_F}$ we define $\nabla F(x) := \frac{\partial F}{\partial x}(x)^\top$.

The above MINLP (1) can conceptionally be formulated as a pure integer optimization problem $\min_{y \in Y} J(y)$ by use of the *value function* $J(y)$ that is defined via the following nonlinear program (NLP):

$$\begin{aligned} J(y) \coloneqq \min_{x \in X} \quad & f(x,y) \\ \text{s.t.} \quad & g(x,y) \leq 0, \\ & h(x,y) = 0. \end{aligned} \tag{2}$$

The MIQP approximation of the MINLP will be based on a function

$$J_{\mathrm{QP}}(y; \bar{x}, \bar{y}, B) \approx J(y)$$

that depends on the chosen linearization point $(\bar{x}, \bar{y})$ and Hessian matrix $B \succeq 0$ and is defined by the following parametric quadratic program (QP):

$$\begin{aligned} J_{\mathrm{QP}}(y; \bar{x}, \bar{y}, B) \coloneqq \min_{x \in X} \quad & f_{\mathrm{L}}(x, y; \bar{x}, \bar{y}) + \frac{1}{2} \begin{pmatrix} x - \bar{x} \\ y - \bar{y} \end{pmatrix}^\top B \begin{pmatrix} x - \bar{x} \\ y - \bar{y} \end{pmatrix} \\ \text{s.t.} \quad & g_{\mathrm{L}}(x, y; \bar{x}, \bar{y}) \leq 0, \\ & h_{\mathrm{L}}(x, y; \bar{x}, \bar{y}) = 0. \end{aligned} \tag{3}$$

To focus on the main algorithmic idea, we make an optimistic **feasibility assumption**: *The NLP problem* (2) *and the QP problem* (3) *are feasible for any* $y, \bar{y} \in \bar{Y}, \bar{x} \in \bar{X}, B \succeq 0$, *and admit a minimizer with finite objective value.* This property can e.g. be achieved by the use of L_1 penalized slack variables for potentially infeasible constraints in the original problem formulation (1).

Note that, if $B = 0$ and under the convexity assumption, the piecewise linear convex function J_{QP} is an underestimator of J, i.e., $J_{\mathrm{QP}}(y; \bar{x}, \bar{y}, 0) \leq J(y)$ for all $y \in \bar{Y}$, as in outer approximation. For $B \succ 0$ no such guarantee exists, though we expect the quadratic approximation to often be better than the linear one. A favourable choice for least squares objectives $f(x, y) = \|F(x, y)\|_2^2$ is the Gauss-Newton Hessian $B_{\mathrm{GN}} = 2\nabla F(\bar{x}, \bar{y}) \nabla F(\bar{x}, \bar{y})^\top$ which makes the QP objective equal to $\|F_{\mathrm{L}}(x, y; \bar{x}, \bar{y})\|_2^2$.

First Version of the S-MIQP Algorithm. The S-MIQP algorithm iterates in the integer space starting with an initial guess $y_0 \in Y$. At iteration k, with a given integer candidate y_k, it first evaluates the NLP (2) giving both the value $J(y_k)$ and a subgradient $\nabla J(y_k)$ that can be computed as in Generalized Benders Decomposition [3]. It stores the current NLP solution tuple $(k, x_k, y_k, J(y_k), \nabla J(y_k))$ in a growing data structure $\mathcal{D}_k$. Now, after having evaluated the auxiliary NLP already at k integer points, giving $J(y_1), \ldots, J(y_k)$ along with their subgradients, we linearize the full problem at the best point found so far, the *incumbent solution* $y_{b(k)}$. Here, for uniqueness, $b(k) \leq k$ denotes the lowest index that achieves the minimal objective $J(y_{b(k)}) = \min_{i \in \{0,1,\ldots,k\}} J(y_i)$. Crucially, we impose additional linear level constraints that exclude all non-optimal points among the other visited points, y_i with $i \in \mathbb{I}_k := \{0, 1, \ldots, k\} \setminus \{b(k)\}$. Thus, with some choice of Hessian $B_k \succeq 0$, we solve the following MIQP in each iteration:

$$
\begin{aligned}
(x_k^{\mathrm{QP}}, y_k^{\mathrm{QP}}) := \quad & \underset{x \in X,\, y \in Y}{\arg\min} && f_{\mathrm{L}}(x, y; x_{b(k)}, y_{b(k)}) + \frac{1}{2} \begin{pmatrix} x - x_{b(k)} \\ y - y_{b(k)} \end{pmatrix}^\top B_k \begin{pmatrix} x - x_{b(k)} \\ y - y_{b(k)} \end{pmatrix} \\
& \text{s.t.} && g_{\mathrm{L}}(x, y; x_{b(k)}, y_{b(k)}) \leq 0, \\
& && h_{\mathrm{L}}(x, y; x_{b(k)}, y_{b(k)}) = 0, \\
& && J(y_i) + \nabla J(y_i)^\top (y - y_i) \leq J(y_{b(k)}),\ i \in \mathbb{I}_k.
\end{aligned} \tag{4}
$$

The level constraints form a polyhedral *Benders region* $\mathbb{B}_k := \{y \in \mathbb{R}^{n_y} | J(y_i) + \nabla J(y_i)^\top (y - y_i) \leq J(y_{b(k)}), i \in \mathbb{I}_k\}$ in integer space that excludes all points y for which we have an approximate certificate - that is exact in the convex case - that they are worse than the incumbent solution. Conversely, the Benders region does not exclude any point for which we do not have a non-optimality certificate in the convex case. This second property is preserved in the final version of the algorithm. Note that the Voronoi-based region used in a preliminary version of the sequential MIQP algorithm as proposed in [4] does not have this property and is inferior from both a theoretical and a practical perspective.

In some iterations, we also solve a mixed integer linear program (MILP) which helps us to generate a lower bound LB_k, to check if termination is possible and to potentially deliver a new trial point. The MILP is solved in two cases: either

if $J(y_k) = J(y_{b(k-1)})$, which can be detected before any MIQP solution, or if the MIQP solution y_k^{QP} equals one of the previously found points $y_0, \ldots, y_k$ (which can only be one of the previously optimal points due to the level constraints). The MILP is formulated as follows.

$$
\begin{aligned}
(\eta_k^{\mathrm{LB}}, x_k^{\mathrm{LB}}, y_k^{\mathrm{LB}}) := \underset{\eta\in\mathbb{R},\, x\in X,\, y\in Y}{\arg\min} \quad & \eta \\
\text{s.t.} \quad & \eta \geq f_{\mathrm{L}}(x, y; x_{b(k)}, y_{b(k)}), \\
& 0 \geq g_{\mathrm{L}}(x, y; x_{b(k)}, y_{b(k)}), \\
& 0 = h_{\mathrm{L}}(x, y; x_{b(k)}, y_{b(k)}), \\
& \eta \geq J(y_i) + \nabla J(y_i)^\top (y - y_i), \quad i \in \mathbb{I}_k.
\end{aligned} \tag{5}
$$

It is interesting to remark that this MILP is a mix of Outer Approximation [2] at the best point and Generalized Benders Decomposition for all other points, which ensures that linearized equality constraints appear only once. The MILP solution gives a new lower bound $\mathrm{LB}_k := \eta_k^{\mathrm{LB}}$, and the algorithm stops if the solution y_k^{LB} of the MILP equals one of the previously found optimal points, as in this case holds $J(y_{b(k)}) = \mathrm{LB}_k$. Otherwise, the algorithm continues by setting $y_{k+1} := y_k^{\mathrm{LB}}$ for the next candidate solution to be explored in the next iteration. In case that no MILP had to be solved, the algorithm sets $y_{k+1} := y_k^{\mathrm{QP}}$. One can show that the algorithm is well defined and stops after finitely many iterations at the globally optimal solution, if the convexity assumption holds.

Final S-MIQP Algorithm with Gradient Correction. In the general nonconvex case, the algorithm presented so far would not always be well defined: it can happen that the incumbent solution $y_{b(k)}$ is excluded from the feasible set of the MIQP if it violates, for some $i \in \mathbb{I}_k$, the level inequality $J(y_i) + \nabla J(y_i)^\top (y_{b(k)} - y_i) \leq J(y_{b(k)})$. To avoid this problem, we replace $\nabla J(y_i)$ by a *corrected gradient* $g_{(i,k)}^{\mathrm{corr}}$ that minimizes the weighted distance to $\nabla J(y_i)$ (with weighting matrix $W \succ 0$), within the set of *admissible gradients* $\mathbb{G}_{(i,k)} := \{g \in \mathbb{R}^{n_y} \mid J(y_i) + g^\top (y_{b(k)} - y_i) \leq J(y_{b(k)})\}$ as follows:

$$
g_{(i,k)}^{\mathrm{corr}} := \underset{g \,\in\, \mathbb{G}_{(i,k)}}{\arg\min} \quad \|g - \nabla J(y_i)\|_W^2. \tag{6}
$$

It is easy to see that, under the convexity assumption, in iteration k holds that $g_{(i,k)}^{\mathrm{corr}} = \nabla J(y_i)$ for all $i \in \mathbb{I}_k$. Thus, gradient correction does not impair the algorithm's capability to find the globally optimal solution in the convex case. However, in practice, the gradient correction alone might lead to very small regions $\mathbb{B}_k$ due to the minimality of the correction. In order to enlarge $\mathbb{B}_k$, we introduce a constant value $\rho \geq 1$ to amplify all gradients as $g_{(i,k,\rho)}^{\mathrm{ampl}} := \rho \cdot g_{(i,k)}^{\mathrm{corr}}$. One can show that gradient amplification can only increase the Benders region $\mathbb{B}_k$ in the MIQP and can only reduce the value of the lower bound in the MILP. Thus, the final nonconvex version of the S-MIQP algorithm replaces, in iteration k, the gradients $\nabla J(y_i)$ by $g_{(i,k,\rho)}^{\mathrm{ampl}}$ in all level inequalities, i.e., in both the MIQP and the MILP. Under the convexity assumption, for any fixed $\rho \geq 1$, one can show

that the algorithm will stop at a globally optimal solution. Without the convexity assumption, one can only show that the algorithm will never revisit a previously visited point, such that it stops after finitely many iterations if the integer set Y is finite.

Additional Practical Considerations. In practice, the generation of tight lower bounds can take many MILP solutions that do not improve the incumbent solution. Thus, one might want to stop much earlier than until the certificate $J(y_{b(k)}) = \mathrm{LB}_k$ is obtained, which is anyway not an exact certificate in the nonconvex case. In particular, in case of MINLP problems with least squares objectives that enable the use of a Gauss-Newton (GN) Hessian in the MIQP, a useful heuristic stopping criterion might be to stop when the MIQP solution equals the incumbent solution, i.e., when $y_k^{\mathrm{QP}} = y_{b(k)}$.

Another important topic is the choice of the initial guess $y_0 \in Y$. One option is to use the following two-step procedure. First, one uses an NLP solver to find the solution $(\bar{x}, \bar{y})$ of the relaxed MINLP. In compact notation, this is equivalent to setting $\bar{y} := \arg\min_{y \in \bar{Y}} J(y)$. Second, one chooses a Hessian B - ideally the GN-Hessian - and formulates and solves the MIQP (4) without level constraints. In compact notation, this amounts to setting $y_0 := \arg\min_{y \in Y} J_{\mathrm{QP}}(y; \bar{x}, \bar{y}, B)$. The integer solution found by this procedure, i.e., by solving a Gauss-Newton MIQP obtained by linearization at the relaxed solution, was proposed in [1] and found to be of very good heuristic value in challenging mixed integer optimal control problems arising from the control of thermal energy systems in buildings.

References

[1] A. Bürger, C. Zeile, A. Altmann-Dieses, S. Sager, and M. Diehl, *A Gauss–Newton-based decomposition algorithm for nonlinear mixed-integer optimal control problems*, Automatica, vol. 152, p. 110967, 2023.

[2] R. Fletcher and S. Leyffer, *Solving Mixed Integer Nonlinear Programs by Outer Approximation*, Mathematical Programming, vol. 66, pp. 327–349, 1994.

[3] A. Geoffrion, *Generalized Benders Decomposition*, Journal of Optimization Theory and Applications, vol. 10, pp. 237–260, 1972.

[4] A. Ghezzi, L. Simpson, A. Bürger, C. Zeile, S. Sager, and M. Diehl, *A Voronoi-based mixed-integer Gauss-Newton algorithm for MINLP arising in optimal control*, Proceedings of the European Control Conference (ECC), 2023.

A new lower bounding approach for the double row facility layout problem

Anja Fischer

(joint work with Frank Fischer, Angelika Wiegele)

In the Double Row Facility Layout Problem (DRFLP) one is given a set of n departments, their lengths $\ell_i, i \in [n] := \{1, \ldots, n\}$, (all widths are the same) and pairwise weights $w_{ij} = w_{ji}, i, j \in [n], i < j$, between them. One looks for a non-overlapping arrangement of the departments along two sides of a path such that the weighted sum of the pairwise center-to-center distances is minimized.

It has applications in factory planning, in the design of office buildings and in semiconductor design, see, e. g. [3]. The DRFLP is an extension of the well-studied Single Row Facility Layout Problem (SRFLP) where the input is the same, but all departments are arranged on exactly one side of a path. For the SRFLP medium-sized to large instances can be solved to optimality and one can derive good lower bounds for very large instances rather fast [1, 7]. The integer-programming as well as the semi-definite programming models for the SRFLP exploit the fact that, given non-negative weights, there always exists an optimal solution without spaces between neighboring departments and so one looks for a best permutation of these [1, 7]. In contrast to this, it seems to be much more challenging to derive high-quality lower bounds quickly or to even solve the DRFLP exactly. One reason for this is that an optimal DRFLP solution may contain free space between adjacent departments in the same row. So, in the solution of the DRFLP we do not only have to determine the row assignment of the departments, the orders of the departments in both rows but the precise positions of the departments in each row as well. In the literature, often big-M-type models are used to couple the row and order information with the positions, see, e. g. [2, 3, 8]. Unfortunately, the linear relaxations of these big-M-type models are rather weak. Apart from this, it is challenging to determine good DRFLP lower bounds because due to the double row structure the distance between two departments might be zero. In [4] some non-trivial lower bounds on the optimal DRFLP solution values are derived exploiting the relation of the DRFLP to certain scheduling problems. The best exact approach for the DRFLP in [5] is able to solve instances with up to 16 departments in less than 12 hours.

In this talk we present a new approach to combine the most successful models for the SRFLP with models to handle the local arrangements of adjacent departments including their free spaces. Indeed, we combine the well-known betweenness model by Amaral [1], where the variables are related to linearized products of ordering variables, with position variables and several further ones. The betweenness variables $x_{ijk} \in \{0,1\}, i,j,k \in [n], |\{i,j,k\}| = 3$, express here whether the center of department j lies between the centers of departments i and k. Apart from this we use two types of distance variables. The standard distance variables express the pairwise distances of the departments which are important for the calculation of the objective function. Indeed, we somehow underestimate the true distances by not taking free spaces into account. But the hope is that the betweenness model describes the global distance structure of the departments sufficiently well. One challenge in the calculation of the distance between two departments using betweenness variables is that there are usually departments in *both* rows that lie completely between two departments, but also few departments that are only partially in between. Unfortunately this might lead to large (relative) errors in the calculation of distances of departments which are close. In order to get good lower bounds on the local distances we set up small subproblems. These contain only a small number of the departments, indeed $D \subset [n]$ where in our test we use $|D| = \{3,4\}$. Using a big-M-type model we can calculate the exact distances of

the departments in D given a specific configuration. Then each of these subproblems provides a lower bound on certain local distances and the best lower bounds are then combined in a Lagrangian relaxation approach. Solving the Lagrangian relaxation leads to a non-smooth convex optimization problem. Deriving a solution value and an associated subgradient requires the solution of a large extended betweenness model and a possibly huge number of the small local subproblems. Since it is well-known that for larger instances even solving the linear relaxation of the betweenness model is challenging but gives strong bounds and exactly solving these problems is time-consuming and since we are mainly interested in deriving lower bounds, we neglect the integrality here and only solve the other subproblems exactly. As a solution method for the convex optimization problem we use the asynchronous parallel proximal bundle method presented in [6] which is summarized in the talk as well. Using this method has the advantage that we can fully exploit modern computer architectures. Indeed, in classic synchronous algorithms the slowest computation time of a subproblem determines the overall running time, which typically is the betweenness model in our case. In contrast, the new method can proceed without waiting for and knowing the solutions of all subproblems, reducing the slow-down due to synchronization significantly. At the end we present preliminary, but encouraging computational results.

References

[1] A. R. S. Amaral, *A new lower bound for the single row facility layout problem*, Discrete Applied Mathematics **157(1)** (2009), 183–190.

[2] A. R. S. Amaral, *A mixed-integer programming formulation for the double row layout of machines in manufacturing systems*, International Journal of Production Research **57(1)** (2019), 34–47.

[3] A. R. S. Amaral, *A mixed-integer programming formulation of the double row layout problem based on a linear extension of a partial order*, Optimization Letters **15(4)** (2021), 1407–1423.

[4] M. Dahlbeck, A. Fischer, F. Fischer, *Decorous combinatorial lower bounds for row layout problems*, European Journal of Operational Research **286(3)** (2020), 929–944.

[5] A. Fischer, F. Fischer, P. Hungerländer, *New exact approaches to row layout problems*, Mathematical Programming Computation **11(4)** (2019), 703–754.

[6] F. Fischer, *An Asynchronous Proximal Bundle Method*, Optimization Online (2022), `https://optimization-online.org/2022/03/8840/`

[7] P. Hungerländer, F. Rendl, *Semidefinite relaxations of ordering problems*, Mathematical Programming **140(1)** (2013), 77–97.

[8] L. D. Secchin, A. R. S. Amaral, *An improved mixed-integer programming model for the double row layout of facilities*, Optimization Letters **13(1)** (2019), 193–199.

Convex envelopes of bounded monomials on two-variable cones

Pietro Belotti

We consider a function $f : \mathbb{R}^n_+ \to \mathbb{R}_+$ defined as $f(\boldsymbol{x}) = \prod_{i\in N} x_i^{a_i}$ where $N = \{1, 2, \ldots, n\}$, $n > 1$, $a_i > 0 \forall i \in N$, a closed compact set $\mathcal{D} \subseteq \mathbb{R}^n_+$, and a scalar interval $[\ell, u]$, with $0 < \ell < u < +\infty$. We seek the convex hull of the set $F(\mathcal{D}) = \{(\boldsymbol{x}, z) \in (X \cap \mathcal{D}) \times \mathbb{R} : z = f(\boldsymbol{x})\}$, where

$$X = \{\boldsymbol{x} \in \mathbb{R}^n_+ : \ell \le f(\boldsymbol{x}) \le u\}.$$

If ℓ and u are infinite and $\mathcal{D}$ is a bounding box over $\boldsymbol{x}$, the convex envelope is polyhedral [9]. For example, if $n = 2$ we have $B = \{(x_1, x_2, z) \in \mathbb{R}^3_+ : z = x_1x_2, (x_1, x_2) \in [\lambda_1, \mu_1] \times [\lambda_2, \mu_2]\}$; then the following four inequalities [6]

$$\begin{array}{rcl@{\qquad}rcl} z &\ge& \lambda_2 x_1 + \lambda_1 x_2 - \lambda_2\lambda_1 & z &\ge& \mu_2 x_1 + \mu_1 x_2 - \mu_2\mu_1 \\ z &\le& \lambda_2 x_1 + \mu_1 x_2 - \lambda_2\mu_1 & z &\le& \mu_2 x_1 + \lambda_1 x_2 - \mu_2\lambda_1 \end{array} \tag{1}$$

form the convex hull of B [2]. An extension to the case $n = 3$ has also been studied [7]. The convex hull is, in general, not polyhedral when z has finite lower and/or upper bounds ℓ, u. In the bilinear case, i.e., $n = 2$, $a_1 = a_2 = 1$, the convex hull of

$$B' = \{(x_1, x_2, z) \in B : \ell \le z \le u\}$$

is tighter than (1) if the bounds on z are tighter than those on x_1 and x_2, i.e., $\ell > \lambda_1\lambda_2$ or $u < \mu_1\mu_2$. A family of infinitely many linear inequalities has been developed for this case [3], and it has been proven [1] that the convex hull of B' is the union of three sets, all second-order cone representable. The cases $a_1 = 1 \le a_2$ (convex hull) and $a_1, a_2 \ge 1$ (lower envelope) have also been studied [8].

The main focus of this abstract is the convex hull of $F(\mathcal{D})$ with

$$\mathcal{D} = W_{ij} := \{\boldsymbol{x} \in \mathbb{R}^n_+ : px_i \le x_j \le qx_i\}, \tag{2}$$

with $0 < p < q$, for two indices $i, j \in N$. The proofs can be found at [4].

Convex hull of $F(\mathbb{R}^n_+)$. Denote $\beta = \sum_{i\in N} a_i$, then for $z_0, \gamma \in \mathbb{R}$ define the cone

$$\mathcal{K} = \left\{(\boldsymbol{x}, z) \in \mathbb{R}^n_+ \times \mathbb{R} : (z - z_0)^\beta \le \gamma \textstyle\prod_{i\in N} x_i^{a_i}\right\}.$$

The vertex of $\mathcal{K}$ is $(\mathbf{0}, z_0)$. Also define $F(\mathcal{D})^{\le} := \{(\boldsymbol{x}, z) \in (X \cap \mathcal{D}) \times \mathbb{R} : z \le f(\boldsymbol{x})\}$. If $\beta = 1$ then $F(\mathcal{D})^{\le}$ is a convex cone intersected with $S := \{(\boldsymbol{x}, z) \in \mathbb{R}^n_+ \times \mathbb{R} : \ell \le z \le u\} = \mathbb{R}^n_+ \times [\ell, u]$. For a tight relaxation, parameters z_0 and γ must satisfy

$$\begin{array}{l} \{(\boldsymbol{x}, z) \in F(\mathcal{D})^{\le} : z = \ell\} = \{(\boldsymbol{x}, z) \in \mathcal{K} : z = \ell\}; \\ \{(\boldsymbol{x}, z) \in F(\mathcal{D})^{\le} : z = u\} = \{(\boldsymbol{x}, z) \in \mathcal{K} : z = u\}, \end{array}$$

and therefore $z_0 = \frac{u^{\frac{1}{\beta}}\ell - \ell^{\frac{1}{\beta}}u}{u^{\frac{1}{\beta}} - \ell^{\frac{1}{\beta}}}$ and $\gamma = \left(\frac{u-\ell}{u^{\frac{1}{\beta}} - \ell^{\frac{1}{\beta}}}\right)^\beta$. Note that $\beta \ge 1 \Leftrightarrow (z_0 \le 0, \gamma \ge 1)$. Also, if $\beta = 1$, $z_0 = 0$, $\gamma = 1$, and it is easy to verify that $\mathcal{K} \cap S \equiv F(\mathbb{R}^n_+)^{\le}$, while $\ell = 0$ implies $z_0 = 0$, i.e., the vertex of $\mathcal{K}$ is the origin.

Lemma 1. *$F(\mathbb{R}^n_+) \subseteq \mathcal{K}$ if and only if $\beta \ge 1$.*

The structure of both upper envelope and lower envelope of f over $X \cap \mathcal{D}$ discussed here changes radically at $\beta = 1$. For instance, $F(\mathbb{R}^n_+)^{\leq}$ is convex for $\beta \leq 1$ and nonconvex for $\beta > 1$.

Proposition 1. *If $\beta \geq 1$, then $conv(F(\mathbb{R}^n_+)) = \mathcal{K} \cap S$.*
If $\beta \leq 1$, then $conv(F(\mathbb{R}^n_+)) = F(\mathbb{R}^n_+)^{\leq}$.

Upper envelope over $X \cap W_{ij}$. From now on, we consider $\mathcal{D} = W_{ij}$ defined in (2); the above result on the upper envelope is substantially unchanged, save an extra inequality for $n = 2$.

Proposition 2. *If $\beta \geq 1$ and $n > 2$, the upper envelope of f over $X \cap W_{ij}$ is*

$$H = \{(\boldsymbol{x}, z) \in \mathbb{R}^n_+ \times \mathbb{R} : z \leq u, px_i \leq x_j \leq qx_i, \textstyle\prod_{k\in N} x_k^{a_k} \geq \ell, \\ (z - z_0)^\beta \leq \gamma \textstyle\prod_{k\in N} x_k^{a_k}\}.$$

Proposition 3. *If $\beta \leq 1$ and $n > 2$, the upper envelope of f over $X \cap W_{ij}$ is*

$$H = \{(\boldsymbol{x}, z) \in \mathbb{R}^n_+ \times \mathbb{R} : z \leq u, px_i \leq x_j \leq qx_i, \textstyle\prod_{k\in N} x_k^{a_k} \geq \ell \\ z \leq \textstyle\prod_{k\in N} x_k^{a_k}\}.$$

Lower envelope over $X \cap W_{ij}$. We build on a property of the monomial function f for general $n \geq 2$ to derive a few results leading to the lower envelope of f over $X \cap W_{ij}$ for $n = 2$. Define the level set $C_\xi = \{\boldsymbol{x} \in \mathbb{R}^n_+ : f(\boldsymbol{x}) = \xi\}$ and the two sets $P_{ij} = \{\boldsymbol{x} \in \mathbb{R}^n_+ : x_j = px_i\}$ and $Q_{ij} = \{\boldsymbol{x} \in \mathbb{R}^n_+ : x_j = qx_i\}$. There is a bijection from $C_\xi \cap P_{ij}$ to $C_\xi \cap Q_{ij}$ that joins all pairs of points by parallel lines.

It is easy to show that for $n = 2$ and $a_1 = a_2 = 1$, $C_\xi \cap P_{ij} = \{\check{\boldsymbol{x}}\}$ and $C_\xi \cap Q_{ij} = \{\hat{\boldsymbol{x}}\}$ where

$$\check{\boldsymbol{x}} = \left(\sqrt{p\xi}, \sqrt{\frac{1}{p}\xi}\right), \quad \hat{\boldsymbol{x}} = \left(\sqrt{q\xi}, \sqrt{\frac{1}{q}\xi}\right).$$

The line through $\check{\boldsymbol{x}}$ and $\hat{\boldsymbol{x}}$ has slope $\frac{1/\sqrt{p}-1/\sqrt{q}}{\sqrt{p}-\sqrt{q}} = -(pq)^{-\frac{1}{2}}$, independent of ξ. This holds for general positive exponents $(a_1, a_2) \neq (1, 1)$, in which case the coefficients of x_1 and x_2 are proportional to $\xi^{\frac{1}{\beta}}$ rather than $\sqrt{\xi}$. This generalizes to $n \geq 2$.

Lemma 2. *Given $\boldsymbol{a} \in \mathbb{R}^n_+$, $i, j \in N$, $\xi \in \mathbb{R}$, $p, q \in \mathbb{R}_+$ with $i \neq j$ and $0 < p < q$, there exist $d_i < 0$ and $d_j > 0$ such that for any $\check{\boldsymbol{x}}$ that satisfies $\check{x}_j = p\check{x}_i$ and $\prod_{i\in N} \check{x}_i^{a_i} = \xi$, there exists a unique solution $(\bar{s}, \bar{\boldsymbol{x}}) \in \mathbb{R}_+ \times \mathbb{R}^n_+$ to the system*

$$\begin{array}{ll} \bar{x}_j = q\bar{x}_i & (\bar{x}_i, \bar{x}_j) = (\check{x}_i + \bar{s}d_i, \check{x}_j + \bar{s}d_j) \\ \bar{x}_k = \check{x}_k \quad \forall k \notin \{i, j\} & \prod_{i\in N} \bar{x}_i^{a_i} = \xi. \end{array}$$

Although $\bar{s}$ does depend on $\check{\boldsymbol{x}}$, the key fact is that direction (d_i, d_j) defines a bijection from P_{ij} to Q_{ij} by joining pairs of points that have same value of the function $\prod_{i\in N} x_i^{a_i}$. This suggests that on a direction orthogonal to (d_i, d_j), i.e., $(d_j, -d_i)$, we can define a lower-bounding function that matches the values on P_{ij} and Q_{ij} and that, for $n = 2$, is the lower envelope of $f(\boldsymbol{x})$.

Proposition 4. *The function $f'_\ell(\boldsymbol{x}) = \lambda(d_j x_i - d_i x_j)^{a_i+a_j} \prod_{k\in N\setminus\{i,j\}} x_k^{a_k}$, with $\lambda = p^{a_j} / (d_j - d_i p)^{a_i+a_j}$,*

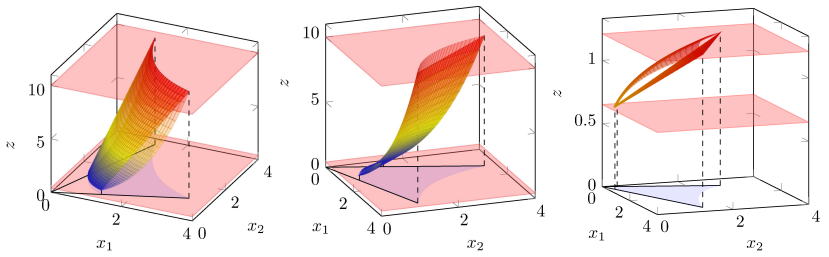

FIGURE 1. Upper envelope $f'_u(\boldsymbol{x})$ and lower envelope $f'_\ell(\boldsymbol{x})$ of function $f(\boldsymbol{x}) = x_1^{a_1} x_2^{a_2}$ for $\beta \geq 1$; and lower envelope $f''_\ell(\boldsymbol{x})$.

(1) *matches the value of* $f(\boldsymbol{x})$ *for* $\boldsymbol{x} \in P_{ij} \cup Q_{ij}$*;*
(2) *is a minorant of* $f(\boldsymbol{x})$ *for* $\boldsymbol{x} \in W_{ij}$*;*
(3) *is, for* $n = 2$ *and* $\beta = a_1 + a_2 \geq 1$*, the lower envelope of* $f(\boldsymbol{x})$ *in* W_{ij}*.*

Proposition 5. *The function* $f''_\ell(\boldsymbol{x}) = \zeta(d_j x_i - d_i x_j)^{\frac{a_i+a_j}{\beta}} \left(\prod_{k\in N\setminus\{i,j\}} x_k^{a_k}\right)^{\frac{1}{\beta}} + z_0$*, for* $\beta \leq 1$ *and with* $\zeta = \lambda^{1/\beta} \frac{u-\ell}{u^{1/\beta} - \ell^{1/\beta}}$ *and* z_0 *as defined above,*

(1) *matches* $f(\boldsymbol{x})$ *at* $(P_{ij} \cup Q_{ij}) \cap (C_\ell \cup C_u)$*;*
(2) *is a minorant of* f *in* $X \cap W_{ij}$*;*
(3) *is the lower envelope of* f *over* $\mathrm{conv}((P_{ij} \cup Q_{ij}) \cap (C_\ell \cup C_u))$ *for* $n = 2$*.*

The above results give the convex envelope of set $F(\mathcal{D})$ in $n = 2$ dimensions.

REFERENCES

[1] K.M. Anstreicher, S. Burer, and K. Park, *Convex hull representations for bounded products of variables*, Journal of Global Optimization **80** (2021), 757–778.
[2] F.A. Al-Khayyal and J.E. Falk, *Jointly constrained biconvex program- ming*. Mathematics of Operations Research, 8:273–286, 1983.
[3] P. Belotti, A.J. Miller, and M. Namazifar, *Valid inequalities and convex hulls for multilinear functions*. Electronic Notes in Discrete Mathematics, 36:805–812, 2010. ISSN 1571-0653.
[4] P. Belotti, *Convex envelopes of bounded monomials on two-variable cones*. Optimization online, August 2023. `https://optimization-online.org/2023/08/convex-envelopes-of-bounded-monomials-on-two-variable-cones`
[5] R. Horst and H. Tuy, *Global Optimization*. Springer-Verlag, New York, 1993.
[6] G.P. McCormick, *Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems*. Mathematical Programming, 10:146–175, 1976.
[7] C.A. Meyer and C.A. Floudas, *Trilinear monomials with mixed sign domains: Facets of the convex and concave envelopes*. Journal of Global Optimization, 29(2), 2004.
[8] T.T. Nguyen, J.-P.P. Richard, and M. Tawarmalani, *Deriving convex hulls through lifting and projection*, Mathematical Programming, **169**(2) (2018), 377–415
[9] A.D. Rikun. *A convex envelope formula for multilinear functions*. Journal of Global Optimization, 10:425–437, 1997.
[10] M. Tawarmalani and J.-P.P. Richard, *Decomposition techniques in convexification of inequalities*. Technical report, Technical report, 2013.

A Complete characterization of the multilinear polytope of acyclic hypergraphs

Aida Khajavirad
(joint work with Alberto Del Pia)

We consider the multilinear polytope defined as the convex hull of the set of binary points z, satisfying a collection of equations of the form $z_e = \prod_{v \in e} z_v$ for all $e \in E$. The complexity of the facial structure of the multilinear polytope is closely related to the acyclicity degree of the underlying hypergraph. We characterize the classes of acyclic hypergraphs for which the corresponding multilinear polyope has a polynomial-size extended formulation.

A *hypergraph* G is a pair (V, E), where V is a finite set of nodes and E is a set of *edges*, which are subsets of V of cardinality at least two. With any hypergraph $G = (V, E)$ we associate the *multilinear set*:

$$\mathcal{S}_G := \Big\{ z \in \{0,1\}^{V+E} : z_e = \prod_{v \in e} z_v,\ \forall e \in E \Big\},$$

and we refer to its convex hull, as the *multilinear polytope* and denote it by MP_G. Unlike graphs, the notions of cycles and acyclicity in hypergraphs are not unique. The most well-known types of acyclic hypergraphs, in increasing order of generality, are Berge-acyclic, γ-acyclic, β-acyclic, and α-acyclic hypergraphs.

In [1], we prove that MP_G coincides with its standard linearization, if and only if the hypergraph G is Berge-acyclic. This in turn implies that if G is Berge-acyclic, then MP_G is defined by $|V| + (r+2)|E|$ inequalities in the original space. In [1], we introduce *flower inequalities*, a class of facet-defining inequalities for the multilinear polytope, and show that the polytope obtained by adding all such inequalities to the standard linearization coincides with MP_G if and only if the hypergraph G is γ-acyclic. This result implies that if G is γ-acyclic, then MP_G has a polynomial-size extended formulation with at most $|V| + 2|E|$ variables and at most $|V| + (r+2)|E|$ inequalities. Subsequently, in [2], we introduce *running intersection inequalities*, a class of facet-defining inequalities for the multilinear polytope that serve as a generalization of flower inequalities. We prove that for kite-free β-acyclic hypergraphs, a class that lies between γ-acyclic and β-acyclic hypergraphs, the polytope obtained by adding all running intersection inequalities to the standard linearization coincides with MP_G, and it admits a polynomial-size extended formulation with at most $|V| + 2|E|$ variables and at most $|V| + (r+2)|E|$ inequalities. Finally, in [3], we present present a polynomial-size extended formulation for the multilinear polytope of β-acyclic hypergraphs with at most $(r-1)|V| + |E|$ variables and at most $(3r-4)|V| + 4|E|$ inequalities.

At the other end of the spectrum, in [4], the authors prove that a binary polynomial optimization problem is strongly NP-hard over α-acyclic hypergraphs. This result implies that, unless P = NP, one cannot construct a polynomial-size extended formulation for the multilinear polytope of α-acyclic hypergraphs.

References

[1] Del Pia, A. and Khajavirad, A., *The multilinear polytope for acyclic hypergraphs*, SIAM Journal on Optimization **28** (2018), 1049–1076.

[2] Del Pia, A. and Khajavirad, A., *The running intersection relaxation of the multilinear polytope*, Mathematics of Operations Research **46** (2021), 1008–1037.

[3] Del Pia, A. and Khajavirad, A., *A polynomial-size extended formulation for the multilinear polytope of beta-acyclic hypergraphs*, Mathematical Programming, Series A, (2023).

[4] Del Pia, A. and Di Gregorio, S., *On the complexity of binary polynomial optimization over acyclic hypergraphs*, Algorithmica **85** (2023), 2189–2213.

Sensitivity analysis for mixed binary quadratic problems: complexity, structure, and computation

Santanu S. Dey

(joint work with Diego Cifuentes, Jingye Xu)

A mixed binary quadratic problem (MBQP) is a problem of the form:

$$\begin{aligned} f(b) := \quad \min \quad & x^\top Q x + c^\top x \\ \text{s.t.} \quad & Ax \leq b \\ & x \in \{0,1\}^{n_1} \times \mathbb{R}^{n_2}_+, \end{aligned}$$

where Q is a real symmetric $n \times n$ matrix, $A \in \mathbb{R}^{n \times m}$, $c \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$.

In many operational applications, it is necessary to routinely find, within a very limited time window, provably good solutions to challenging MBQPs. An example is the Security-Constrained Unit Commitment (SCUC) problem, solved daily to clear the day-ahead electricity markets. In such operational applications, instances are significantly similar to each other. Specifically, instances typically share the same size and problem structure, with differences only in right-hand sides and objective function. This motives the need to conduct sensitivity analysis on MBQPs.

We first analyze the formal complexity question of conducting sensitivity analysis for MBQPs. Let $\delta \in \mathbb{R}^m$ and let $\Delta f(\delta) = \|f(b+\delta) - f(b)\|$. We want to approximate $\Delta f(\delta)$ as a function of δ. Formally, an algorithm is called (α, β)-approximation for some $\beta \geq \alpha > 0$ if it takes as input: (i) an MBQP instance $(A, b, c, Q, f(b), \delta)$, (ii) the optimal objective value $f(b)$, and (iii) $\delta \in \mathbb{R}^m$, and outputs p satisfying:

$$\alpha \cdot \Delta f(\delta) \leq p \leq \beta \cdot \Delta f(\delta).$$

We prove that it is NP-hard to achieve (α, β)-approximation for any $\beta \geq \alpha > 0$ for general MBQPs.

Next, we leverage Sam Burer's completely-positive (CPP) reformulation [1] of MBQPs by examining its dual, the so-called co-positive (COP) problem, and use this dual to obtain bounds with respect to changing rhs. We show that strong duality holds between the CPP reformulation and its COP dual if the feasible region of the MBQP is bounded or if the objective function of the MBQP is

convex. When the feasible region is unbounded and the objective function is a non-convex quadratic, we show using examples that strong duality may not hold.

We next show that the dimension of the set of optimal solutions of the COP dual is at least m, where the original MBQP has m constraints. The choice of optimal solution of the dual affects the quality of bounds we obtain when changing the rhs. Finally, we provide an algorithmic approach to find "best values" of optimal dual solutions and present preliminary computational results on sensitivity analysis for MBQPs.

References

[1] S. Burer, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Mathematical Programming **120** (2009), 479–495.

Data-driven distributional robustness over time: How to learn uncertainties with robust decisions

Frauke Liers

(joint work with K. Aigner, A. Bärmann, K. Braun, S. Pokutta, O. Schneider, K. Sharma, S. Tschuppik)

Classical stochastic Optimization (SO) typically requires knowledge about the probability distribution of uncertain parameters. As the latter is often unknown, Distributionally Robust Optimization (DRO) provides a strong alternative that determines the best guaranteed solution over a set of distributions (ambiguity set).

In this talk, we present a DRO approach that iteratively incorporates such information over time, [1]. Specifically, we provide an online learning algorithm that solves DRO problems with limited initial knowledge about the uncertainty, but which can leverage additional incoming data. This allows the optimal solutions to adapt to the uncertainty and gradually reduce the cost of protection. To this end, we use scenario observations arriving as a data stream to construct and update the ambiguity sets.

The two key differences in our work which distinguish it from regular online optimization are (i) use of DRO while learning from data and (ii) solving the DRO problem approximately. The first ensures that our solutions are robust to uncertainty in the knowledge of the true probability distribution. The second allows us to obtain robust solutions without solving the DRO problem exactly at each step. Specifically, the key contributions of our work are:

Online Learning Algorithm for DRO. We provide an online algorithm to solve the DRO problem. It also learns the uncertainty from scenario observations over time, shrinking the ambiguity sets. This allows for rapid computation of the DRO solutions along with their adaptation. Thus, reducing the cost of protection.

Stochastic Consistency. We also prove that the solution of the DRO problem converges to the SO problem. Since our online algorithm solves the DRO problem, thus it too converges to the solution of the SO problem.

High Probability Regret Bounds. We prove that the cumulative regret between the solutions generated by our online method and the exact DRO solution at each time step shrinks at a rate of $\mathcal{O}(\log T/\sqrt{T})$ with high probability.
Flexibility of Uncertainty Models. We consider 3 different ambiguity models:

- confidence intervals,
- ℓ_2-norm sets
- kernel based ambiguity sets.

These allow our approach to adapt to the application.

Computational Results. We provide a computational study on mixed-integer benchmark instances and on real world problem examples. Specifically, we compare on the MIPLIB and QPLIB libraries and further illustrate our results with two realistic applications from telecommunications and routing. We demonstrate that our online method leads to significantly reduced computation times with only marginal sacrifices in solution quality.

References

[1] K. Aigner, A. Bärmann, K. Braun, F. Liers, S. Pokutta, O. Schneider, K. Sharma, S. Tschuppik. *Data-Driven Distributionally Robust Optimization over Time*, INFORMS Journal on Optimization, https://doi.org/10.1287/ijoo.2023.0091.

SDP hierarchies for distance-avoiding sets on compact spaces

Bram Bekker

(joint work with Olga Kuryatnikova, Fernando Mário de Oliveira Filho, Juan C. Vera)

Witsenhausen's problem [6] asks for the largest measurable set on a unit sphere such that no two points in the set are orthogonal. More precisely, let $S^{n-1} = \{x \in \mathbb{R}^n \colon \|x\| = 1\}$ be the n-dimensional unit sphere, ω be the standard surface measure, and ω_n the total measure of the sphere. We define

$$\alpha_n = \sup\{\omega(I)/\omega_n \colon I \subset S^{n-1} \text{ does not contain orthogonal pairs}\}.$$

The Double Cap Conjecture claims that the largest such set is achieved by taking the union of two open antipodal spherical caps of angular size $\pi/4$, i.e.

$$\{x \in S^{n-1} \colon \|e \cdot x\| > \cos(\pi/4)\},$$

for some fixed point $e \in S^{n-1}$. Optimality of this set was conjectured by Kalai [3, Conjecture 2.8].

We model this problem as an independent set problem on a graph. Let $G = (S^{n-1}, E)$ be a graph with vertex set S^{n-1}. Two points $x, y \in S^{n-1}$ share an edge if and only if $x \cdot y = 0$, i.e. if and only if they are orthogonal. The sets on the sphere avoiding orthogonal pairs are exactly the independent sets of this graph.

A popular starting point for finding upper bounds to independent set problems is the Lovász ϑ-number, is given by

$$\begin{aligned}\vartheta(G) = \sup\ & \textstyle\int_{S^{n-1}} \int_{S^{n-1}} A(x,y) d\omega(x) d\omega(y) \\ & \textstyle\int_{S^{n-1}} A(x,x) d\omega(x) = 1, \\ & A(v,w) = 0 \quad \text{for all } (v,w) \in E, \\ & A \text{ continuous and PSD.}\end{aligned}$$

We call a symmetric L^2-function $A : (S^{n-1})^2 \to \mathbb{R}$ a *kernel* and denote the set of kernels by $L^2_{\text{sym}}(S^{n-1})$, and we call a continuous kernel A *positive semidefinite* (PSD) if and only if every finite submatrix of A is positive semidefinite.

The Lovász ϑ-number often gives quite weak bounds, so the question arises how to strengthen it. We have described a new approach based on results from copositive programming. The discrete setting goes back to the copositive programming bounds described by de Klerk and Pasechnik [4]. This approach was generalized to the setting of another type of graph, so-called compact packing graphs, in the PhD-thesis of Kuryatnikova [5]. We now extend these copositive programming bounds to the setting of distance-avoiding sets, where the independent sets are no longer discrete.

Earlier results by DeCorte, Oliveira and Vallentin [2] show that if we replace the PSD-cone by the *completely positive* cone

$$\text{CP}(S^{n-1}) = \text{cc}\{f \otimes f \colon f \in L^2(S^{n-1}), f \geq 0\},$$

where cc denotes the closure of the convex hull, we obtain exactly α_n. That is, let

$$\begin{aligned}\vartheta_{\text{CP}}(G) = \sup\ & \textstyle\int_{S^{n-1}} \int_{S^{n-1}} A(x,y) d\omega(x) d\omega(y) \\ & \textstyle\int_{S^{n-1}} A(x,x) d\omega(x) = 1, \\ & A(v,w) = 0 \quad \text{for all } (v,w) \in E, \\ & A \text{ continuous and } A \in \text{CP}(S^{n-1}),\end{aligned}$$

then $\alpha_n = \theta_{\text{CP}}$. However, optimizing over the completely positive cone even in the discrete setting is usually not tractable, and so we have to find tractable relaxations.

A matrix is *completely positive* if and only if it is a convex combination of rank 1 matrices of the form xx^{T}, with $x \geq 0$. We denote the cone of these matrices by $\text{CP}([n])$, where $[n]$ is the index set of A. Let $\text{COP}([n])$ be its conic dual, which we call the cone of *copositive* matrices. Note that a matrix A is in $\text{COP}([n])$ if and only if $x^{\mathsf{T}} A x \geq 0$ for all $x \geq 0$.

Pólya's theorem now claims that $x^{\mathsf{T}} A x > 0$ for all $x \geq 0$ if and only if there exists some $r \in \mathbb{N}$ such that

$$(\mathbb{1}^{\mathsf{T}} x)^r x^{\mathsf{T}} A x$$

has only non-negative coefficients. Using the euclidean inner-product $\langle \cdot, \cdot \rangle$ we can rewrite this as

$$(\mathbb{1}^{\mathsf{T}} x)^r x^{\mathsf{T}} A x = \langle A \otimes \mathbb{1}^{\otimes r}, x^{\otimes (r+2)} \rangle$$

where $A \otimes \mathbb{1}^{\otimes r}(v_1, \ldots, v_{r+2}) = A(v_1, v_2)\mathbb{1}(v_3) \cdots \mathbb{1}(v_{r+2}) = A(v_1, v_2)$, and similar for $x^{\otimes (r+2)}$.

If we replace the euclidean inner product by the L^2-inner product, and let A be a continuous kernel, the expression on the right-hand side still makes sense. Because we prefer to work with symmetric functions, we furthermore symmetrize the function $A \otimes \mathbb{1}^{\otimes r}$ over all permutations of its $r+2$ arguments by using the operator

$$\mathcal{R}_{\mathcal{S}_{r+2}}(T)(v_1, \ldots, v_{r+2}) = \frac{1}{(r+2)!} \sum_{\pi \in \mathcal{S}_{r+2}} T(v_{\pi^{-1}(1)}, \ldots, v_{\pi^{-1}(r+2)}),$$

where $\mathcal{S}_{r+2}$ denotes the permutation group on $r+2$ items. This leaves the expression above unchanged. Note that the function $\mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbb{1}^{\otimes r})$ holds the coefficients of the polynomials in the discrete setting, so it makes sense to define the cones

$$\mathcal{C}_r(V) = \{A \in L^2_{\mathrm{sym}}(S^{n-1}) : \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbb{1}^{\otimes r}) \geq 0\}.$$

Our L^2-version of Pólya's theorem now becomes

Theorem 1. *If $A \in L^{\mathrm{sym}}(S^{n-1})$ is such that $\langle A, Z \rangle \geq 0$ for all*

$$Z \in \bigcup_r \mathcal{C}_r(S^{n-1}),$$

then A is completely positive.

It turns out that cuts given by $\mathcal{C}_r(S^{n-1})$ are implementable as positive semidefinite constraints, and so we obtain a tractable hierarchy

$$\begin{aligned} \vartheta_r(G) = \sup \; & \textstyle\int_{S^{n-1}} \int_{S^{n-1}} A(x,y) d\omega(x) d\omega(y) \\ & \textstyle\int_{S^{n-1}} A(x,x) d\omega(x) = 1, \\ & A(v,w) = 0 \quad \text{for all } (v,w) \in E, \\ & A \text{ continuous and SDP, } A \in \mathrm{C}_r(S^{n-1}), \end{aligned}$$

where we still require A to be SDP.

We were able to show that this hierarchy converges to exactly α_n as r goes to ∞. We also extended the Lasserre hierarchy and the k-point bound to this setting, and showed that both are stronger than this copositive hierarchy, obtaining convergence results for both. However, neither results in tractable upper bounds.

Using the level $r = 1$ of the copositive hierarchy, we were able to calculate the best upper bounds known for Witsenhausen's problem in dimensions 3 to 8, see the table below. The lower bound comes from the Double Cap Conjecture, the percentage gap closed assumes the gap between the previous best upper bounds and the lower bound is 100%.

Dimension	Lower bound	Previous best upper bound	New best upper bound	Percentage gap closed
3	0.2928	0.3015	0.2977	43%
4	0.1816	0.2168	0.1943	64%
5	0.1161	0.1677	0.1346	64%
6	0.0755	0.1338	0.0981	61%
7	0.0498	0.1174	0.0758	62%
8	0.0331	0.0998	0.0612	58%

References

[1] B. Bekker, O. Kuryatnikova, F. M. de Oliveira Filho, and J. C. Vera, *Optimization hierarchies for distance-avoiding sets in compact spaces*, https://doi.org/10.48550/arXiv.2304.05429

[2] E. DeCorte, F.M. de Oliveira Filho, and F. Vallentin, Complete positivity and distance-avoiding sets, *Mathematical Programming, Series A* (2020) 72pp, DOI: 10.1007/s10107-020-01562-6, arXiv:1804:09099.

[3] G. Kalai, *Some old and new problems in cominatorial geometry I: around Borsuk's problem,* in: Surveys in combinatorics 2015, London Mathematical Society Lecture Note Series 424, Cambridge University Press, Cambridge, 2015, pp. 147–174

[4] E. de Klerk, and D. V. Pasechnik, *Approximation of the Stability Number of a Graph via Copositive Programming*, SIAM Journal on Optimization, **4** (2002), 875–892

[5] O. Kuryatnikova, *The many faces of positivity to approximate structured optimization problems*, Ph.D. Thesis, Tilburg University, (2019)

[6] H. S. Witsenhausen, *Spherical Sets Without Orthogonal Point Pairs*, The American Mathematical Monthly, **10** (1974) 1101–1102, https://doi.org/10.1080/00029890.1974.11993744

Projected Eigenvector Cuts for Solving Sparse Semidefinite Programs

Jeff Linderoth

(joint work with Oktay Günlük, James Luedtke, Andrea Lodi)

Consider the following semidefinite programming relaxation of a nonconvex quadratically-constrained quadratic program, first proposed by Shor [1].

$$\text{(P)} \qquad z_P := \min_{x,X} \ \langle X, Q^0\rangle + c_0^\top x$$
$$\text{s.t. } \langle X, Q^k\rangle + c_k^\top x + d_k \le 0, \quad k = 1, \dots, m$$
$$Y := \begin{bmatrix} 1 & x^\top \\ x & X \end{bmatrix} \succeq 0. \tag{1}$$

We are especially interested in the case where the matrices Q^k for $k = 0, 1, \dots, m$ are sparse, and we define the set $E := \{ij : \exists k \text{ with } Q^k_{ij} \neq 0\}$ as the set of i, j pairs in which one of the matrices Q^k has a nonzero element. We wish to create a relaxation of our problem (P) in which we introduce decision variables X_{ij} only for $ij \in E$.

To that end, suppose we have solved a relaxation of (P) where the semidefinite constraint (1) has not been enforced, and we have a solution $\hat{Y}$. Our goal is to

identify, if possible, a linear inequality

$$\langle C, Y \rangle \geq 0 \tag{2}$$

valid for the convex cone of positive semidefinite matrices $\mathcal{S}^+_{n+1}$ which is violated by $\hat{Y}$. The cone of semidefinite matrices $\mathcal{S}^+_n$ is self-dual, so inequality (2) is valid for $\mathcal{S}^+_{n+1}$ if and only if the cut coefficients $C \in \mathcal{S}^+_{n+1}$. We wish to generate a cut that has C supported only on E, therefore we obtain the following constrained separation problem:

$$\text{(SEP)} \qquad \min_{C \in \mathcal{S}^+_{n+1}} \{\langle C, \hat{Y} \rangle \;\mid\; \mathrm{Tr}(C) = 1, C_{ij} = 0 \quad \forall ij \notin E\},$$

where the constraint $\mathrm{Tr}(C) = 1$ is a normalization condition.

A natural question to ask is what relaxation bound can be obtained using only projected inequalities coming from repeatedly solving the separation problem (SEP). The following short lemma establishes that one can obtain the original Shor bound z_P.

Lemma 1. *Let*

$$X = (P \times \mathbb{R}^k) \cap Q$$

where $P \subset \mathbb{R}^n$ and $Q \subset \mathbb{R}^{n+k}$. Then,

$$\mathrm{proj}_{\mathbb{R}^n}(X) \;=\; P \cap \mathrm{proj}_{\mathbb{R}^n}(Q)$$

Proof. If $x \in \mathrm{proj}_{\mathbb{R}^n}(X)$ then for some $w \in \mathbb{R}^k$ we have $(x, w) \in P \times \mathbb{R}^k$ and $(x, w) \in Q$. Therefore, $x \in P$ and $x \in \mathrm{proj}_{\mathbb{R}^n}(Q)$.

If $x \in P \cap \mathrm{proj}_{\mathbb{R}^n}(Q)$, then for some $w \in \mathbb{R}^k$ we have $(x, w) \in Q$ and $x \in P$. As $(x, w) \in P \times \mathbb{R}^k$, the claim follows. □

However, to obtain the strength of the Shor bound requires solving a (structured) semidefinite programming problem for separating a (sparse) point $\hat{Y}$ from $\mathcal{S}^+_{n+1}$. To avoid solving an SDP to generate a cut, one can instead solve a sequence of linear programs that restrict the cut generation search space. Let $C^1, \ldots, C^q$ be a given set of PSD matrices. Then a restriction of the separation problem (SEP) is

$$\begin{aligned}
\text{(LP-SEP)} \qquad & \min \langle C, \hat{Y} \rangle \\
& \text{s.t. } C_{ij} = \sum_{t=1}^{q} \lambda_t C^t_{ij}, \quad \forall ij \in E \\
& \qquad 0 = \sum_{t=1}^{q} \lambda_t C^t_{ij}, \quad \forall ij \notin E \\
& \qquad \sum_{j=1}^{n+1} C_{jj} \leq 1 \\
& \qquad \lambda_t \geq 0, \quad t = 1, \ldots, q.
\end{aligned}$$

The dual (LP-SEP) has the form:

$$\begin{aligned} \max\ & \theta \\ \text{s.t. } & \Pi_{ij} = \hat{Y}_{ij} \quad \forall ij \in E, i \neq j \\ & \langle \Pi, C^t \rangle \geq 0 \quad \forall t = 1, \ldots, q \\ & \Pi_{jj} + \theta = \hat{Y}_{jj} \quad \forall j = 1, \ldots, n+1 \\ & \theta \leq 0. \end{aligned}$$

The dual problem can be interpreted as choosing a matrix Π that matches $\hat{Y}$ on the entries in E except the diagonals, is valid for all given PSD matrices, $\langle \Pi, C^t \rangle \geq 0\ \forall t \in [q]$, and minimizes the maximum amount added to the diagonals in order to do so.

Suppose the separation problem has been solved and yielded dual solution $\hat{\Pi}$. Thus, if one wishes to generate a new C^t to include in the formulation to improve the violation of the cut found, one can attempt to find a PSD matrix $\hat{C}$ (with no restrictions on sparsity) with $\langle \hat{\Pi}, \hat{C} \rangle < 0$ and add this to the list of C^t. This could be done, for example, by doing an eigenvalue decomposition on the matrix Π and using eigenvectors associated with negative eigenvalues to define the cut (i.e., if μ is such an eigenvector, $C = \mu\mu^\top$ would suffice). In the limit of this procedure, if $\hat{\Pi} \succeq 0$ (so no more such cuts can be found), this yields an exact separation of the sparse PSD cut. Continuing work will be to implement this computational procedure, including the consideration of relaxations of the PSD constraint that enforce the positive-semidefnite condition only on certain $k \times k$ minors of Y.

References

[1] N. Z. Shor, *Quadratic Optimization Problems*, Soviet Journal of Circuits and Systems Sciences **25** (1987), 1–11.

Convexification techniques for fractional programming

Mohit Tawarmalani
(joint work with Taotao He, Siyue Liu)

Fractional programming problems arise in many optimization models including those involving consumer choice functions, financial and performance ratios, and matrix eigenvalues. Luce-type choice functions have been used in feature selection, assortment optimization, and facility location while financial and performance ratios are used to measure return on investment, performance of restricted properties and other graph properties such as density and connectedness.

Seminal results in fractional programming have shown techniques to maximize $\frac{1+a^\intercal x}{1+b^\intercal x}$ over a polytope [2] or a combinatorial set [4]. At the same time, it is known that constructing the concave envelope of a general fractional function over

a polytope is not tractable. The current relaxation schemes in mixed-integer nonlinear programming create relaxations with a non-zero gap between the relaxation optimal value and the optimal value of the rational function.

In this talk, which is based on [3] we construct a new class of relaxations for fractional programming problems that relies on a precise correspondence between two convex hulls. The first convex hull convexifies functions $f_i(x)$ for $i \in 1, \ldots, m$ over $x \in X$ while the second convex hull convexifies $\frac{\left(1, f_1(x), \ldots, f_m(x)\right)}{1+\sum_{i=1}^{m} \alpha_i f_i(x)}$ over the same domain assuming that the denominator is sign-invariant. This correspondence describes a formulation for one of the convex hulls given that for the other and how points can be separated from one set given a separation procedure for the other. As such, this result generalizes various existing results in the literature.

We discuss applications in various contexts. First, we derive a hierarchy of relaxations converging to the convex hull of fractional functions over 0-1 hypercube using relaxation hierarchies for multilinear polytopes. We show computationally that, even at the first level, this relaxation is significantly tighter than the traditional relaxation schemes. Our results also give insights into hardness results in fractional programming. For example, we provide a simpler proof of the NP-Hardness of optimizing a linear perturbation of a 0-1 fractional function using the correspondence with bilinear optimization.

We develop new relaxations for a ratio of quadratic functions over an ellipsoid using semidefinite programming reproducing an earlier result of [1]. More generally, we show that copositive programming can be used to minimize a ratio of two quadratic function over an arbitrary polytope. Finally, we relate the convexification of $\left(\frac{1}{x+a_1}, \ldots, \frac{1}{x+a_m}, x\right)$ with that the convex hull of $(x, \ldots, x^{m-1})$ over an interval such that the signs of denominators do not vary. This set has applications in distillation configuration design and optimizing exergy for separation processes.

References

[1] Beck, Amir, and Marc Teboulle, *A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid*, Mathematical programming 118, no. 1 (2009): 13-35.

[2] Charnes, Abraham, and William W. Cooper, *Programming with linear fractional functionals*, Naval Research logistics quarterly 9, no. 34 (1962): 181–186.

[3] He, Taotao, and Siyue Liu and Mohit Tawarmalani, *Convexification techniques for fractional programming*, working paper, 2023.

[4] Megiddo, Nimrod, *Combinatorial optimization with rational objective functions*, In Proceedings of the tenth annual ACM symposium on Theory of computing, pp. 1-12. 1978.

Reporter: Daniel Brosch

Participants

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

Prof. Dr. Kurt M. Anstreicher
Department of Business Analytics
The University of Iowa
Iowa City, IA 52242-1994
UNITED STATES

Bram Bekker
Faculty of Mathematics and
Informatics
Delft University of Technology
P. O. Box 356
2628 BL Delft
NETHERLANDS

Prof. Dr. Pietro Belotti
Dipartimento di Elettronica,
Informazione e Bioingegneria
Politecnico di Milano
Piazza Leonardo da Vinci, 32
20133 Milano
ITALY

Dr. Merve Bodur Cevik
University of Toronto
Dept. of Mechanical and Industrial
Engineering
5 King's College Road
Toronto M5S 3G8
CANADA

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

Prof. Dr. Christoph Buchheim
Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund
GERMANY

Dr. Dominik H. Cebulla
Institut für Mathematische Optimierung
TU Braunschweig
Universitätsplatz 2
38106 Braunschweig
GERMANY

Dr. Martina Cerulli
Dipartimento di Matematica
Università di Salerno
84084 Fisciano SA
ITALY

Diego Cifuentes
School of Industrial and Systems
Engineering
Georgia Institute of Technology
765 Ferst Drive
Atlanta, GA 30332-0205
UNITED STATES

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

Prof. Dr. Santanu S. Dey
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
765 Ferst Drive
Atlanta, GA 30332
UNITED STATES

Prof. Dr. Moritz Diehl
Institut für Mikrosystemtechnik (IMTEK) und Mathematisches Institut
Albert-Ludwigs-Universität Freiburg
Georges-Köhler-Allee 102
79110 Freiburg i. Br.
GERMANY

Prof. Dr. Marcia Fampa
Universidade Federal do Rio de Janeiro
Rio de Janeiro 21941-914
BRAZIL

Dr. Anja Fischer
FB Wirtschaftswissenschaften
Universität Dortmund
Vogelpothsweg 87
Postfach 500500
44221 Dortmund
GERMANY

Dr. Frank Fischer
Institut für Informatik
Johannes Gutenberg-Universität Mainz
55099 Mainz
GERMANY

Prof. Dr. Laura Galli
Dipartimento di Informatica
- Operations Research -
Università di Pisa
Largo B. Pontecorvo, 3
56127 Pisa
ITALY

Prof. Dr. Falk M. Hante
Institut für Mathematik
Fachbereich Angewandte Mathematik
Humboldt-Universität zu Berlin
10099 Berlin
GERMANY

Prof. Dr. Christoph Helmberg
Fakultät für Mathematik
Technische Universität Chemnitz
Reichenhainer Straße 39
09126 Chemnitz
GERMANY

Dorothee Henke
Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund
GERMANY

Prof. Dr. Volker Kaibel
Fakultät für Mathematik
Otto-von-Guericke-Universität Magdeburg
Universitätsplatz 2
39106 Magdeburg
GERMANY

Dr. Aida Khajavirad
Department of Industrial and Systems Engineering, Lehigh University
200 West Packer ave
Bethlehem 18015
UNITED STATES

Dr. Fatma Kilinc-Karzan
Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213-3890
UNITED STATES

Prof. Dr. Christian Kirches
Institute for Mathematical Optimization
Technische Universität Braunschweig
Universitätsplatz 2
38106 Braunschweig
GERMANY

Dr. Jan Kronqvist
Department of Mathematics
KTH
10044 Stockholm
SWEDEN

Prof. Dr. Simge Küçükyavuz
Department of Industrial Engineering
and Management Sciences
Tech Institute
Northwestern University
2145 Sheridan Road
Evanston, IL 60208-2730
UNITED STATES

Prof. Dr. Jon Lee
Department of Industrial and Operations
Engineering, University of Michigan
1205 Beal Avenue
Ann Arbor, MI 48109-2117
UNITED STATES

Dr. Sven Leyffer
Mathematics and Computer Science
Division
Argonne National Laboratory
9700 South Cass Avenue
Argonne, IL 60439-4844
UNITED STATES

Prof. Dr. Leo Liberti
LIX CNRS, École Polytechnique,
Institut Polytechnique de Paris
91128 Palaiseau Cedex
FRANCE

Prof. Dr. Frauke Liers
Department of Data Science
Universität Erlangen-Nürnberg
Cauerstraße 11
91058 Erlangen
GERMANY

Prof. Dr. Jeffrey T. Linderoth
Department of Industrial Engineering
The University of Wisconsin-Madison
1513 University Avenue
Madison, WI 53706-1572
UNITED STATES

Paul Manns
Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund
GERMANY

Julius Martensen
Fakultät für Mathematik
Otto-von-Guericke-Universität
Magdeburg
Universitätsplatz 2
39016 Magdeburg
GERMANY

Prof. Dr. Maximilian Merkert
Institut für Mathematische Optimierung
TU Braunschweig
Universitätsplatz 2
38106 Braunschweig
GERMANY

Dr. Ruth Misener
Department of Computing
Imperial College London
Huxley Building
London SW7 2AZ
UNITED KINGDOM

Prof. Dr. Marc Pfetsch
Fachbereich Mathematik
TU Darmstadt
Dolivostr. 15
64293 Darmstadt
GERMANY

Christoph Plate
Fakultät für Mathematik
Otto-von-Guericke-Universität
Magdeburg
Universitätsplatz 2
39016 Magdeburg
GERMANY

Prof. Dr. Andreas Potschka
Institut für Mathematik
Technische Universität Clausthal
Erzstr. 1
38678 Clausthal-Zellerfeld
GERMANY

Diane Puges
Institut für Mathematik
Alpen-Adria-Universität Klagenfurt
Universitätsstr. 65-67
9020 Klagenfurt
AUSTRIA

Dr. Jan Rolfes
Analytics and Mixed-Integer
Optimization
Department of Data Science
FAU Erlangen-Nürnberg
91058 Erlangen
GERMANY

Florian Rösel
Department Mathematik
Universität Erlangen-Nürnberg
Cauerstr. 11
91058 Erlangen
GERMANY

Prof. Dr. Sebastian Sager
Fakultät für Mathematik
Otto-von-Guericke-Universität
Magdeburg
Universitätsplatz 2
39016 Magdeburg
GERMANY

Prof. Dr. Nick Sahinidis
Georgia Institute of Technology
755 Ferst Dr NW
Atlanta, GA 30332
UNITED STATES

Jan Schwiddessen
Institut für Mathematik
Alpen-Adria-Universität Klagenfurt
Universitätsstr. 65-67
9020 Klagenfurt
AUSTRIA

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

Dr. Matteo Tacchi
GIPSA-lab & CNRS
11 rue des Mathématiques
Grenoble Campus BP46
38402 Saint-Martin-d'Hères Cedex
FRANCE

Prof. Dr. Mohit Tawarmalani
Krannert School of Management
Purdue University
4095 Rawls Hall
100, S. Grant Street
West Lafayette IN 47907-2076
UNITED STATES

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

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

Liding Xu
LIX CNRS, École Polytechnique,
Institut Polytechnique de Paris
91128 Palaiseau
FRANCE

Dr. Shudian Zhao
Department of Mathematics
KTH
10044 Stockholm
SWEDEN