FGS 2024

French-German-Spanish Conference on Optimization

June 18-21, 2024

Universidad de Oviedo. Gijón (Asturias)

Book of abstracts

www.unioviedo.es/fgs2024

Sponsors

Departamento de Matemáticas
Universidad de Oviedo

EPI Gijón
Universidad de Oviedo

visita gijón Convention Bureau

accenture

Fundación Cajastur

SeMA
Sociedad Española
de Matemática Aplicada

AMPL

FGS 2024

French-German-Spanish Conference on Optimization

Universidad de Oviedo

Gijón

June 18-21, 2024

PRESENTATION

The first edition of the of the French-German Congress of Optimization started back in 1980 in Oberwolfach (Germany) as a meeting of French and German mathematicians who shared a common interest in different aspects of mathematical optimization. Since this edition, researchers of other countries have participated in this congress and, starting in 1998 a third country is invited to participate in the organization.

On this occasion, we have had the honour to be chosen to host it in Spain, so this is a French-German-Spanish conference (FGS2024, https://www.unioviedo.es/fgs2024/). It takes place in the Campus of Gijón of the Universidad de Oviedo, in the Spanish region of Asturias.

In this book you can find the one-page-abstracts of the one hundred and thirty-four talks that will be presented along the conference. The congress includes twelve plenary speakers, thirteen mini-symposia and six thematic sessions, where theoretical aspects, practical optimization methods and applications are discussed. Together with the classical topics in optimization such as mixed or integer programming, convex optimization, algorithms, or optimal control, in this edition we can also find an increasing number of contributions dealing with machine learning, neural networks, data mining and other applications to artificial intelligence, a field where optimization plays a crucial role in the design and understanding of efficient algorithms.

We would like to thank all the people that have made this congress possible: all the authors and coauthors of presentations, the scientific committee, the technical secretariat of the Foundation Universidad de Oviedo, the students that have volunteered, and our sponsors.

The local organizing committee from the Universidad de Oviedo

Contents

Committee

Local

Mariano José Mateos Alberdi
Universidad de Oviedo

Pedro Alonso Velázquez
Universidad de Oviedo

Rafael Gallego Amez
Universidad de Oviedo

Jorge Jiménez Meana
Universidad de Oviedo

Antonio Palacio Muñiz
Universidad de Oviedo

Set Pérez González
Universidad de Oviedo

Maria Luisa Serrano Ortega
Universidad de Oviedo

Jesús Suárez Pérez del Río
Universidad de Oviedo

Technical Secretariat

Adriana Suárez Paredes
Fundación Universidad de Oviedo

Vanessa Cuenco Díaz
Fundación Universidad de Oviedo

Committee

Scientific

Anne Auger
INRIA Saclay - Île-de-France

Eduardo Casas
Universidad de Cantabria

Antonin Chambolle
Université Paris – Dauphine

Elena Fernández
Universidad de Cádiz

Enrique Fernández-Cara
Universidad de Sevilla

Didier Henrion
LAAS-CNRS Toulouse

Marco Antonio López-Cerdá
Universidad de Alicante

Alexander Martin
Friedrich-Alexander-Universität Erlangen-Nürnberg

Helmut Maurer
University of Münster

Sabine Pickenhain
Brandenburgische Technische Universität

Jérôme Renault
Toulouse School of Economics

Dolores Romero-Morales
Copenhagen Business School

Michael Ulbrich
Technische Universtät München

Andrea Walter
Humboldt-Universität zu Berlin

Hasnaa Zidani
INSA-Rouen

Communications

List of

List of Abstracts

Guest Speakers

Individual Contributions

List of Minisymposia

Optimal control of ODEs: theory and applications
Organizers: Sabine Pickenhain, Valeriya Lykina

Optimal control theory and applications
Organizers: Olivier Cots

Optimization and stabilization of infinite-dimensional dynamical systems
Organizers: Behzad Azmi, Sergio S. Rodrigues

Optimization methods for inverse problems and beyond
Organizers: Neil Dizon, Jyrki Jauhiainen, Tuomo Valkonen

Shape optimization and uncertainty
Organizers: Winnifried Wollner, Kathrin Welker

Theoretical approaches to modern machine learning methods
Organizers: Riccardo Bonalli, Ziad Kobeissi, Mathieu Laurière

Uncertain optimization and related topics
Organizers: Carlos Escudero, César Gutiérrez, Miguel Sama

Speakers

Guest

Aris Daniilidis
Technische Universität Wien

Sourur Elloumi
École Nationale Supérieure de Techniques Avancées, Paris

Gersende Fort
Université de Toulouse

Emmanuel Trélat
Laboratoire Jacques-Louis Lions, CNRS Paris

Christoph Brune
Universiteit Twente

Gabriele Eichfelder
Technische Universität Ilmenau

Max Klimm
Technische Universität Berlin

Enrique Zuazua
Friedrich-Alexander Universität

Juan Enrique Martínez-Legaz
Universitat Autònoma de Barcelona

Emilio Carrizosa
Universidad de Sevilla

Anna Doubova
Universidad de Sevilla

Karl Kunisch
Universität Graz and Radon Institute, Austrian Academy of Sciences

The data-weight duality for deep learning inverse problems

Christoph Brune

c.brune@utwente.nl University of Twente, The Netherlands

Scientific machine learning combines the approximation capabilities of machine learning with modeling and numerical analysis of PDEs leading to digital twins where data-driven and model-driven techniques are combined. In this context, deep kernel representations have the potential to form an integral link between scientific computing, inverse problems, uncertainty quantification, and machine learning towards the nonlinear analysis of deep neural operators.

Reproducing kernel Hilbert spaces (RKHS) have been successful in various areas of machine learning, like kernel SVMs, relating model and feature representations. Recently, Barron spaces have been used to prove bounds on the generalization error for neural networks. Unfortunately, Barron spaces cannot be understood in terms of RKHS due to the strong nonlinear coupling of the weights in neural network representations [1]. This can be solved by using the more general Reproducing Kernel Banach spaces (RKBS).

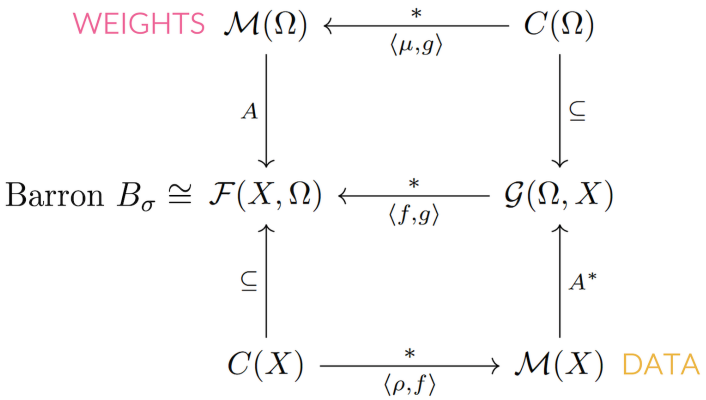


Fig. 1 Barron Space Duality for Neural Networks through Reproducing Kernel Banach Spaces

As a key result, we can show that the dual space of such RKBSs, is again an RKBS where the roles of the data and parameters are interchanged, forming an adjoint pair of RKBSs including a reproducing kernel [2].

The duality provides a natural construction of saddle point optimization for active learning and scale space flows of neural networks [3] and enables sparse architecture search, model exploration, model-order reduction, or deep kernel learning following sparse representer theorems. Impactful applications for such deep kernel representations include high-dimensional biomedical flow imaging or structure-preserving dynamics in robotic systems.

This work is based on joint collaborations with Martin Burger, Tim Roith, Len Spek, Tjeerd Jan Heeringa, and Felix Schwenninger.

References

[1] Tjeerd Jan Heeringa, Tim Roith, Christoph Brune, and Martin Burger. Learning a Sparse Representation of Barron Functions with the Inverse Scale Space Flow, December 2023. arXiv:2312.02671 [cs, math, stat]. URL: http://arxiv.org/abs/2312.02671, doi:10.48550/arXiv.2312.02671.

[2] Tjeerd Jan Heeringa, Len Spek, Felix Schwenninger, and Christoph Brune. Embeddings between Barron spaces with higher order activation functions, May 2023. arXiv:2305.15839 [cs, math, stat]. URL: http://arxiv.org/abs/2305.15839.

[3] Len Spek, Tjeerd Jan Heeringa, Felix Schwenninger, and Christoph Brune. Duality for Neural Networks through Reproducing Kernel Banach Spaces, March 2023. arXiv:2211.05020 [cs, math]. URL: http://arxiv.org/abs/2211.05020.

Mathematical optimization models in explainable machine learning. Hop on / hop off tour

Emilio Carrizosa
ecarrizosa@us.es Universidad de Sevilla, Spain

1. Introduction

In the burgeoning field of Machine Learning, Mathematical Optimization is crucial to define procedures which, on top of enjoying a high accuracy, take into account other aspects such as Transparency or Fairness. In this talk, a personal overview of the topic, with main focus on the work being developed by our team, will be given.

Transparency and Fairness can be induced in a classification/regression framework by appropriately adding constraints (soft or hard) which either allows us to control the features involved or for the classification/regression model to predict similar results, either in terms of accuracy or in terms of predicted value, for different subgroup of individuals. Illustrations in Sparse Optimal Trees and Support Vector Machines will be discussed.

We will also describe new models on Counterfactual Analysis. In Supervised Classification, Counterfactual Analysis means associating with each record a so-called counterfactual explanation, i.e., an instance close to the record (closeness measured by an appropriate and context-dependent metric) and whose probability of being in the positive class, according to the classifier given, is high. While the literature focuses on the problem of finding one counterfactual for one record, in this talk we address the more general setting in which a group of counterfactual explanations is sought for a group of instances, or several counterfactuals are to be defined for one record. Explainability will be addressed through sparsity, by controlling as an extra criterion in this multiple-objective problem the number of features involved.

Interpretable models in Multiple-Objective Inverse Optimization will be also described: Assuming pairs of contexts (parameters) and decisions are given, one seeks the (linear) functions and their associated weights such that the decisions given are optimal under the context given. Identifying weights and eventually the most relevant features in the contexts space will be expressed as Mathematical Optimization problems, whose structure will be discussed.

Acknowledgements

Research funded by PID2022-137818OB-I00/AEI/10.13039/501100011033/ FEDER, EU, and EC H2020 MSCA RISE NeEDS (Grant agreement ID 822214).

References

[1] Rafael Blanquero, Emilio Carrizosa, Cristina Molero-Río, Dolores Romero Morales. Sparsity in optimal randomized classification trees. *European Journal of Operational Research*, 284(1): 255–272, 2020.

[2] Emilio Carrizosa, Cristina Molero-Río, Dolores Romero Morales. Mathematical optimization in classification and regression trees. *TOP*, 29: 5–33, 2021.

[3] Emilio Carrizosa, Jasone Ramírez-Ayerbe, Dolores Romero Morales. Mathematical optimization modelling for group counterfactual explanations. *European Journal of Operational Research*, 2024.

Unilateral analysis, orientation and determination

Aris Daniilidis
aris.daniilidis@tuwien.ac.at (VADOR E105-04), TU Wien, Austria

The norm of the gradient $\|\nabla f(x)\|$ measures the maximum descent of a smooth function f at x. For nonsmooth convex functions, this is expressed by the remoteness of the subdifferential (that is, the distance of $\partial f(x)$ to the origin), while for general real-valued functions defined on metric spaces by the notion of metric slope $s_f(x)$ due to De Giorgi. More generally, an axiomatic definition of *descent modulus* is possible, for functions defined on general spaces (not necessarily metric), encompassing both the (metric) notion of steepest descent as well as the (probabilistic) notion of average descent for functions defined on probability spaces. A large class of functions are completely determined by their descent modulus and corresponding critical values. This result is already surprising in the smooth case: a one-dimensional information (norm of the gradient) turns out to be almost as powerful as the knowledge of the full gradient mapping. In the nonsmooth case, the key element for this determination result is the *break of symmetry* induced by a downhill orientation, in the spirit of the definition of the metric slope.

Acknowledgements

Research supported by the FWF Grant P 36344-N (*Unilaterality and asymmetry in Variational Analysis*)

References

[1] Tahar Boulmezaoud, Philippe Cieutat, Aris Daniilidis, Gradient flows, second-order gradient systems and convexity, *SIAM J. Optim.* **28** (2018), 2049–2066.

[2] Aris Daniilidis, David Salas Determination of functions by metric slopes *Proc. Amer. Math. Soc.* **150** (2022), 4325–4333.

[3] Aris Daniilidis, Dmitriy Drusvyatskiy, The slope robustly determines convex functions *Proc. Amer. Math. Soc.* **151** (2023), 4751–4756.

[4] Aris Daniilidis, Laurent Miclo, David Salas, Descent modulus and applications `hal-03861013` (Preprint 45p.)

Some questions related to geometric inverse problems

Anna Doubova
doubova@us.es Universidad de Sevilla, Dpto. EDAN and IMUS, Spain

This conference delves into geometric inverse problems for partial differential equations. Our focus lies on determining a portion of the domain where the equations hold true, based on external measurements.

We will consider real-world applications and will explore two crucial aspects: uniqueness and numerical reconstruction. First, we will investigate how initial and boundary data influence solution's uniqueness. Among others, we will deal with one-dimensional inverse problems for the Burgers equation and related nonlinear systems, where heat effects, non-constant density and fluid-solid interaction are taken into account. The goal is to determine the size of the spatial interval based on specific boundary observations of the solution. We will explore both analytical and numerical solutions to these problems, employing powerful tools like Carleman estimates and insights from existing research (see [5], [4]). Additionally, we will provide methods to approximate the interval sizes. (This work is a collaboration with J. Apraiz, E. Fernández-Cara and M. Yamamoto [1]).

Next, some purely numerical approaches will be presented. We will showcase recent advancements and open questions concerning the reconstruction of the unknown domain. This involves solving a carefully chosen optimization problem (see [2]). We will also explore meshless technique based on the method of fundamental solutions (see [3]), demonstrating its effectiveness in tackling problems in both 2D and 3D settings.

Acknowledgements

The author was partially supported by MICINN, under Grant PID2020-114976GB-I00.

References

[1] J. Apraiz, A. Doubova, E. Fernández-Cara, M. Yamamoto. Some inverse problems for the Burgers equation and related systems. *Communications in Nonlinear Science and Numerical Simulation* 107, 106113, 2022.

[2] M. Kern. Numerical Methods for Inverse Problems. *New York, Wiley*, 2016.

[3] V.D. Kupradze, M.A. Aleksidze. The method of functional equations for the approximate solution of certain boundary value problems. *USSR Computational Mathematics and Mathematical Physics*, 4, 82–126, 1964.

[4] X. Huang, O. Imanuvilov, M. Yamamoto. Stability for inverse source problems by Carleman estimates. *Inverse Problems*, 36(12), 125006, 2020.

[5] M. Yamamoto. Carleman estimates for parabolic equations and applications. *Inverse Problems*, 25(12), 123013, 75 pp, 2009.

Robust optimization of uncertain multiobjective problems

Gabriele Eichfelder
gabriele.eichfelder@tu-ilmenau.de Technische Universität Ilmenau, Germany

1. Introduction

In many applications one has to deal with various difficulties at the same time, like uncertain data or several competing objective functions. For instance, for the integration of neighborhood networks into overarching distributing energy networks, a pure optimization of the neighborhood networks under an externally defined weighting of the relevant targets does not adequately model the problem. Moreover, uncertainties in the form of fluctuations or other disturbances can appear and a found solution has to be robust against that. A robust approach to uncertain multiobjective optimization corresponds to solving a set-valued optimization problem. However, it is a very difficult task to solve these optimization problems even for specific cases.

In this talk, after giving an introduction to the topic, we present a parametric multiobjective optimization problem for which the optimal solutions are strongly related to the robust solutions of the uncertain multiobjective problem. With this approach we can approximate the robust solution set with desired accuracy.

2. The uncertain multiobjective problem

We study uncertain multiobjective optimization problems under the following assumptions: Let $\Omega \subseteq \mathbb{R}^n$ and $\mathcal{U} \subseteq \mathbb{R}^k$ be nonempty and closed sets, with $\mathcal{U}$ compact. The set $\mathcal{U}$ is called the uncertainty set. Let $f : \Omega \times \mathcal{U} \to \mathbb{R}^m$ be a given continuous vector-valued objective function, i.e., we have m competing objective functions which have to be optimized at the same time.

The uncertain multiobjective optimization problem is then defined by

$$\left\{ \begin{array}{l} \min\limits_{x} \ f(x,u) \\ \text{s.t.} \ \ x \in \Omega \end{array} \ \middle| \ u \in \mathcal{U} \right\}, \tag{$\mathcal{UP}$}$$

i.e., ($\mathcal{UP}$) is formally represented as a parametric family of multiobjective optimization problems.

In the single-objective case, i.e., for $m = 1$, the classical robust counterpart problem in the sense of Ben-Tal and Nemirowski is given by

$$\min_{x \in \Omega} \ \sup_{u \in \mathcal{U}} f(x,u),$$

with the well-known epigraphical reformulation

$$\begin{array}{ll} \min\limits_{x,y} & y \\ \text{s.t.} & f(x,u) \leq y \quad \text{for all} \ \ u \in \mathcal{U}, \\ & x \in \Omega, \ y \in \mathbb{R}. \end{array}$$

For the uncertain multiobjective problem, following these ideas from the single-objective case, we obtain the so-called min-max set-based approach and thus consider the associated set-valued mapping $F_{\mathcal{U}} : \Omega \rightrightarrows \mathbb{R}^m$ defined as $F_{\mathcal{U}}(x) := \{f(x,u) \mid u \in \mathcal{U}\}$. For every $x \in \Omega$, the set $F_{\mathcal{U}}(x)$ represents the set of all possible vector costs associated to the decision x. The robust counterpart of ($\mathcal{UP}$) is then the set optimization problem

$$\min_{x \in \Omega} \ F_{\mathcal{U}}(x). \tag{$\mathcal{RC}$}$$

Based on [2], we present a generalization of the epigraphical reformulation for the uncertain multiobjective problem ($\mathcal{UP}$). This will be a semi-infinite multiobjective problem which can then be studied with classical techniques from the literature. For obtaining this reformulation we use vectorization results developed in [1] to characterize the optimal solutions of the set optimization problem ($\mathcal{RC}$).

Acknowledgements

This research is funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) under grant 03EI4013B and is a joint work with Ernest Quintana, Technische Universität Ilmenau, Germany.

References

[1] Gabriele Eichfelder, Ernest Quintana, Stefan Rocktäschel. A Vectorization Scheme for Nonconvex Set Optimization Problems. *SIAM Journal on Optimization*, 32(2), 1184–1209, 2022.

[2] Gabriele Eichfelder, Ernest Quintana. Set-based Robust Optimization of Uncertain Multiobjective Problems via Epigraphical Reformulations. *European Journal of Operational Research*, 313(3), 871–882, 2024.

Global solution of quadratic programs with binary variables and extensions

Sourour Elloumi
sourour.elloumi@ensta-paris.fr UMA, ENSTA Paris, Institut Polytechnique de Paris, and CEDRIC, Conservatoire National des Arts et Métiers, 75003 Paris, France

A quadratic program in binary variables has very special properties. It can be linearized, i.e. reformulated into a linear program in binary variables, by increasing the number of variables. It can also be convexified, i.e. reformulated into a convex quadratic program in binary variables, by a simple calculation of an extreme eigenvalue. Whether linearized or convexified, the reformulated problem can then be solved by a branch-and-bound based on continuous relaxation. These two paradigms have been known for several decades, but neither of them can efficiently achieve the optimum, outside of small or low-density instances. In a series of works starting from [1, 2, 4], we show that Quadratic Convex Reformulation allows the two paradigms to be put together to strengthen both of them and allow the global resolution of much larger instances. In these reformulations we seek a Mixed-Integer quadratic problem, equivalent to the first one, but with the additional property that its continuous relaxation is a convex problem. Furthermore, we want the bound by continuous relaxation of the new problem to be as tight as possible. We show that this reformulation is acheavable from the solution of a semidefinite programming problem.

We then extend this approach to the case of polynomial (or multilinear) optimization with binary variables in [5]. We first use a usual way of transforming the multilinear to an extended quadratic reformulation. Then, we use specific quadratic convex reformulation to this quadratic problem.

We also use similar ideas in [3] for mixed-integer quadratically constrained problems with general integer and bounded variables.

References

[1] Alain Billionnet and Sourour Elloumi. Using a mixed integer quadratic programming solver for the unconstrained quadratic 0–1 problem. *Mathematical Programming*, 109(1):55–68, 2007.

[2] Alain Billionnet, Sourour Elloumi, and Amélie Lambert. Extending the QCR method to general mixed-integer programs. *Mathematical programming*, 131(1–2):381–401, 2012.

[3] Alain Billionnet, Sourour Elloumi, and Amélie Lambert. Exact quadratic convex reformulations of mixed-integer quadratically constrained problems. *Mathematical Programming*, 158(1–2):235–266, 2016.

[4] Alain Billionnet, Sourour Elloumi, and Marie-Christine Plateau. Improving the performance of standard solvers for quadratic 0-1 programs by a tight convex reformulation: The qcr method. *Discrete Applied Mathematics*, 157(6):1185–1197, 2009.

[5] Sourour Elloumi, Amélie Lambert, and Arnaud Lazare. Solving unconstrained 0–1 polynomial programs through quadratic convex reformulation. *Journal of Global Optimization*, 80(2):231–248, 2021.

Stochastic Approximation beyond gradient

Gersende Fort
gersende.fort@math.univ-toulouse.fr CNRS, Institut de Mathématiques de Toulouse, France

In Machine Learning, many methods rely on Optimization, including its stochastic versions introduced for example to tackle non closed form expressions of the objective function or to reduce the computational cost. In 1951, H. Robbins and S. Monro introduced the method named "Stochastic Approximation" which is a root-finding method when the objective function is defined by an intractable expectation: it defines a sequence of iterates by using a Monte Carlo approximation of the expectation. Then, this method was generalized to solve a root-finding problem when only stochastic oracles of the objective field are available.

Stochastic Gradient algorithms are the most popular examples of the class of Stochastic Approximation methods. Nevertheless, Stochastic Approximation also contains far more general algorithms said "beyond gradient" since roughly, they consist in solving a minimization problem by using a vector field which is not a gradient field. These "beyong gradient" Stochastic Approximation methods often come with the additional difficulty that the stochastic oracles are biased approximations of the vector field. They occur in Computational Statistics (for example, some stochastic versions of Expectation Maximization are an instance of this beyond gradient case) and in Machine Learning as well (for example, some Temporal Difference algorithms for the estimation of the value function in Reinforcement Learning, are another instance).

This talk will first detail examples of such Stochastic Approximation methods beyond gradient. We will then present how a general theory can be derived, general enough to encompass as many as possible instances of Stochastic Approximation: we will emphasize the theory devoted to finite time analysis and will discuss how to choose design parameters of the algorithm in order to reach an epsilon-stationary point. We will finally show how to improve the original Stochastic Approximation scheme by plugging a variance reduction technique .

Approximation of binary second order cone programs of packing type

Max Klimm[1], **Marc E. Pfetsch**[2], **Rico Raber**[1], **Martin Skutella**[1]

1. klimm@math.tu-berlin.de Technische Universität Berlin, Germany
2. pfetsch@mathematik.tu-darmstadt.de Technische Universität Darmstadt, Germany
3. raber@math.tu-berlin.de Technische Universität Berlin, Germany
4. skutella@math.tu-berlin.de Technische Universität Berlin, Germany

We consider binary second order cone programs of packing type where a linear objective is optimized under m second order cone packing constraints and all decision variables are binary. Specifically, we consider mathematical programs of the form

$$\begin{aligned} \text{maximize} \quad & p^\top x \\ \text{subject to} \quad & c_k^\top x + \left\| A_k x + b_k \right\| \leq d_k \quad \text{for all } k \in \{1, \dots, m\}, \\ & x \in \{0,1\}^n, \end{aligned}$$

where $p \in \mathbb{Q}^n_{\geq 0}$, and for each $k \in \{1, \dots, m\}$ we have $A_k \in \mathbb{Q}^{\ell_k \times n}$ and $b_k \in \mathbb{Q}^{\ell_k}_{\geq 0}$ for some $\ell_k \in \mathbb{N}$ as well as $c_k \in \mathbb{Q}^n_{\geq 0}$ and $d_k \in \mathbb{Q}_{\geq 0}$. We require that $A_k^\top A_k$ and $A_k^\top b_k$ have only non-negative entries for all $k \in \{1, \dots, m\}$.

We show that when m is part of the input, these problems cannot be approximated within a factor of $1/(m+1)^{1-\varepsilon}$ for any $\varepsilon > 0$, unless P = NP. We then propose approximation algorithms based on different algorithmic principles that almost match this approximation factor: a pipage rounding technique that solves fractional relaxations of the problems and modifies the solutions so that few fractional variables remain, a greedy approach, and a randomized rounding technique. While all algorithms have similar theoretical approximation guarantees in the order of $1/m$, we also test the algorithms on realistic instances that arise in the context of gas transportation networks. This empirical study reveals in particular that taking the best of the proposed algorithms produces highly competitive solutions that yield on average 96 % of the value of an optimal solution.

Acknowledgements

We acknowledge funding through the DFG CRC/TRR 154, Subproject A07.

Solving the Hamilton Jacobi Bellman equation of optimal control: towards taming the curse of dimensionality

Karl Kunisch

karl.kunisch@uni-graz.at *University of Graz and Radon Institute, Austrian Academy of Sciences*

I consider optimal control problems on an infinite time horizon of the form

$$\begin{cases} \min\limits_{u(\cdot)\in U_{ad}} J(u(\cdot),x) := \int_0^\infty \ell(y(t)) + \frac{\gamma}{2}|u(t)|^2\, dt \\ \text{subject to} \quad \dot{y}(t) = f(y(t)) + Bu(t)\,, \quad y(0) = x. \end{cases} \tag{P}$$

It is assumed that $\ell(0) = f(0) = 0$ and thus solving (P) is intimately related to optimal stabilisation of the dynamical system $\dot{y}(t) = f(y(t)$ by means of the control u. As it stands, (P), at first, might be viewed as a special case of the abstract optimization problem

$$\min\ f(z) \quad \text{subject to} \quad e(z) = 0,$$

where we consider the time-dependent control-state pair (y, u) as the variable z with respect to which this constrained optimization problem needs to be solved. As a result we would get the optimal control u^* as a function of time.

Here, however, we are interested in a closed loop representation of the optimal control, i.e. we aim for expressing the optimal solution in feedback form as a function of the state $u^*(t) = u^*(y(t))$. In engineering practice this is motivated by the inherent stability properties of the closed loop approach.

Obtaining the optimal control in feedback form requires knowledge of the optimal value function

$$V(x) := \min_{u(\cdot)\in U_{ad}} J(u(\cdot), x),$$

which, in a viscosity sense is a solution to the Hamilton-Jacobi-Bellman (HJB) equation

$$\min_{u\in U_{ad}} \{\nabla V(x)^\top (f(x) + Bu) + \ell(x) + \frac{\gamma}{2}|u|^2\} = 0\,, \quad V(0) = 0, \quad \nabla V(0) = 0.$$

Once V is available, the optimal control in feedback from is given by $u(t) = -B^* grad V(y(t))$. The HJB equation is a hyperbolic equation and its dimension is that of the state-space of the nonlinear system in (P) ! Thus solving the HJB equation is a formidable task since one is confronted with a curse of dimensionality.

I give a brief survey of current solution strategies to partially cope with this challenge.

Subsequently I describe two approaches in some detail. The first one is a data driven technique, which approximates the solution to the HJB equation and its gradient from an ensemble of open loop solves.

The second technique also circumvents the direct solution of the HJB equation. It is based on a succinctly chosen learning ansatz, with subsequent approximation of the feedback gains by neural networks or polynomial basis functions.

Acknowledgements

This work relies on collaborations with B.Azmi (Uni-Konstanz), S.Dolgov (University of Bath), D.Kalise(Imperial College), D.Vasquez-Varas (Ricam Austrian Academy Sciences) , and D.Walter (Humbold Universität Berlin).

References

[1] Bezhad Azmi, Dante Kalise, and Karl Kunisch. Optimal feedback law recovery by gradient-augmented sparse polynomial regression. *Journal of Machine Learning Research*, 22(48):1–32, 2021. doi:http://jmlr.org/papers/v22/20-755.html.

[2] Sergey Dolgov, Dante Kalise, and Karl Kunisch. Tensor decompositions for high-dimensional Hamilton-Jacobi-Bellman equations. *SIAM Journal on Scientific Computing*, 43(3):A1625–A1650, 2021. doi:10.1137/19M1305136.

[3] Karl Kunisch and Donato Vásquez-Varas. Optimal polynomial feedback laws for finite horizon control problems. *Computers & Mathematics with Applications*, 148:113–125, 2023. doi:10.1016/j.camwa.2023.08.004.

[4] Karl Kunisch and Donato Vásquez-Varas. Smooth approximation of feedback laws for infinite horizon control problems with non-smooth value functions. submitted, 2024.

[5] Karl Kunisch, Donato Vásquez-Varas, and Daniel Walter. Learning optimal feedback operators and their sparse polynomial approximations. *Journal of Machine Learning Research*, 24(301):1–38, 2023. URL: https://jmlr.org/papers/v24/22-1010.html.

[6] Karl Kunisch and Daniel Walter. Semiglobal optimal feedback stabilization of autonomous systems via deep neural network approximation. *ESAIM: COCV*, 27:16, 2021. doi:10.1051/cocv/2021009.

[7] Karl Kunisch and Daniel Walter. Optimal feedback control of dynamical systems via value-function approximation. *Comptes Rendus Mécanique*, 351(S1):1–37, 2023. doi:10.5802/crmeca.199.

On five types of Voronoi diagrams

Juan Enrique Martínez-Legaz

juanenrique.martinez.legaz@uab.cat Universitat Autònoma de Barcelona, Spain

In addition to the classical Voronoi diagrams in $\mathbb{R}^n$, we delve into other variants: higher-order cells, farthest cells based on both Euclidean and Bregman distances, and power cells. Despite their diversity, all these Voronoi diagrams share a common feature: their cells are closed convex sets.

For the classical case, we study the Voronoi inverse problem. Given a closed convex set $F \subseteq \mathbb{R}^n$ and a point $s \in F$, we study the family of sets $T \subseteq \mathbb{R}^n$ that contain s and satisfy the condition that the corresponding Voronoi cell $V_T(s)$ of s with respect to T is precisely F. We explore the relationship between the elements of this family and the linear representations of F. Additionally, we provide explicit formulas for maximal and minimal elements within this family. Our investigation also involves studying the closure operator that assigns to each closed set T containing s the largest set $\mathbb{F} \subseteq \mathbb{R}^n$ containing s such that $V_{\mathbb{F}}(s) = V_T(s)$.

Higher order Voronoi cells associate to each k-element subset of the set of sites the set of points for which the given subset consists of the k closest sites. We delve into the structure of k-order Voronoi cells and illustrate our theoretical findings through a case study involving two-dimensional higher-order Voronoi cells for four points.

The farthest Voronoi cell of a point s, denoted as $F_T(s)$, comsits of all the points farther from s than from any other site. We explore farthest Voronoi cells and diagrams corresponding to arbitrary (possibly infinite) sets. Specifically, for a given arbitrary set T, we characterize those $s \in T$ such that $F_T(s)$ is nonempty and analyze the geometrical properties of $F_T(s)$. Additionally, we characterize those sets T whose farthest Voronoi diagrams form tessellations of the Euclidean space, as well as those that can be expressed as $F_T(s)$ for some $T \subseteq \mathbb{R}^n$ and some $s \in T$. Some of the results obtained can be extended to the case when, instead of the Euclidean distance, the Bregman distance associated with a differentiable strictly convex function is considered.

Finally, we study power cells. Given a set $T \subseteq \mathbb{R}^n$ and a nonnegative function r defined on T, the power of $x \in \mathbb{R}^n$ with respect to the sphere with center $t \in T$ and radius $r(t)$ is $p_r(x,t) := \|x-t\|^2 - r^2(t)$, with $\|\cdot\|$ denoting the Euclidean distance. The corresponding power cell of $s \in T$ is the set

$$C_T^r(s) := \{x \in \mathbb{R}^n : p_r(x,s) \leq p_r(x,t), \text{ for all } t \in T\}.$$

We investigate the structure of these power cells and explore assumptions on r that allow for generalizing known results on classical Voronoi cells.

The results presented in this talk are contained in joint works with E. Allevi, M. A. Goberna, E. Naraghirad, R. Riccardi, V. Roshchina, M. Tamadoni Jahromi, M. I. Todorov and V. N. Vera de Serio.

Acknowledgement

The author was partially supported by Grant PID2022-136399NB-C22 from MICINN, Spain, and ERDF, "A way to make Europe," European Union.

References

[1] E. Allevi, J. E. Martínez-Legaz, R. Riccardi. On the basic properties and the structure of power cells. *J. Optim. Theory Appl.*, to appear.

[2] M. A. Goberna, J. E. Martínez-Legaz, M. I. Todorov. On farthest Voronoi cells. *Linear Algebra Appl.*, 583:306–322 (2019).

[3] M. A. Goberna, J. E. Martínez-Legaz, V. N. Vera de Serio. The Voronoi inverse mapping. *Linear Algebra Appl.*, 504:248–271 (2016).

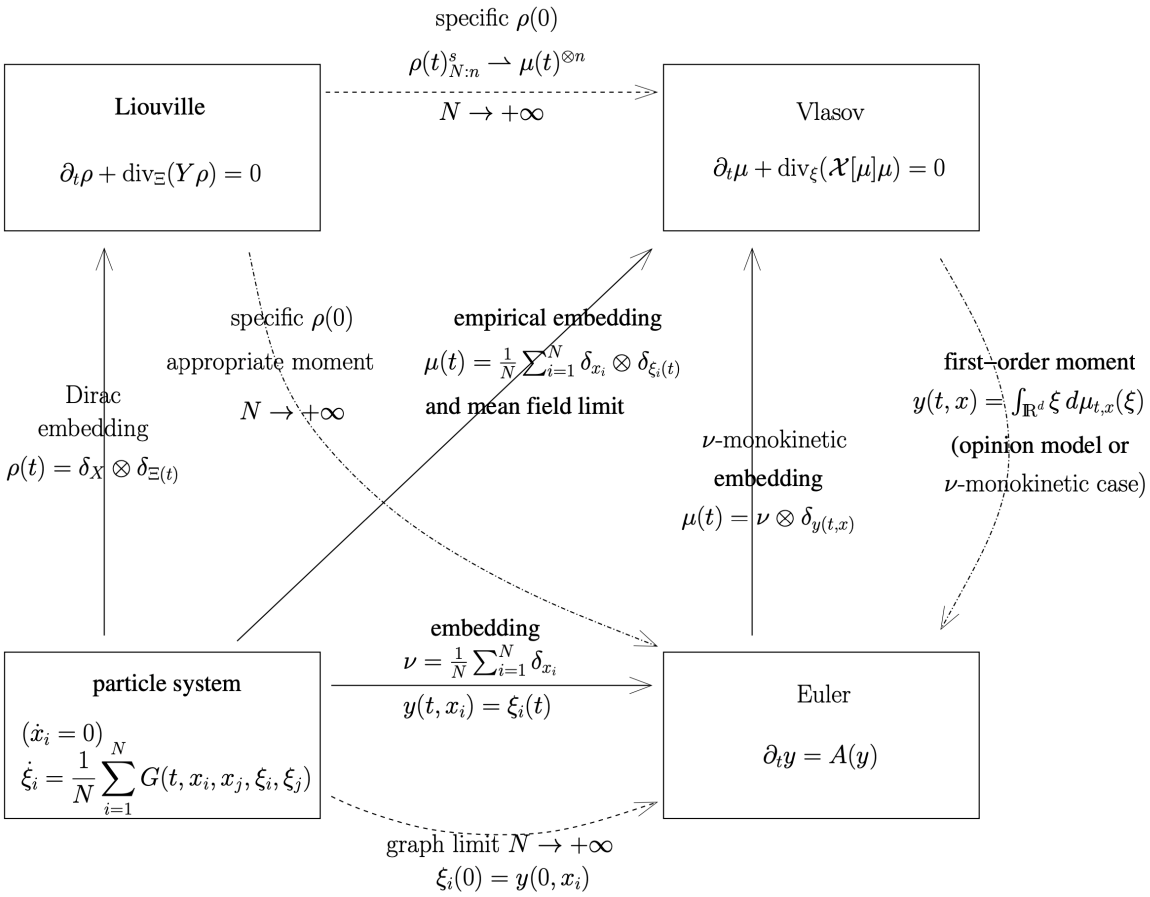
[4] J. E. Martínez-Legaz, V. Roshchina, M. I. Todorov. On the structure of higher order Voronoi cells. *J. Optim. Theory Appl.*, 183:24–49 (2019).

[5] J. E. Martínez-Legaz, M. Tamadoni Jahromi, E. Naraghirad. On farthest Bregman Voronoi cells. *Optimization*, 71:937-947 (2022).

From microscopic to macroscopic scale equations: mean field, hydrodynamic and graph limits

Emmanuel Trélat

emmanuel.trelat@sorbonne-universite.fr Sorbonne Université, Paris, France



Considering finite particle systems, we elaborate on various ways to pass to the limit as the number of agents tends to infinity, either by mean field limit, deriving the Vlasov equation, or by hydrodynamic or graph limit, obtaining the Euler equation. We provide convergence estimates. We also show how to pass from Liouville to Vlasov or to Euler by taking adequate moments. Our results encompass and generalize a number of known results of the literature.

As a surprising consequence of our analysis, we show that sufficiently regular solutions of any quasilinear PDE can be approximated by solutions of systems of N particles, to within $1/\log\log(N)$.

Control and Machine Learning

Enrique Zuazua[1,2,3]

1. enrique.zuazua@fau.de Chair for Dynamics, Control, Machine Learning and Numerics, Alexander von Humboldt Professorship, Department of Mathematics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany
2. Chair of Computational Mathematics, Fundación Deusto, Av. de las Universidades, 24, 48007 Bilbao, Basque Country, Spain
3. Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid - Spain

In this lecture we shall present some recent results of our group on the interplay between control and Machine Learning, and more precisely, Supervised Learning and Universal Approximation.

We adopt the perspective of the simultaneous or ensemble control of systems of Residual Neural Networks (ResNets). Roughly, each item to be classified corresponds to a different initial datum for the Cauchy problem of the ResNets, leading to an ensemble of solutions to be driven to the corresponding targets, associated to the labels, by means of the same control.

We present a genuinely nonlinear and constructive method, allowing to show that such an ambitious goal can be achieved, estimating the complexity of the control strategies.

This property is rarely fulfilled by the classical dynamical systems in Mechanics and the very nonlinear nature of the activation function governing the ResNet dynamics plays a determinant role. It allows deforming half of the phase space while the other half remains invariant, a property that classical models in mechanics do not fulfill.

This viewpoint opens up interesting perspectives to develop new hybrid mechanics-data driven modelling methodologies. We shall discuss a number of promising open problems.

Abstracts

Mixed-integer non-linear programming approach for identifying parameters inconsistencies in load flow calculations

Pierre Arvy
pierre.arvy@artelys.com Artelys España, Spain

For several decades, electrical networks have experienced continuous expansion in size and scale, facilitating the transmission of electrical power from sources to end-users. The growing complexity of these networks necessitates the use of tools like Load Flow (LF) calculations, which are crucial for power systems engineering [1], playing a critical role in ensuring grid reliability and efficiency. This numerical analysis of electric power within interconnected systems determines steady-state voltages, currents, and power flows across network components.

Given the significance of LF calculations, the power systems community has invested substantial efforts in enhancing the convergence of the methods used to solve them [4], including the associated optimization problem known as Optimal Power Flow [2]. Despite these efforts, real-world scenarios, especially in power system planning, often involve imperfect or erroneous network parameters, leading to challenging convergence issues for existing methods [3]. The physical problem studied may be mathematically infeasible due to complex inconsistencies in the parameters, yet the divergence of methods is not a clear certificate of this. Consequently, power system engineers usually engage in iterative processes to pinpoint the moment at which divergence occurs and identify the responsible parameters, making it a time-consuming endeavor.

This research project explores a novel approach centered around Mixed-Integer Non-Linear Programming (MINLP) to detect inconsistencies or errors in the parameters of a power network, on which LF calculations are applied and divergent. The MINLP is obtained by utilizing the non-linear LF equations as constraints and transforming the various network parameters involved in these equations into penalized variables. This transformation enables parameter adjustments -or corrections- while the MINLP minimizes them using binary variables. This introduces additional variables and constraints into the optimization problem, following a conventional form in operations research. Solving the MINLP identifies inconsistencies in the power network's parameters, corresponding to the adjusted values.

The study extends with the implementation of an open-source tool (available on GitHub here), integrated with the PowSyBl Java framework, widely used in the power systems community for LF analyses. It relies on an implementation of the previously discussed MINLP using the AMPL modeling language, and conducts a solution to this non-linear and non-convex optimization problem using the commercial solver Artelys Knitro. Interior-point algorithms are employed for solving the non-linear relaxation of the problem, along with the non-linear branch-and-bound algorithm of the solver for handling the non-continuous problem.

Finally, the research project outlines the validation procedure of the implemented tool. Validation through the detection of inconsistencies in power network parameters is based on unit tests manually generated from both open-source and real-world use cases. Besides, the tool successfully identifies inconsistencies in an operational scenario involving a real-world case with parameter inconsistencies. The average computation time of the tool is also analyzed, demonstrating its practical usability by power systems operators. Furthermore, based on observations and results obtained, the study presents future research directions to enhance the detection of parameter inconsistencies in power networks. These directions include using other formulations of the objective function of the presented MINLP, exploring alternative load flow formulations, and integrating validated physical inequalities as constraints.

References

[1] Olukayode Afolabi, Warsame Ali, P. Cofie, John Fuller, Pamela Obiomon, and Emmanuel Kolawole. Analysis of the load flow problem in power system planning studies. *Energy and Power Engineering*, 07:509–523, 01 2015. doi:10.4236/epe.2015.710048.

[2] Stephen M. Frank, Ingrida Steponavice, and Steffen Rebennack. Optimal power flow: a bibliographic survey i. *Energy Systems*, 3:221 – 258, 2012. doi:10.1007/s12667-012-0056-y.

[3] Jigar Patel, Daivat Desai, Vaibhav Patel, Dishang D. Trivedi, and Santosh C. Vora. Analyzing effect of bad measurement data on load flow and state estimation in power system. In *2015 5th Nirma University International Conference on Engineering (NUiCONE)*, pages 1–6, 2015. doi:10.1109/NUICONE.2015.7449618.

[4] B. Stott. Review of load-flow calculation methods. *Proceedings of the IEEE*, 62(7):916–929, 1974. doi:10.1109/PROC.1974.9544.

On the solution of shape optimization for the Navier-Stokes problem with the stick-slip condition

Vladimír Arzt[1], **Petr Beremlijski**[2]

1. vladimir.arzt@vsb.cz VSB-Technical University of Ostrava, Czechia
2. petr.beremlijski@vsb.cz VSB-Technical University of Ostrava, Czechia

1. Introduction

The shape optimization for the Navier-Stokes problem with the stick-slip boundary condition can be modelled as a bilevel problem, where the upper-level optimization task is described as the minimization of a given objective function with respect to the control variables that control the shape of the boundary (the objective function also depends on state variables). The lower-level optimization task is the Navier-Stokes problem with the boundary condition for a given boundary shape described by the control variable. Its solution is the state variables (velocity and pressure). The shape optimization problem can also be modelled as the minimization of the composite function generated by the objective and the control-state mapping. It can be shown that this composite function is generally non-differentiable (nonsmooth).

2. Mathematical problem

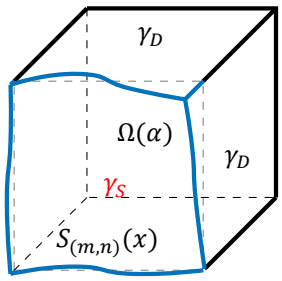
To solve the nonsmooth optimization problem (2.1), we have to use methods that work with the calculus of Clarke.

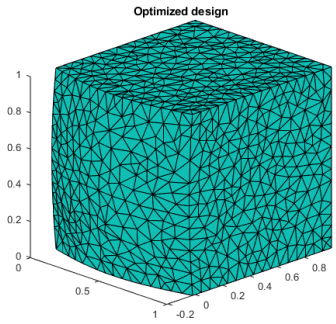
$$\begin{cases} \text{Find } \alpha^* \in \mathcal{U} \text{ such that } \forall \alpha \in \mathcal{U} \\ \mathcal{J}(\alpha^*, (\mathbf{u}^{\alpha^*}, p^{\alpha^*})) \leq \mathcal{J}(\alpha, (\mathbf{u}^{\alpha}, p^{\alpha})) \end{cases} \tag{2.1}$$

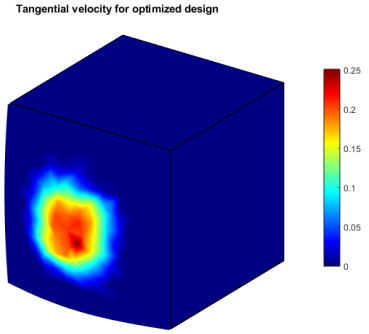
First, we use the bundle trust method proposed by Schramm and Zowe. In each step of the iteration process, we must find the solution of the state problem, i.e., the Navier-Stokes problem with the stick-slip condition, and compute one arbitrary Clarke subgradient of the nonsmooth composite function. Finite differences are used for the approximation of this subgradient. Second, we use the limited memory discrete gradient bundle method, where the discrete gradient is used instead of finite differences.

The state Navier-Stokes problem is approximated by FEM P1-bubble/P1 pair and contains two nonlinearities. The first caused by the convective term is linearized by Oseen or Newton iterations (see [3]), and the second caused by the nonlinear stick-slip condition is solved by a semi-smooth Newton method based on active/inactive sets with preconditioned BiCGstab as an inner solver (see [1]).

As one of the shape optimization problems, we look for a shape Ω such that the tangential velocity on γ_S fits the function $\mathbf{u}_e(x, y)$ (for more details see [2]).







Acknowledgements

The authors were supported by the European Union under the REFRESH – Research Excellence For REgion Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition.

References

[1] Vladimír Arzt. Algorithms for flow problems with stick-slip boundary conditions in three space dimensions. Master's thesis, VŠB-TU Ostrava, Czech Republic, 2021. url: https://dspace.vsb.cz/bitstream/handle/10084/143852/ARZ0009_FEI_N0541A170007_S02_2021.pdf?sequence=1&isAllowed=y [online].

[2] Jaroslav Haslinger, Raino Mäkinen, and Jan Stebel. Shape optimization for stokes problem with threshold slip boundary conditions. *Discrete and Continuous Dynamical Systems - Series S*, 10:1281–1301, 12 2017. doi:10.3934/dcdss.2017069.

[3] Radek Kučera, Vladimír Arzt, and Jonas Koko. Mini element for the navier-stokes system in 3d: vectorized codes and superconvergence. *Informatica*, accepted.

Control and estimation for the design of a smart electrostimulator using Ding et al model

Toufik Bakir[1], **Bernard Bonnard**[2]

1. *toufik.bakir@u-bourgogne.fr Laboratoire ImViA, Université de Bourgogne, France*
2. *bernard.bonnard@u-bourgogne.fr Laboratoire IMB, Université de Bourgogne, France*

In this article, we present the design of a smart electrostimulator for muscle rehabilitation or reinforcement, using fast computations, in order to control the muscular force or to regulate to a force level. The Ding and al. model is used to predict and to optimize the muscular force response to functional electrical stimulation [4]. We analyze the estimation of the Ding and al. parameters using an approximation of the force response [1] which depends upon the 6 parameters of the Ding's model and we derive optimization scheme which bypass the time computational expensive integration of the dynamics of the Ding and al. equations. Two of the parameters are depending upon the fatigue and we present a numerical sensor for their estimation.

References

[1] Toufik Bakir, Bernard Bonnard, Sandrine Gayrard, Jeremy Rouot. Finite Dimensional Approximation to Muscular Response in Force-Fatigue Dynamics using Functional Electrical Stimulation. *Automatica*, 2022.

[2] Toufik Bakir, Bernard Bonnard, Loïc Bourdin, Jeremy Rouot. Pontryagin-typeconditions for optimal muscular force response to functional electrical stimulations. *Journal of Optimization Theory and Applications*, pp. 184:581–602, 2020.

[3] Toufik Bakir, Bernard Bonnard, Loïc Bourdin, Jeremy Rouot. Direct and indirect methods to optimize the muscular force response to a pulse train of electrical stimulation. *ESAIM : Proceedings and Survey, EDP Sciences*, vol. 71, pp. 1–10, 2021.

[4] Jun Ding, Anthony S Wexler, and Stuart A Binder-Macleod.. Development of a mathematical model that predicts optimal muscle activation patterns by using brief trains. *Journal of applied physiology*, vol. 88, pp 917–925, 2000.

[5] Emma Wilson. Force response of locust skeletal muscle. Southampton University, Ph.D. thesis, 2021.

[6] Brian D Doll ,Nicholas A Kirsch, and Nitin Sharma. Optimization of a stimulation train based on a predictive model of muscle force and fatigue. *IFAC- PapersOnLine*, vol. 48(20), pp. 338-342, 2015.

[7] Sandrine Gayrard. Fast Optimization Scheme for the Muscular Response to FES Stimulation to Design a Smart Electrostimulator. Phd Thesis, Université de Bourgogne, 2023.

The chemostat system with mutation: asymptotic analysis, stabilization, and optimal control

Claudia Alvarez-Latuz[1], **Térence Bayen**[2], **Jérôme Coville**[3]

1. claudia.alvarez-latuz@univ-avignon.fr *Avignon Université, Laboratoire de Mathématiques d'Avignon (EA 2151), F-84018 Avignon, France.*
2. terence.bayen@univ-avignon.fr *Avignon Université, Laboratoire de Mathématiques d'Avignon (EA 2151), F-84018 Avignon, France.* Corresponding author.
3. jerome.coville@inrae.fr *UR 546 Biostatistique et Processus Spatiaux, INRAE, Domaine St Paul Site Agroparc, F-84000 Avignon, France.*

In this presentation, we consider the so-called chemostat system with mutation

$$\left|\begin{array}{ll} \dot{x}_i(t) & = (\mu_i(s(t)) - u)x_i + \varepsilon(Tx(t))_i, \qquad 1 \le i \le n, \\ \dot{s}(t) & = -\sum_{j=1}^n \mu_j(s(t))x_j(t) + u(s_{in} - s(t)). \end{array}\right. \tag{1}$$

This system describes the evolution of n species (with concentrations $x_i(t)$) that compete over a same resource (called substrate with concentration $s(t)$). Each species grows over the common resource according to its own kinetics μ_i which is a real-valued non-negative continuous functions over $\mathbb{R}_+$ satisfying $\mu_i(s) = 0$. In this system, u is a constant and denotes the dilution rate (but it also often taken as a control like in biotechnology) and s_{in} is the input substrate concentration (that can also be controlled in practice). The main new feature that interests us (w.r.t. the chemostat system [5]) is the possible mutation between species (like in gene tranfer) that is modeled in (1) via the so-called transition matrix T (quasi-positive irreducible) corresponding to these interactions. The parameter $\varepsilon > 0$ then corresponds to the mutation rate for these interactions between species. One key issue is to predict the asymptotic behavior as time $t \to +\infty$. This question appears to be central in the study of the optimal control of (1) given some objective function. Remind that without mutation effect, i.e., when $\varepsilon = 0$, the competitive exclusion principle predicts that, generically, only one species survives as $t \to +\infty$ since (1) reduces to the classical chemostat system (see [3, 5]).

Mutations between species in the chemostat model are important to model whenever $n \gg 1$. This leads to a system like (1), but more general mutation terms can also be considered. Mutation is encountered typically in cultured microalgae that serve for biofuels. Whenever mutation is taken into account, (1) is called *a regular perturbation* of the original chemostat system without mutation. However, the study of the asymptotic analysis of (1) is more involved than for the chemostat system : in general, even if $\varepsilon > 0$ is arbitrarily small, the asymptotic behavior of a regular perturbation of a dynamical system may differ a lot from the asymptotic behavior of the original system (without perturbation). In [1], it has been shown that (1) has a unique coexistence steady-state[1] globally asymptotically stable (apart the wash-out equilibrium corresponding to extinction of species, i.e., $x = 0$ and $s = s_{in}$) provided that the dilution rate u is small enough.

Here, we analyze another (and more realistic) issue corresponding to the case where the mutation rate ε is such that $\varepsilon \downarrow 0$. The goal of this presentation is to show new results toward this direction. Global stability of the coexistence steady-state is obtained provided that the mutation parameter is small enough. Our analysis is based on the Malkin-Gorshin Theorem (see [4]) and a global stability result related to regular perturbations of dynamical systems [6]. We shall also give preliminary results concerning a similar study for the PDE-based model when a pool of species is considered (corresponding to a trait) and also some results concerning the optimization of microbial production when mutation is taken into account and $t \mapsto u(t)$ is taken as a control function in (1). The results that are proposed in this presentation are based on undergoing works extending the recent articles [1, 2].

References

[1] T. Bayen, H. Cazenave-Lacroutz, J. Coville. *Stability of the chemostat system including a linear coupling between species.* Discrete and Continuous Dynamical Systems-B, 28(3), 2104–2129, 2023.

[2] T. Bayen, J. Coville, F. Mairet. *Stabilization of the chemostat system with mutations and application to microbial production.* Optimal Control Applications and Methods, 44(6), 3342-3360, 2023.

[3] J. Harmand, C. Lobry, A. Rapaport, T. Sari. *The Chemostat: Mathematical Theory of Microorganism Cultures.* Wiley-ISTE, 2017.

[4] T. Sari, B.S. Kalitin. *B-stability and its applications to the Tikhonov and Malkin-Gorshin theorems.* Differential Equations, Kluwer Academic Publishers-Plenum Publishers, Edinburgh, 37, 11–16, 1996.

[5] H.L. Smith, P. Waltman. *The theory of the chemostat: dynamics of microbial competition*. Vol. 13, Cambridge university press, 1995.

[6] H.L. Smith, P. Waltman. *Perturbation of a globally stable steady state.* Proc. Amer. Math. Soc., 127(2), 447–453, 1999.

[1] i.e., every species is present asymptotically

Newton-type methods for detecting robotic arm position using camera

Michaela Bailová[1], **Petr Beremlijski**[2]

1. michaela.bailova@vsb.cz VSB-Technical University of Ostrava, Czechia
2. petr.beremlijski@vsb.cz VSB-Technical University of Ostrava, Czechia

1. Introduction

This contribution introduces a digital twin of an industrial camera for the calibration of robotic manipulators in manufacturing processes. We approximate the camera's characteristics to facilitate precise calibration by using a full camera model (for more details, see to [1]). The main goal of the contribution is to present optimization methods for finding a camera's location based on the captured image of the calibration pattern. Such coordinates are then used to navigate the camera back to the prescribed position.

2. Mathematical problem

The calibration uses a predefined calibration square with five significant points, see Fig. 1, to locate the camera accurately in 3D space.

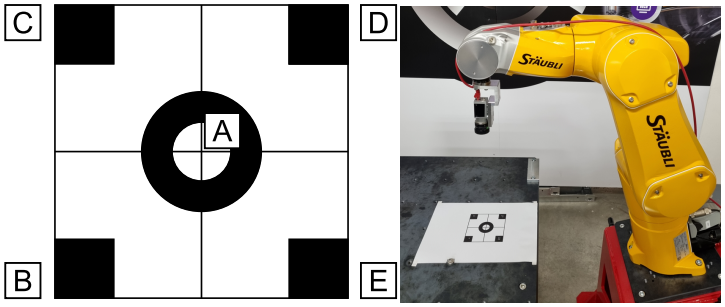


Fig. 1 Geometric pattern displayed by the camera and the robot.

The objective is to determine a position and orientation where the points computed by the camera model closely match the significant points in the captured image. This problem translates to an unconstrained quadratic programming problem

$$\min_{\vec{x}\in\mathbb{R}^6} f(\vec{x}) := \min_{\vec{x}\in\mathbb{R}^6} \sum_{i=1}^{10} (F_i(\vec{x}) - \overline{F}_i)^2.$$

Here, the mapping $F : \mathbb{R}^6 \to \mathbb{R}^{10}$ describes the function assigning the camera's position and orientation to the image of the calibration square on the camera display chip, as determined by the camera model. $\overline{F}$ represents the vector of prescribed x and y-coordinates of the calibration points.

This contribution presents and compares solutions to this industrial problem using Newton's method, hybrid Newton's methods, and a combination of neural networks and Newton's method.

While Newton's method offers local convergence, its limitation is requiring an initial iteration close to the solution. Thus, combining Newton's method with metaheuristic methods like differential evolution proves beneficial. Alternatively, approximating the inversion of the function F for a given image with a neural network and employing the resulting value of the inverse as the starting iteration for Newton's method is another possibility to solve the problem.

The main advantages of these methods are their robustness and speed.

Acknowledgements

The authors were supported by the European Union under the REFRESH – Research Excellence For REgion Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition.

References

[1] Jaromir Konecny, Petr Beremlijski, Michaela Bailova, Zdenek Machacek, Jiri Koziorek, and Michal Prauzek. Industrial camera model positioned on an effector for automated tool center point calibration. *Scientific Reports*, 14(1), 2024. doi:10.1038/s41598-023-51011-5.

Sharp Hoffman constant of the argmin mapping in the linear setup

Jesús Camacho[1], **María Josefa Cánovas**[2], **Helmut Gfrerer**[3], **Juan Parra**[4]

1. j.camacho@umh.es Miguel Hernández University of Elche, Spain
2. Miguel Hernández University of Elche, Spain
3. Johannes Kepler University, Austria
4. Miguel Hernández University of Elche, Spain

The main goal of this talk is to provide a point-based (only involving the problem's data) formula for the Hoffman constant of the argmin mapping in linear optimization, understood as the sharp Lipschitz constant restricted to its domain. The work is developed in the parametric context of right-hand side perturbations of the constraint systems. Historically, this constant has been of general interest to the scientific community, which is evidenced by the outstanding research on it. Although we can find some upper estimates of such quantity in the literature, we provide in this talk the first exact expression of it.

To achieve the main objective, we introduce new general tools of their own interest, as the concept of well-connected piecewise convex mappings. We isolate the nice behavior of such mappings to derive a crucial equality between the Hoffman constant (which is a global stability measures) and the supremum of calmness moduli (which are of local nature). Additionally, a directional stability approach will be introduced for computational purposes.

Sparsity and L^1-optimal control for linear control systems

Pierre Cavaré[1], **Marc Jungers**[2], **Jérôme Lohéac**[3]

1. pierre.cavare@univ-lorraine.fr Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France
2. marc.jungers@univ-lorraine.fr Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France
3. jerome.loheac@univ-lorraine.fr Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France

For an initial state $x^0 \in \mathbb{R}^n$, a target $x^1 \in \mathbb{R}^n$ and a final time $T > 0$ given, we consider the problem of state transfer for linear time-invariant systems whose plant model is given by

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ x(0) = x^0, \quad x(T) = x^1. \end{cases}$$

When x^1 is reachable from x^0 in time T, it is common to look for maximum hands-off controls [4, 8, 9], i.e. feasible control(s) whose support is of minimal Lebesgue measure. We call this problem (P_0). Here, we authorize controls in the set of Radon measures and, more precisely, impulsive controls [2, 7] whose support is of null Lebesgue measure. Thus, we can discredit the problem (P_0) by showing that there always exists a feasible impulsive control that we can approach with a sequence of feasible controls in $L^1([0,T],\mathbb{R}^m)$.

For optimization purposes, we want to minimize the weights of impulses of a such control and we show that this new problem is linked to the problem (P_1) of L^1-minimization for controls in $L^1([0,T],\mathbb{R}^m)$. In particular, optimal impulsive controls steer x^0 to x^1 in time T with a smaller (or, at least, the same) "L^1-cost" than optimal piecewise continuous controls. As with (P_0), there is no gap between these two problems and an optimal impulsive solution can be approached by a sequence of feasible controls of $L^1([0,T],\mathbb{R}^m)$. The main result of this talk is that there exists an impulsive solution which is a linear combination of at most nm Dirac impulses (where n and m are respectively the dimension of the state and of the control).

Then, by using the Pontryagin's maximum principle [1, 5], we show that any minimizer of (P_1) is necessarily of purely impulsive form as soon as A is nonsingular and all pairs (A, B_i) are controllable (where B_i is the i-th column of B).

The remainder of this talk is dedicated to the implementation of an algorithm to approach an optimal impulsive control. We build it by using the coordinate descent method [6] and Bregman iterations [3].

References

[1] V. Alexéev, S. Fomine, and V.M. Tikhomirov. *Optimal control*. Mir, Moscou, 1979.

[2] A. Bressan and F. Rampazzo. On differential systems with vector-valued impulsive controls. *Boll. Unione Mat. Ital., VII. Ser., B*, 2(3):641–656, 1988.

[3] J. Darbon, D. Goldfarb, S. Osher, and W. Yin. Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. *SIAM J. Imaging Sci.*, 1(1):143–168, 2008. doi:10.1137/070703983.

[4] T. Ikeda and K. Kashima. On sparse optimal control for general linear systems. *IEEE Trans. Autom. Control*, 64(5):2077–2083, 2019. doi:10.1109/TAC.2018.2863220.

[5] A.D. Ioffe and V.M. Tikhomirov. *Theory of Extremal Problems*. Russian edition: Nauka, Moscow. North-Holland Publishing Company, Amsterdam, 1979.

[6] Y. Li and S. Osher. Coordinate descent optimization for l^1 minimization with application to compressed sensing; a greedy algorithm. *Inverse Probl. Imaging*, 3(3):487–503, 2009. doi:10.3934/ipi.2009.3.487.

[7] J. Lohéac, E. Trélat, and E. Zuazua. Nonnegative control of finite-dimensional linear systems. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 38(2):301–346, 2021. doi:10.1016/j.anihpc.2020.07.004.

[8] M. Nagahara. Sparse control for continuous-time systems. *International Journal of Robust and Nonlinear Control*, 33(1):6–22, 2023. doi:10.1002/rnc.5858.

[9] M. Nagahara, D. Nešić, and D.E. Quevedo. Maximum hands-off control: a paradigm of control effort minimization. *IEEE Trans. Autom. Control*, 61(3):735–747, 2016. doi:10.1109/TAC.2015.2452831.

About the Bang-Bang Principle in a non-smooth setting

Ruben Chenevat[1], **Bruno Cheviron**[2], **Alain Rapaport**[3], **Sébastien Roux**[3]

1. ruben.chenevat@inrae.fr MISTEA, Univ. Montpellier, INRAE, France.
2. bruno.cheviron@inrae.fr G-EAU, INRAE, Montpellier, France
3. {alain.rapaport,sebastien.roux}@inrae.fr MISTEA, INRAE, Montpellier, France

1. Introduction

The Bang-Bang Principle is well established for non-autonomous linear time-optimal problems (see for example [3]). In this work, we revisit this result for problems with non-smooth dynamics, but continuous with respect to the state variable. In this context, we show that there exists an optimal trajectory composed of bang-bang arcs where the dynamics is differentiable, and singular arcs may occur at non-differentiable locus of the dynamics. Moreover we can characterize these singulars arcs using the Pontryagin Principle for non-smooth dynamics (see [2]).

2. Some details

More precisely we consider the following dynamics in a region D_i :

$$\dot{X}(t) = f(t, X(t), u(t)) = A_i(t)X(t) + B(t)u(t) \tag{2.1}$$

for some partition $\bigsqcup_{i\in\mathcal{I}} D_i$ of $\mathbb{R}^n$ with $\mathring{D}_i \neq \emptyset$, where A_i and B are $n\times n-$ and $n\times r-$ matrix-valued functions, and $u(\cdot)$ is a mesurable control taking values in a compact set of $\mathbb{R}^r$. We also assume that for all (t,u) the function $x \mapsto f(t,x,u)$ is continuous.
We consider a compact target set $C \subset \mathbb{R}^n$, and we investigate the minimal time problem to reach C from an initial state $X(0) = X_0 \in \mathbb{R}^n$ that belongs to the attainability set of C.

We shall show that there is an optimal solution of the minimal time problem subject to (2.1) composed of a succession of bang-bang arcs over time intervals where the state trajectory belongs to a region D_i, and possible singular arcs over time intervals where the state trajectory evolves in a boundary of a region D_j.

2.1. Example

We will highlight this result with an example coming from an optimization problem for an irrigation model with water scarcity (see [1]). In this example, we prove with the help of the non-smooth PMP that an optimal solution must contain a singular arc for a subset of initial conditions.

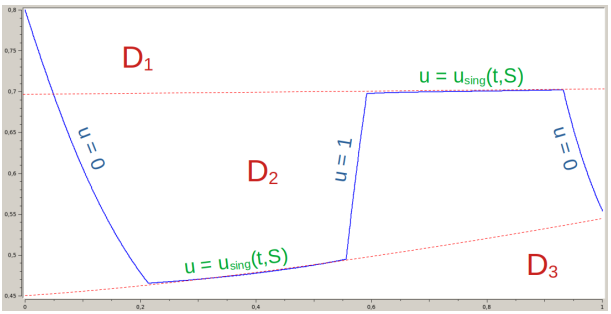


Fig. 1 Optimal "bang - singular - bang - singular - bang" trajectory for the soil moisture.

Acknowledgements

This work is supported by Labex Numev and #DigitAg Convergence Institute.

References

[1] K. Boumaza, N. Kalboussi, A. Rapaport, S. Roux, C. Sinfort. *Optimal control of a crop irrigation model under water scarcity*. Optimal Control Applications and Methods, 42(6) 1612-1631, 2021.

[2] F. Clarke. *Functional analysis, calculus of variation and optimal control*. Springer, 2013.

[3] H. Hermès, J. P. Lasalle. *Functional analysis and time optimal control*. Academic Press, Inc., 1969.

Robust optimal design of renewable energy systems

Elisa Corbean[1], **Stefan Ulbrich**[2]

1. corbean@mathematik.tu-darmstadt.de Technical University of Darmstadt, Germany
2. ulbrich@mathematik.tu-darmstadt.de Technical University of Darmstadt, Germany

Renewable energy systems are the basis for the current endeavor of eliminating CO_2 emissions and the expansion and construction of new systems is an indispensable action within this process. Therefore, finding the optimal design of renewable energy systems and related process components represents an interesting task that can be tackled by optimization techniques.

Many energy system descriptions based on renewable energy sources rely on location-specific weather data to account for the renewable energy potential at different sites. However, the optimal design of renewable energy systems should also be robust with respect to fluctuations of these meteorological parameters due to the inherent volatility of natural phenomena. Therefore, robust optimization techniques are employed to compute robust optimal system designs.

The source of complexity of such optimization problems is two-fold. On the one hand, depending on the level of detail of the process description, nonlinearities may arise. On the other hand, incorporating hourly resolved weather data over longer periods into the model description results in large-scale optimization problems. Therefore, finding efficient ways of computing the robust optimal system design is crucial.

In this work, we investigate the computation of robust optimal solutions for renewable energy system models of different levels of detail. Strategies for improving the efficient computation of the robust counterpart are based on time-series aggregation methods in order to reduce the model size of the system description.

Furthermore, the impact of incorporating knowledge gained from time-series data on the choice of the uncertainty set will be investigated.

Finally, numerical results will be presented in order to showcase the investigated approach using realistic data for various locations.

Acknowledgements

Funding by the Hessian Ministry of Higher Education, Research, Science and the Arts - cluster project Clean Circles is gratefully acknowledged.

Optimal control for species selection in continuous culture bioprocesses

Walid Djema[1], **Térence Bayen**[2], **Jean-Luc Gouzé**[3]

1. walid.djema@inria.fr *Centre Inria d'Université Côte d'Azur, Biocore team, Sophia Antipolis, France.* Corresponding author
2. terence.bayen@univ-avignon.fr *Laboratoire de Mathématiques d'Avignon (EA 2151), Avignon Université, 84018 Avignon, France.*
3. jean-luc.gouze@inria.fr *Centre Inria d'Université Côte d'Azur, Macbes team, Sophia Antipolis, France.*

Investigating competition among various cell populations, especially different microorganisms sharing same environment and nutrients, constitutes a major area of research (see [1]). This is particularly relevant for cultures where implementing an optimal control strategy can significantly influence the outcome, favoring species with desirable characteristics [2, 3]. For instance, selecting microalgae species with the higher lipid content for biofuel production, or those that enhance light penetration in dense cultures, can significantly improve the culture yields.

In this presentation, we shall discuss some novel challenges in species selection within chemostat systems (photobioreactors), involving one or two control variables. More precisely, we consider a chemostat system involving two species and two substrates. We are interested in studying the selection of the first species in minimal time : given a threshold (contamination rate), our goal is to drive the system in a minimum time from any initial condition to a target where species two is negligible w.r.t. the first one (while keeping the proportion of the first species significant). Such an optimal control problem is in the spirit of [4–8], but, here, the main novelty is that two substrates are considered, thus extending the study in [4]. Moreover, several control variables are considered which correspond to parameters that can be controlled in practice : in addition to the usual control variable representing the dilution rate, two scalar controls are also considered representing the input substrate concentrations associated with both substrates present in the culture.

We investigate the minimum time control problem where only the dilution rate varies (while the two other controls are fixed to a constant value). We first discuss the reachability of the target under hypotheses on the data defining the control system. Next, our goal is to synthesize an optimal control in open loop and if possible in a feedback form (robust w.r.t. uncertainties inherent in the control of bioprocesses).

By applying the Pontryagin Maximum Principle (PMP), we characterize optimal trajectories under several conditions showing in particular the occurrence of singular arcs and highlighting some turnpike-like features. These conditions are validated through numerical simulations based on direct optimization methods, implemented in the bocop software. In particular, numerical simulations are used to check second-order necessary conditions along singular arcs in order to confirm the solution found theoretically, thanks to the PMP.

Finally, we also discuss another minimum time control problem where now the control is the input substrate concentration associated with both substrates (and the dilution rate is fixed to a constant value). Similar results concerning the minimum time synthesis are obtained using numerical simulations via a direct optimization method. Comparisons between the two problems are also provided.

References

[1] Smith, H.L. and Waltman, P., 1995. The theory of the chemostat: dynamics of microbial competition (Vol. 13). Cambridge university press.

[2] Bernard, O., 2011. Hurdles and challenges for modelling and control of microalgae for CO2 mitigation and biofuel production. Journal of Process Control, 21(10), pp.1378-1389.

[3] Bonnefond, H., Grimaud, G., Rumin, J., Bougaran, G., Talec, A., Gachelin, M., Boutoute, M., Pruvost, E., Bernard, O. and Sciandra, A., 2017. Continuous selection pressure to improve temperature acclimation of Tisochrysis lutea. PloS one, 12(9), p.e0183547.

[4] Bayen, T. and Mairet, F., 2014. Optimization of the separation of two species in a chemostat. Automatica, 50(4), pp.1243-1248.

[5] Djema, W., Bayen, T. and Bernard, O., 2022. Optimal Darwinian Selection of Microorganisms with Internal Storage. Processes, 10(3), p.461.

[6] Djema, W., Giraldi, L., Maslovskaya, S. and Bernard, O., 2021. Turnpike features in optimal selection of species represented by quota models. Automatica, 132, p.109804.

[7] Caillau, J.B., Djema, W., Gouzé, J.L., Maslovskaya, S. and Pomet, J.B., 2022. Turnpike property in optimal microbial metabolite production. Journal of Optimization Theory and Applications, 194(2), pp.375-407.

[8] Fierro, I., Djema, W., Bernard, O., 2023. Optimal control of microalgae culture accounting for photoinhibition and light attenuation. IFAC WC23, Yokohama, Japan., IFAC-PapersOnLine, 56(2), pp.7222-7227.

A new approach to 3D printing: compatible TOSets with POSets

P. Abascal[1], **F. Fueyo**[2], **J. Jiménez** [3], **A. Palacio** [4], **M. L. Serrano** [5]

1. abascal@uniovi.es Universidad de Oviedo, Spain
2. fueyofernando@uniovi.es Universidad de Oviedo, Spain
3. meana@uniovi.es Universidad de Oviedo, Spain
4. palacioantonio@uniovi.es Universidad de Oviedo, Spain
5. mlserrano@uniovi.es Universidad de Oviedo, Spain

1. Introduction

In 3D printing, the importance of optimizing the process becomes evident as it leads to both energy savings and increased productivity. For this reason, being a highly costly process, a metallurgical company reached out to us to conduct a study on the optimization of this type of printing. Generally, existing studies model the process based on finite elements [3, 5]. However, we now propose a new approach to modeling by working with total order relations compatible with a given order relation defined over the set of parts to be printed. A total order relation is said to be compatible with an order relation when the former is contained within the latter. Building upon this compatibility between relations, the 3D printing process is significantly simplified as the order of printing parts is fundamental. The underlying motivation lies in the necessity of applying genetic algorithms to a series of total order relations during the process to find more efficient relations, making it essential to generate those that are compatible with a given order relation. This work presents results to demonstrate the existence of these orders and several algorithms are shown to generate them. One of the most significant allows us to quickly and efficiently generate compatible total orders from a randomly generated total order, which is a crucial aspect when employing genetic algorithms.

2. Some relevant results

Concepts related to order relations can be found in classic texts such as [1] or more contemporary ones like [2] or [4].

Given that total order relations defined on a set can be characterized by a permutation of the elements of the set or their adjacency matrices, results are obtained in terms of permutations of these elements. Among the most relevant results of this work, we can highlight:

Theorem 2.1 *Let (S, R) be an ordered indexed set and σ a permutation of elements of I. The following statements are equivalent:*

a) *σ is compatible with the relation R.*

b) *$M_R^\sigma \otimes M_{T_\sigma}^\sigma = M_R^\sigma$.*

c) *M_R^σ is lower triangular.*

d) *All elements a_i are totally compatible in the permutation σ.*

Theorem 2.2 *Given an indexed POSet (S, R) with n elements, there is always a compatible permutation.*

Theorem 2.3 *Given an indexed ordered set (S, R) with n elements, if $N_i = Card\{a_j \,/\, a_j R a_i\}$ for $i \in \{1, \dots n\}$, a permutation $\sigma = (\sigma(1), \sigma(2), \dots, \sigma(n))$ such that $N_{\sigma(j)} \leq N_{\sigma(i)}$ if $i < j$ is compatible with the relation R.*

Acknowledgements

This work has been partially supported by the collaborative project FUO-115-22.

References

[1] Birkhoff G.: Lattice Theory, American Mathematical Society, Providence. (1948)

[2] Caspard N., Leclerc B., Monjardet B.: Finite Ordered Sets. Concepts, Results and Uses, Encyclopedia of Mathematics and Its Application, 144. Cambridge University Press, Cambridge (2012)

[3] Chowdhury S., Yadaiah N., Prakash C., Ramakrishna S., Dixit S., Gupta L R., Buddhi D.: Laser powder bed fusion: a state-of-the-art review of the technology, materials, properties & defects, and numerical modelling, J. Market. Res. **20**, pp. 2109-2172. (2022) https://doi.org/10.1016/j.jmrt.2022.07.121

[4] Garg V. K.: Introduction to Lattice Theory with Computer Science Applications. Wiley, Hoboken (2015)

[5] Olleak A., Xi Z.: Efficient LPBF process simulation using finite element modeling with adaptive remeshing for distortions and residual stresses prediction, Manuf. Lett. **24**, 140-144 (2020). https://doi.org/10.1016/j.mfglet.2020.05.002. (ISSN 2213-8463)

Analyzing the sensitivity of integer linear programs via optimization oracles

Erik Jansen[1], **Marc E. Pfetsch**[2]

1. jansen@mathematik.tu-darmstadt.de Technische Universität Darmstadt, Germany
2. pfetsch@mathematik.tu-darmstadt.de Technische Universität Darmstadt, Germany

We investigate the behavior of an integer linear program when the objective function changes. We are interested in determining the set of objective vectors under which the given integer solution stays optimal. The set of all such objective vectors forms the normal cone at the solution, i.e., for all objectives in this cone, the solution stays optimal. This cone is described by all active constraints of the integer hull. Since the integer hull is not known in general and expensive to calculate, we are looking for an efficient way to calculate such cones or rather its polar, the radial cone. Therefore, the efficient calculation of radial cones at given vertices can be valuable tool for analyzing the stability of integer solutions and in general integer programs.

We will present an approach which guarantees to calculate all facet-defining inequalities of the radial cone. The algorithm combines existing algorithms modified to fit the purpose. Specifically, it uses 'Investigating Polyhedra by Oracle' from Matthias Walter [5] and the Beneath-Beyond Method [1,3] which incrementally calculate convex hulls. Similar ideas were presented for the analysis of linear programs from Huggins [2] and the calculation of convex hulls from Lassez [4], but, as far as we know, not extended to integer programs.

The general idea of our algorithm is to utilize an optimization oracle to calculate – preferably neighboring – solutions and incrementally extend a current sub cone of the radial cone with these solutions. Initially, this cone consist solely of the solution of interest as its origin. To avoid exhaustive calculation of vertices of the integer hull, we choose the linear objective vectors in such a way that we only extend our cone with relevant vertices. This process is repeated until no further relevant vertex can be found and we obtain the radial cone

For the extension of the sub cone, similar to the Beneath-Beyond Method, we need incidence information in form of its incidence/facial graph. The rules for updating the graph to correctly mimic the extension will be discussed and needed modification to the original Beneath-Beyond update steps will be clarified. Furthermore, we address the handling of potential problematic cases such as unbounded directions and non-full-dimensional polyhedra.

Since we only require the existence of an optimization oracle, we can calculate the radial cone for all linear optimization problems over polyhedra. This extends the algorithm's utility to mixed-integer problems as well. As we have to expect up to exponential many oracle calls for degenerate vertices of the integer hull, we discuss the utilization of gathered information for decreasing computation time and present general computational results.

Acknowledgements

Funding by the Hessian Ministry of Higher Education, Research, Science and the Arts - cluster project Clean Circles is gratefully acknowledged.

References

[1] Herbert Edelsbrunner. *Constructing Convex Hulls*, pages 139–176. Springer Berlin Heidelberg, Berlin, Heidelberg, 1987. doi:10.1007/978-3-642-61568-9_8.

[2] Peter Huggins. ib4e: A software framework for parametrizing specialized lp problems. In Andrés Iglesias and Nobuki Takayama, editors, *Mathematical Software - ICMS 2006*, pages 245–247, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[3] Michael Joswig. Beneath-and-beyond revisited, 2002. arXiv:math/0210133.

[4] Catherine Lassez and Jean-Louis Lassez. Quantifier elimination for conjunctions of linear constraints via a convex hull algorithm. 1992. URL: https://www2.cs.duke.edu/donaldlab/Books/SymbolicNumericalComputation/103-119.pdf.

[5] Matthias Walter. Investigating polyhedra by oracles and analyzing simple extensions of polytopes, 2016. URL: http://nbn-resolving.de/urn:nbn:de:gbv:ma9:1-7737.

Geometric optimization of a lithium-ion battery model

Grégoire Allaire[1,a], **Romain de Loubens**[2,b], **Richard Joly**[1,2,c]

1. *CMAP, Ecole Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France*
2. *TotalEnergies OneTech, 7-9 boulevard Thomas Gobert, 91120 Palaiseau, France*

a. *grégoire.allaire@polytechnique.fr*
b. *romain.de-loubens@totalenergies.com*
c. *richard.joly@totalenergies.com*

1. Introduction

The homogenized model of Doyle, Fuller and Newman [4] (DFN) - serves as the foundation for numerous academic and industrial softwares [5] used to simulate the operation of lithium-ion batteries during charging and discharging. Studying a battery cell $\Omega = \Omega_{\text{anode}} \cup \Omega_{\text{separator}} \cup \Omega_{\text{cathode}}$ subject to a constant current discharge, our goal is to optimize each electrode interface by minimizing a shape dependent objective function.

2. Some details

The multiscale equations of the DFN-model, derived from the theory of porous electrodes, form a coupled system of two unsteady diffusion equations for the macroscale lithium-ion concentration in the liquid electrolyte (c_e) and the microscale solid-phase lithium concentration (c_s),

$$\epsilon_e \frac{\partial c_e}{\partial t} - \nabla \cdot \left(\mathbf{D}_{e,\text{eff}} \nabla c_e\right) = -\frac{t_+^0}{F} \nabla \cdot \mathbf{i}_e + \frac{a_s i_{se}}{F} \qquad \text{in } \Omega \times [0, T_{\max}] \tag{2.1}$$

$$\frac{\partial c_s}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(D_s r^2 \frac{\partial c_s}{\partial r} \right) \qquad \text{in } B_R \times [0, T_{\max}] \tag{2.2}$$

along with two steady-state elliptic equations for the macroscale ionic potential of electrolyte (φ_e) and the macroscale solid-phase electric potential (ϕ_s),

$$-\nabla \cdot \mathbf{i}_e + a_s i_{se} = 0, \quad \mathbf{i}_e = -\mathbf{\Lambda}_{e,\text{eff}} \nabla \varphi_e + \frac{2RT(1 - t_+^0)(1 + \delta_e)}{F} \mathbf{\Lambda}_{e,\text{eff}} \nabla \ln c_e \qquad \text{in } \Omega \times [0, T_{\max}] \tag{2.3}$$

$$\nabla \cdot \left(\mathbf{\Lambda}_{s,\text{eff}} \nabla \phi_s\right) = a_s i_{se} \qquad \text{in } \Omega_{\text{anode}} \cup \Omega_{\text{cathode}} \times [0, T_{\max}] \tag{2.4}$$

The source term of these equations i_{se} is a function depending on all aforementioned variables and is expressed with the Butler-Volmer relation. To these equations, boundary conditions [5] are added, as well as continuity conditions at the interfaces. The solid-phase microstructure of each porous electrode is composed of an active material that is simply represented by a ball B_R of radius R. Finally, $\epsilon_e, a_s, \mathbf{\Lambda}_{e,\text{eff}}, \mathbf{\Lambda}_{s,\text{eff}}$ and $\mathbf{D}_{e,\text{eff}}$ are the homogenized coefficients of the model, and D_s, t_+^0 and δ_e are some intrinsic material parameters.

We propose here an implementation of the so-called pseudo-3D (P3D) version of the DFN model, which combines finite-element and finite-difference methods using the FreeFem and C++ languages. Coupled with the application of geometric optimization tools [1] to the electrode interfaces, this implementation allows us to compute a shape gradient through the adjoint method and to apply a gradient flow algorithm [2] to minimize a performance function under geometric constraints. We impose for instance in this minimization process the non-mixing constraint [3] between the anodic domain and the cathodic domain.

These computations represent a first step towards a complete geometric and topological optimization of the battery cell, including the optimization of the porous microstructure within each electrode.

References

[1] Grégoire Allaire, Charles Dapogny, and François Jouve. Shape and topology optimization. In *Handbook of numerical analysis*, volume 22, pages 1–132. Elsevier, 2021.

[2] Florian Feppon, Grégoire Allaire, and Charles Dapogny. Null space gradient flows for constrained optimization with applications to shape optimization. *ESAIM: Control, Optimisation and Calculus of Variations*, 26:90, 2020.

[3] Florian Feppon, Grégoire Allaire, Charles Dapogny, and Pierre Jolivet. Body-fitted topology optimization of 2d and 3d fluid-to-fluid heat exchangers. *Computer Methods in Applied Mechanics and Engineering*, 376:113638, 2021.

[4] Thomas F Fuller, Marc Doyle, and John Newman. Simulation and optimization of the dual lithium ion insertion cell. *Journal of the electrochemical society*, 141(1):1, 1994.

[5] Vasilena Taralova, Oleg Iliev, and Yalchin Efendiev. Derivation and numerical validation of a homogenized isothermal Li-ion battery model. *J. Eng. Math.*, 101:1–27, 2016. doi:10.1007/s10665-015-9842-6.

The bipartite quadric polytope with partitioned subtotal constraints and its application to pooling problems

Robert Burlacu[1], **Jan Krause**[2], **Christof Kümpel**[3], **Oskar Schneider**[4]

1. robert.burlacu@utn.de, University of Technology Nuremberg, Germany
2. jan.krause@utn.de, University of Technology Nuremberg, Germany
3. christof.kuempel1@gmail.com, KPMG Deutschland, Germany
4. oskar.schneider@fau.de, Horváth, Germany

The bipartite version of the well-known boolean quadric polytope [3] was studied in [5]. This polytope arises in bilinear programming contexts, i.e. in quadratic pograms with two distinct sets of variables X and Y, such that only products between X- and Y-variables are present. We study a similar, but more structured version of the latter polytope, which we call *the bipartite quadric polytope with partitioned subtotal constraints*. Among others, we further involve a subdivision on the X-variables as in [1] and introduce several additional constraints. In particular, we require the sum of the variables of a subdivided X-variable set to be in a given interval and we require the total sum of the X-variables to be equal to one. Elementary polyhedral properties are discussed, including vertex characteristics and valid resp. facet-defining inequalities. Besides the adaption of the known RLT-inequalities [4] we find a new class that we call *subset-m inequalities*. Furthermore, we investigate certain variants of the pooling problem [2] that contain the mentioned polytope as a substructure and demonstrate the potential of found cutting planes in these applications.

Acknowledgements

This research was supported by the Bavarian Research Foundation within the project SHIELD.

References

[1] Andreas Bärmann, Alexander Martin, and Oskar Schneider. The bipartite boolean quadric polytope with multiple-choice constraints. *SIAM Journal on Optimization*, 33:2909–2934, 2023.

[2] Ruth Misener and Christodoulos A Floudas. Advances for the pooling problem: modeling, global optimization, and computational studies. *Appl. Comput. Math*, 8(1):3–22, 2009.

[3] Manfred Padberg. The boolean quadric polytope: some characteristics, facets and relatives. *Mathematical programming*, 45(1-3):139–172, 1989.

[4] Hanif D. Sherali and Amine Alameddine. A new reformulation-linearization technique for bilinear programming problems. *Journal of Global Optimization*, 2:379–410, 1992.

[5] Piyashat Sripratak, Abraham P Punnen, and Tamon Stephen. The bipartite boolean quadric polytope. *Discrete Optimization*, 44(100657), 2022.

On robust mathematical programs with vanishing constraints with uncertain data

Priyanka Bharati[1], **Vivek Laha**[2]

1. priyankabharati09275@gmail.com Banaras Hindu University, India
2. laha.vivek333@gmail.com Banaras Hindu University, India

Abstract

The main objective of this presentation is to explore mathematical programs that incorporate data uncertainty in the vanishing constraints (UMPVC) and to solve them by using a robust optimization framework to deal with the worst-case scenario. To begin with, we derive robust Fritz-John conditions for the UMPVCs and introduce extended no nonzero abnormal multiplier constraint qualification to obtain robust Karush-Kuhn-Tucker conditions. We also identify the robust strong stationary points of the UMPVC and attain sufficient optimality conditions under generalized convexity assumptions. We also identify robust weak stationary points of the UMPVC using a tightened nonlinear programming approach to seek necessary and sufficient robust optimality conditions. The robust version of several constraint qualifications (CQ), like Abadie CQ, Mangasarian-Fromovitz CQ, and linearly independent CQ, are introduced to handle the uncertainties associated with the special structure of the vanishing constraints. Several algorithms are given to apply the results and various examples are presented to illustrate the algorithms.

Homogeneous approach to the subdifferential of the supremum function

Marco A. López-Cerdá

marco.antonio@ua.es Universidad de Alicante, Spain

We start by providing alternative characterizations of the normal cone to the effective domain of the supremum of an arbitrary family of convex functions. These results are then applied to give new formulas for the subdifferential of the supremum function, which use both the active and nonactive functions at the reference point. In contrast with previous works, the main feature of our subdifferential characterization is that the normal cone to the effective domain of the supremum (or to infinite-dimensional sections of this domain) does not appear. The talk also includes a new type of optimality conditions for convex optimization. The results presented in this talk were established in a joint research project with R. Correa and A. Hantoute.

Double control problem: domains and coefficients for elliptic equations

Juan Casado-Díaz[1], **Manuel Luna-Laynez**[2], **Faustino Maestre**[3]

1. jcasadod@us.es Universidad de Sevilla, Spain
2. mllynez@us.es Universidad de Sevilla, Spain
3. fmaestre@us.es Universidad de Sevilla, Spain

A classical problem in the optimal design of materials consists in finding the optimal arrangement of two materials in the sense that it minimizes a certain cost functional [1, 6, 7]. We consider two isotropic conductive materials with diffusion constants $0 < \alpha < \beta$. From mathematical point of view the problem is given by

$$\min_{\omega\subset\Omega \text{ measurable}} \int_\omega F(x,u)\,dx, \qquad \begin{cases} -\mathrm{div}\big((\alpha\chi_\omega + \beta\chi_{\Omega\setminus\omega})\nabla u\big) = f \ \text{ in } \Omega \\ u = 0 \ \text{ on } \partial\Omega \end{cases} \tag{1}$$

where Ω is a bounded open set of $\mathbb{R}^N$, f a given source and the measurable set $\omega \subset \Omega$ is the control variable which determines where the material α is placed.

Another classical problem in optimal design is when we only have one conductive material, but the control variable is the open set where the diffusion equation is posed. Similarly to (1), the problem can be written by

$$\min_{\omega\subset\Omega \text{ open}} \int_\omega F(x,u)\,dx, \qquad \begin{cases} -\Delta u = f \ \text{ in } \omega \\ u = 0 \ \text{ on } \partial\omega. \end{cases} \tag{2}$$

In this work we are interested in considering the couple problem where as in (1), we look for the optimal distribution of two conductive materials and, similarly to (2), we search the set where the diffusion equations holds. If we consider a constraint on the amounts of the materials used in the mixture, the problem can be formulated as

$$\min_{\omega^\alpha,\omega^\beta} \int_{\omega^\alpha\cup\omega^\beta} F(x,u)\,dx$$
$$\begin{cases} -\mathrm{div}\big((\alpha\chi_{\omega^\alpha} + \beta\chi_{\omega^\beta})\nabla u\big) = f \ \text{ in } \omega^\alpha\cup\omega^\beta \\ u = 0 \ \text{ on } \partial(\omega^\alpha\cup\omega^\beta) \\ \omega^\alpha,\omega^\beta \subset \Omega \ \text{ measurable, } \ \omega^\alpha\cup\omega^\beta \ \text{ open, } \ |\omega^\alpha| \le \kappa^\alpha, \ \ |\omega^\beta| \le \kappa^\beta, \end{cases} \tag{3}$$

with κ^α, κ^β two positive constants.

The lack of classical solutions of (1) and (2) is well-known [5]. In this work, we obtain a relaxed formulation of (3), optimality conditions, and we provide a numerical algorithm to solve it. We show some numerical experiments [4].

Acknowledgements

The authors were partially supported by Grant PID2020-116809GB-I00 of the Government of Spain.

References

[1] G. Allaire. *Shape optimization by the homogenization method.* Appl. Math. Sci. 146, Springer-Verlag, New York, 2002.

[2] M.P. Bendsøe, O. Sigmund. *Topology optimization.* Springer-Verlag, Berlin, 2003.

[3] G. Buttazzo, G. Dal Maso. *Shape optimization for Dirichlet problems: relaxed formulation and optimality conditions.* Appl. Math. Optim. 23 (1991), 17–49.

[4] J. Casado-Díaz, M. Luna-Laynez, F. Maestre. *Control problems in the coefficients and the domain for linear elliptic equations.* Submitted.

[5] F. Murat. *Théorèmes de non existence pour des problèmes de contrôle dans les coefficients.* C.R.A.S Sci. Paris A 274 (1972), 395-398.

[6] F. Murat, L. Tartar. *Calcul des variations et homogénéisation*. In *Les méthodes de l'homogénéisation: theorie et applications en physique*, Eirolles, Paris, 1985, 319–369.

[7] L. Tartar. *The general theory of homogenization. A personalized introduction*. Springer, Berlin Heidelberger, 2009.

Maximum likelihood estimation of the covariance matrix of Gaussian Markov Random Fields over graphs of paths

Juan Baz[1], **Pedro Alonso**[2], **Juan Manuel Peña**[3], **Raúl Pérez-Fernández**[4]

1. *Departamento de Estadística, I.O. y D.M., Universidad de Oviedo, España*
2. *Departamento de Matemáticas, Universidad de Oviedo, España*
3. *Departamento de Matemática Aplicada, Universidad de Zaragoza, España*
4. perezfernandez@uniovi.es *Departamento de Estadística, I.O. y D.M., Universidad de Oviedo, España*

Maximum Likelihood Estimation (MLE) is a paramount technique in statistics for the estimation of an unknown parameter of a probability distribution. This technique selects as the estimation of the parameter the value that maximizes the so-called likelihood function, which simply is the joint density at the observed sample as a function of the unknown parameter. In the most didactical examples shown in statistical courses, this maximization is oftentimes an easy-to-solve optimization process that is usually addressed by solving the likelihood equations (derivating the logarithm of the likelihood function and finding the values where the derivative is equal to zero) and/or searching the boundaries of the parametric space. However, there exist more involved scenarios in which the maximization of the likelihood function is not straightforward and dedicated optimization techniques need to be developed. A common example arises in the context of finite mixture models [5], in which the Expectation-Maximization (EM) algorithm was developed in order to find the MLE of the parameters in a mixture model.

As another apparently-unrelated research problem, algebraic operations with covariance matrices are commonly addressed in the field of statistics. For example, Principal Component Analysis (PCA) [3] is an example of a widely used multivariate technique for data reduction that is based on identifying the eigenvalues of the (sample) covariance matrix. If these algebraic operations are not performed symbolically, it becomes inevitable that a certain approximation error arises. Interestingly, as shown by Koev [4], for some specific matrices several algebraic computations (e.g., the computation of the inverse matrix, triangular factorization, determination of eigenvalues and singular values, and resolution of certain linear systems) may be performed with High Relative Accuracy (HRA), meaning that the relative error of the computations is of the order of machine precision. In particular, this is possible whenever the matrix is non-singular and totally positive [1] (i.e., a matrix such that all its minors are non-negative), assuming a bidiagonal factorization is available.

Recent work [2] has studied population covariance matrices that are totally positive, thus opening the door for the use of HRA techniques. Unfortunately, a sample covariance matrix is not necessarily totally positive even if the sample comes from a random vector with a totally positive population covariance matrix. This implies that all the developed work may not be applicable when dealing with real-life data if the sample covariance matrix turns out not to be totally positive. A possible solution is to resort to the MLE of the population covariance matrix under a restriction that assures that this population covariance matrix is totally positive. In particular, the considered restriction is that presented in [2] requiring that the random vector is a Gaussian Markov Random Field over a graph of paths such that all covariances between adjacent variables on the graph are either (1) non-positive or (2) non-negative. The aforementioned MLE problem requires to solve a constrained optimization problem that turns out to be difficult to solve. A naive solution is to subdivide this optimization problem into $n!$ subproblems (where n is the number of nodes in the graph associated with the Gaussian Markov Random Field), one per each possible graph of paths. Obviously, its computational complexity makes this solution not to be very appealing. Here, some results allowing to express the optimization problem as a shortest Hamiltonian path problem are presented. As a result of the present work, popular statistical techniques such as PCA can be performed with HRA.

Acknowledgements

The financial support of Programa Severo Ochoa of Principality of Asturias (BP21042), Ministerio de Ciencia, Innovación y Universidades del Gobierno de España (PID2022-139886NB-I00 and PID2022-140585NB-I00) and Gobierno de Aragón (E41_23R) is acknowledged.

References

[1] T. Ando. Totally positive matrices. *Linear Algebra and its Applications*, 90:165–219, 1987.

[2] J. Baz, P. Alonso, J.M. Peña and R. Pérez-Fernández. Gaussian Markov Random Field and totally positive matrices. *Journal of Computational and Applied Mathematics*, 430:115098, 2023.

[3] R. Bro and A. K. Smilde. Principal component analysis. *Analytical Methods* 6:2812–2831, 2014.

[4] P. Koev. Accurate computations with totally non-negative matrices. *SIAM Journal on Matrix Analysis and Applications*, 29:731–751, 2007.

[5] G. McLachlan and D. Peel. *Finite Mixture Models*. John Wiley & Sons, New York, 2000.

Decomposition approach for mixed-integer nonlinear optimization of heating networks with heating storages

Marc Pfetsch[1], **Lea Rehlich**[2], **Stefan Ulbrich**[3]

1. pfetsch@mathematik.tu-darmstadt.de Technical University of Darmstadt, Germany
2. rehlich@mathematik.tu-darmstadt.de Technical University of Darmstadt, Germany
3. ulbrich@mathematik.tu-darmstadt.de Technical University of Darmstadt, Germany

To reduce carbon emissions the transformation of the heating sector is of great importance. In particular, district heating networks in combination with renewable energy generation and waste heat play a major role. This includes lowering the operating temperatures and transforming the existing networks to decentralized structures. Incorporating this, new operating strategies are needed.

We consider a global optimization approach that aims at finding cost-optimal operating strategies. The optimization problem is based on nonlinear physical equations for describing the network state and binary variables to determine flow directions in the network pipes. The resulting optimization problem is solved with the solver SCIP. For district heating networks of a practically relevant size the solving process of this mixed-integer nonlinear optimization problem leads to high computational costs. To lower the computational costs we include methods that make use of the underlying network structure to reduce the number of binary variables and the number of nonlinear equations.

To absorb fluctuations in the availability of renewable energy generation and waste heat and ensure a reliable heat supply, the usage of heating storages offers great potential. Including heating storages to the optimization problem leads to a coupling of the complex optimization problem over multiple time steps and an additional increase of the complexity. Therefore we consider a two-stage optimization approach which aims at decomposing the time-coupled problem in two optimization problems, that can be solved more efficiently. In the first step the mixed-integer nonlinear optimization problem is solved for every time step separately. In the second step the costs over all time steps are minimized and the optimal storage strategies are determined.

The developed methods are evaluated with numerical results based on a real district heating network.

Acknowledgements

This work was supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK).

Optimization of functions defined over sets of points in polygons with evolutionary algorithms based on Wasserstein barycenters

Merlin Keller[1], **Rodolphe Le Riche** [2], **Julien Pelamatti** [3], **Babacar Sow**[4], **Sanaa Zannane** [5]

1. merlin.keller@edf.fr EDF R&D, France
2. rodolphe.leriche@emse.fr CNRS/LIMOS, FRance
3. julien.pelamatti@edf.fr EDF R&D, France
4. babacar.sow@emse.fr Ecole des Mines de Saint-Etienne, France
5. zannane.sannane@edfenergy.com EDF R&D, France

1. Introduction

We consider the optimization of functions where the design variables are clouds of points (or equivalently, bags of vectors). The clouds can have different sizes and the objective functions are invariant under arbitrary permutations of the points within the cloud. Furthermore, no information related to the convexity and/or smoothness of the functions is available. Such functions are of practical interest as they appear in several real-life optimization problems, such as the design of a wind farm, where the design variables are the set of turbine positions.

2. Evolutionary algorithm based on Wassertein barycenter

The characteristics mentioned above make it hard to use off-the-shelf algorithms, such as gradient-based methods. For these reasons, in this work we introduce a generic approach for such black-box optimization problems where each cloud is modeled as a discrete uniform measure supported by the cloud points. This allows to define stochastic evolutionary optimization algorithms whose transition operators (crossover and mutation) rely on the use of Wasserstein distances and associated barycenters (see [1] for a review). The crossover operator interpolates between two clouds of points, by calculating their average in the Wasserstein sense. We give an illustration of such an interpolation on clouds in Figure 1. We prove that the Wasserstein barycenter has a contracting effect in the sense that the support of the barycenter is included in the closed convex set of the unions of the initial two supports. The latter effect can lead to a premature convergence to degenerated solutions. Specific mutations, once again based on Wasserstein barycenters, are designed to counteract this effect.

We investigate the performance of variants of our Wasserstein-based optimizers by comparing them to classical evolutionary algorithms on a family of test functions including wind farm layout optimization proxies. We first consider point domains under the form of convex polygons, such as squares and trapezes. Subsequently, the points optimization domain is generalized to involve exclusion zones, defined as subdomains which must not include any point of the optimal solution. Exclusion zones constitute a specific type of constraint which is handled by penalization of the functions or in the definition of the algorithm operators (crossover and mutation). The developed optimization method can also be combined with a surrogate model, for instance a Gaussian process, when the objective function is costly. In this case our approach can be used to optimize the acquisition criterion when the latter is defined over clouds of points.

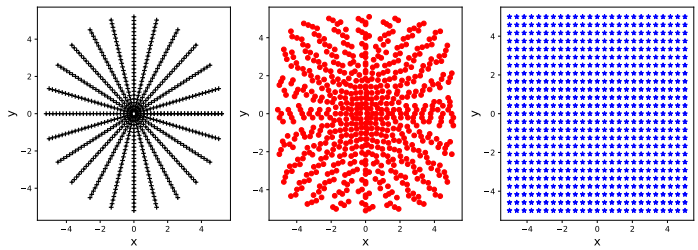


Fig. 1 X_1 and X_2, respectively on bottom left and bottom right, are two initial clouds and X (in the middle) represents their Wasserstein barycenter.

Acknowledgements

This work is funded by the ANR SAMOURAI project ANR-20-CE46-0013.

References

[1] Martial Agueh and Guillaume Carlier. Barycenters in the wasserstein space. *SIAM Journal on Mathematical Analysis*, 43(2):904–924, 2011.

Minisymposia

Control, inverse problems and machine learning
Organizers: Carlos Castro

Numerical analysis in PDE constrained optimization
Organizers: Arnd Rösch, Fredi Tröltzsch

Numerical methods and software for solving optimal control problems
Organizers: Jean-Baptiste Caillau, Joseph Gergaud

Optimal control and machine learning: recent advances
Organizers: Sofya Maslovskaya, Boris Wembe

Optimal control of ODEs: theory and applications
Organizers: Valeriya Lykina, Sabine Pickenhain

Optimal control theory and applications
Organizers: Olivier Cots

Optimization and stabilization of infinite-dimensional dynamical systems
Organizers: Behzad Azmi, Sergio S. Rodrigues

Optimization methods for inverse problems and beyond
Organizers: Neil Dizon, Jyrki Jauhiainen, Tuomo Valkonen

Optimization methods with worst-case complexity guarantees
Organizers: Nikita Doikov, Geovani Nunes Grapiglia

Recent trends in nonsmooth optimization
Organizers: Luise Blank, Andrea Walther

Shape optimization and uncertainty
Organizers: Kathrin Welker, Winnifried Wollner

Theoretical approaches to modern machine learning methods
Organizers: Riccardo Bonalli, Ziad Kobeissi, Mathieu Laurière

Uncertain optimization and related topics
Organizers: Carlos Escudero, César Gutiérrez, Miguel Sama

Control, inverse problems and machine learning

Inverse problem for one-dimensional fluid-solid interaction model

Jone Apraiz[1], **Anna Doubova**[2], **Enrique Fernández-Cara**[3], **Masahiro Yamamoto**[4]

1. jone.apraiz@ehu.eus Universidad del País Vasco, Spain
2. doubova@us.es Universidad de Sevilla, Spain
3. cara@us.es Universidad de Sevilla, Spain
4. myama@next.odn.ne.jp The University of Tokyo, Japan

The analysis and solution of inverse problems has recently increased a lot because of their importance in many applications: elastography and medical imaging, seismology, potential theory, ion transport problems or chromatography and other similar fields.

In this talk we will consider a one-dimensional fluid-solid interaction model governed by the Burgers equation with a time varying interface. It is a preliminary simplified version of other more complicate and more realistic models in higher dimensions. For example, we could consider a system governed by the Navier-Stokes equations around a moving sphere that interacts with the fluid.

We will see the results we have obtained for the inverse problem of determining the shape of the interface from Dirichlet and Neumann data at one end point of the spatial interval. In particular, we will establish uniqueness results and some conditional stability estimates. For the proofs, we have used and adapt some lateral estimates that, in turn, rely on appropriate Carleman and interpolation inequalities (following results in [5]).

The results that will be shown in this talk have been written in a preprint and submitted for publication, [1]. Other related works are [3] about control results for simplified one-dimensional models of fluid-solid interaction and [4] about large time behavior for a simplified n-dimensional model of the same type interaction.

On the other hand, in order to understand the research group's background and approach to this kind of problems, [2] can be read.

This is a joint research project in collaboration with Anna Doubova, Enrique Fernández-Cara and Masahiro Yamamoto.

Acknowledgements

Jone Apraiz was supported by the Grant PID2021-126813NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe" and by the grant IT1615-22 funded the Basque Government. Anna Doubova and Enrique Fernández-Cara were partially supported by MICINN, under Grant PID2020-114976GB-I00. Masahiro Yamamoto was supported by Grant-in-Aid for Scientific Research (A) 20H00117 and Grant-in-Aid for Challenging Research (Pioneering) 21K18142 of the Japan Society for the Promotion of Science.

References

[1] J. Apraiz, A. Doubova, E. Fernández-Cara, M. Yamamoto. Inverse problems for one-dimensional fluid-solid interaction models. Submitted for publication, *Arxiv*, 2401.16546, 2024.

[2] J. Apraiz, A. Doubova, E. Fernández-Cara, M. Yamamoto. Some Inverse Problems for the Burgers Equation and Related Systems. *Communications in Nonlinear Science and Numerical Simulation*, 107:106113-106136, 2022.

[3] A. Doubova, E. Fernández-Cara. Some control results for simplified one-dimensional models of fluid-solid interaction. *Mathematical Models & Methods in Applied Sciences* 15, no. 5, 783–824, 2005.

[4] A. Munnier, E. Zuazua. Large time behavior for a simplified n-dimensional model of fluid-solid interaction. *Communications in Partial Differential Equations*, 30, no. 3, 377–417, 2005.

[5] M. Yamamoto. Carleman estimates for parabolic equations and applications. *Inverse Problems*, 25, 123013, 2009.

Control, inverse problems and machine learning

Optimal control for neural ODE in a long time horizon and applications to the classification and ensemble controllability problems

Jon Asier Bárcena-Petisco

jonasier.barcena@ehu.eus University of the Basque Country UPV/EHU, Spain

In this talk, we study the optimal control, in a long time horizon, of neural ordinary differential equations which are control-affine or whose activation function is homogeneous. When considering the classical regularized empirical risk minimization problem we show that, in long time and under structural assumption on the activation function, the final state of the optimal trajectories has zero training error if the data can be interpolated and if the error can be taken to zero with a cost proportional to the error. These hypotheses are fulfilled in the classification and ensemble controllability problems for some relevant activation and loss functions. Our proofs are mainly constructive combined with a proof by contradiction: we find that in long time horizon if the final error is not zero, we can construct a less expensive control which takes the error to zero. Moreover, we prove that the norm of the optimal control is constant. Finally, we show the sharpness of our hypotheses by giving an example for which the error of the final state of the optimal trajectory, even if it decays, is strictly positive for any time.

The talk is based on the paper [1], and provides answers to open problems posed in [2–4].

Acknowledgements

The paper [1] has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement NO: 694126-DyCon). It is also supported by the Grant PID2021-126813NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", and by the grant IT1615-22 funded the Basque Government.

References

[1] Jon Asier Bárcena-Petisco. Optimal control for neural ode in a long time horizon and applications to the classification and simultaneous controllability problems. *HAL-03299270v4*, 2024.

[2] C. Esteve, B. Geshkovski, D. Pighin, and E. Zuazua. Large-time asymptotics in deep learning. *arXiv preprint arXiv:2008.02491v2*, 2021.

[3] C. Esteve-Yagüe and B. Geshkovski. Sparsity in long-time control of neural odes. *Syst. Control Lett.*, 172:105452, 2023.

[4] D. Ruiz-Balet, E. Affili, and E. Zuazua. Interpolation and approximation via momentum resnets and neural odes. *Syst. Control Lett.*, 162:105182, 2022.

Control, inverse problems and machine learning

Numerical approximation of some inverse source problems associated to the wave equation

Carlos Castro
carlos.castro@upm.es Universidad Politécnica de Madrid, Spain

We analyze the use of two different numerical methods to approximate the solutions of inverse source problems associated to the wave equation. The inverse problem consists in recovering the source term from known boundary information of a single solution for sufficiently large time. It is well-known that a suitable boundary observability inequatity gives the stability of such inverse problems. However, natural discretizations of these inverse problems (as finite elements or finite differences) may not inherit the observability inequality, uniformly with respect to the discretization parameter. In this case, the solution of the discrete inverse problem cannot be used to approximate the solution of the continuous one. In this talk we present two numerical methods that overcome this difficulty. The first one is based on a mixed finite element formulation and can be extended to variable coefficients in one dimension. This extension is delicate and it is based on nonstandard uniform spectral properties of the associated discrete operator. The second is based on a polynomial spectral collocation method and can be extended to the elasticity system. The first part of this work is done in collaboration with S. Micu and the second one with S. Boumimez.

Control, inverse problems and machine learning

A random domain decomposition scheme for parabolic PDE constrained optimization problems

Alberto Domínguez Corella

alberto.corella@fau.de University of Erlangen-Nuremberg

Probability theory has had a significant impact on various areas of mathematics, particularly in algorithms and combinatorics. The use of randomness in algorithms dates back to the Monte Carlo methods [6, 7]. In subsequent years, random algorithms gained strength in optimization, notably with the emergence of techniques such as simulated annealing [5] and genetic algorithms [4], which were used in complex optimization problems. In recent years, stochastic gradient descent has attracted attention in machine learning and artificial intelligence; its development has been crucial in training large-scale models, specially machine learning algorithms, see [1, 2].

In this talk, we discuss a random domain decomposition scheme for parabolic optimal control problems, inspired by mini-batch algorithms used in machine learning. This scheme is based on the papers [3, 8] The method discretizes the parabolic equation using the explicit Euler scheme and replaces the full elliptic operator with a randomly selected elliptic operator acting on a part of the domain at each step. This approach aims to reduce computational time. We will demonstrate the convergence of the scheme, highlighting the trade-off between speed and accuracy. Special attention is paid to optimal control problems with bang-bang minimizers.

Acknowledgements

The author has been funded by means of an Alexander von Humboldt scholarship.

References

[1] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*, pages 177–186. Physica-Verlag/Springer, Heidelberg, 2010.

[2] Léon Bottou, Frank E. Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *SIAM Rev.*, 60(2):223–311, 2018. doi:10.1137/16M1080173.

[3] Monika Eisenmann and Tony Stillfjord. A randomized operator splitting scheme inspired by stochastic optimization methods. *arXiv preprint arXiv:2210.05375*, 2022.

[4] John H. Holland. *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, MI, 1975. An introductory analysis with applications to biology, control, and artificial intelligence.

[5] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983. URL: https://www.science.org/doi/abs/10.1126/science.220.4598.671, arXiv:https://www.science.org/doi/pdf/10.1126/science.220.4598.671, doi:10.1126/science.220.4598.671.

[6] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[7] Nicholas Metropolis and S. Ulam. The Monte Carlo method. *J. Amer. Statist. Assoc.*, 44:335–341, 1949. URL: http://links.jstor.org/sici?sici=0162-1459(194909)44:247<335:TMCM>2.0.CO;2-3&origin=MSN.

[8] D. W. M. Veldman and E. Zuazua. A framework for randomized time-splitting in linear-quadratic optimal control. *Numer. Math.*, 151(2):495–549, 2022. doi:10.1007/s00211-022-01290-3.

Control, inverse problems and machine learning

Controllability and augmented Lagrangian techniques

Enrique Fernández Cara

cara@us.es University of Sevilla, Spain

This contribution deals with the numerical solution of null controllability problems for heat-like equations and systems. In each case, we start from a well known formulation due to Fursikov and Imanuvilov and we apply duality-penalty techniques that lead to an augmented Lagrangian and several iterative algorithms. We prove several convergence results and we illustrate the situation with some numerical experiments.

Control, inverse problems and machine learning

Some chemotaxis PDE-constrained optimal control problems. Analysis and approximation.

Francisco M. Guillen-Gonzalez
guillen@us.es EDAN and IMUS, Universidad de Sevilla, Spain

Chemotaxis PDE problems appear for modeling spatial movement of live populations (cells for instance) which are attracted or repulsed by chemical signals. Other possible effects are also considered as self-diffusion (of cells and chemical), chemical production or consumption by cells, etc. The mathematical analysis of Chemotaxis PDE systems has been largely developed in the last twenty years, with results highly dependent of each system. These results are mainly: finite or infinite time blow-up versus bounded classical solutions, and large time stability of constant states versus pattern formation phenomena. The great part of these results use existence and uniqueness of local in time classical solutions (using Amann's argument [1]), extensibility criteria and a priori estimates (from below for blow-up or from above for global in time solutions), see [2] and the references therein cited.

In the last years, we have considered optimal control problems subjected to chemotaxis PDE problems, where the control acts directly on the chemical signal equation by means of a bilinear term (and only indirectly on the cells equation). In particular, from a mathematical point of view, the bilinear control adds a reaction term with a non-regular coefficient, hence in general the previous classical solution arguments ([1]) cannot be used. Instead, the generalized or weak solutions setting must be considered. From the analytical point of view, three main results have been studied: existence of global optimal solution (via minimizing sequence), first-order necessary optimality conditions (based on a generic Lagrange multiplier Theorem) and regularity of the Lagrange Multiplier problem (via a very-weak vs strong uniqueness result). These results were obtained with some collaborators: M.A. Rodriguez-Bellido [3, 4, 6] (Universidad de Sevilla, Spain), E. Mallea Zepeda [3–5] (Universidad Tarapaca, Chile), E.J Villamizar Roa [5] (Universidad Industrial de Santander, Colombia), P. Braz and C. Perusato [6] (Universidade Federal de Pernambuco, Brazil) and A.L.Correa Vianna Filho [7, 8] (Universidade Federal do Paraná, Brazil).

In the first part of this talk, we will explore the possibility to have local uniqueness of state with respect to the control and the derivate of the corresponding local control-to-state operator, by applying the Implicit Function Theorem in Banach spaces. In these cases, the gradient of the reduced cost functional can be identified, which is applied to arrive at necessary optimality conditions and to describe gradient descent methods.

In the second part, the numerical approximation of Keller-Segel-constrained optimal control is studied, by using the discretize-then-optimize approach. Three type of controls are considered always acting directly on the chemical equation: distributed bilinear control, a linear Neumann or Robin boundary control and a bilinear boundary control. Our aim is to minimize a tracking cost functional driving the cell density near of a given target state. We use an upwind finite volume numerical scheme to approach the state and the corresponding discrete adjoint system, computing the discrete gradient exactly. Then, to minimize the cost functional we employ the so-called ADAM method, which is a descent gradient adapting the pass via the first and second moment.

References

[1] H. Amann, *Linear and Quasilinear Parabolic Problems: Abstract Linear Theory*. Vol. 1, Monogr. Math. 89, Birkhäuser, Basel, 1995.

[2] N. Bellomo, A. Bellouquid, Y. Tao, and M. Winkler. Toward a mathematical theory of Keller-Segel models of pattern formation in biological tissues. *Math. Models Methods Appl. Sci.*, 25 (2015), pp. 1663-1763.

[3] F. Guillén-González, E. Mallea-Zepeda, M.A. Rodriguez-Bellido. Optimal bilinear control problem related to a chemo-repulsion system in 2D domains. ESAIM-Control Optim. Calc. Variat., 26 (2020) 29.

[4] F. Guillén-González, E. Mallea-Zepeda, M.A. Rodriguez-Bellido. A Regularity Criterion for a 3D Chemo-Repulsion System and Its Application to a Bilinear Optimal Control Problem. SIAM J. Control and Optimization, 58, no. 3 (2020), 1457-1490.

[5] F. Guillén-González, E. Mallea-Zepeda, E.J. Villamizar-Roa. On a Bi-dimensional chemo-repulsion model with nonlinear production and a related optimal control problem. ACTA APPLICANDAE MATHEMATICAE, (2020) Vol.170, No.1, 963-979.

[6] P. Braz e Silva, F. Guillén-González, C.F. Perusato, M.A. Rodriguez-Bellido. Bilinear optimal control of the Keller-Segel logistic model in 2D-domains. Applied Mathematics and Optimization, Vol. 87, No. 3, 55-1 55-20 (2023).

[7] A.L. Correa Vianna Filho, F. Guillén-González. An Optimal Control Problem Subject to Strong Solutions of Chemotaxis-Consumption Models. SICON, Vol. 61, No. 5 (2023), pp. 3156-3182.

[8] A.L. Correa Vianna Filho, F. Guillén-González. Optimal control related to weak solutions of a chemotaxis-consumption model. Applied Mathematics and Optimization (2024) 89:98. https://doi.org/10.1007/s00245-024-10109-6

CONTROL, INVERSE PROBLEMS AND MACHINE LEARNING

Finite element error analysis of affine optimal control problems

Nicolai Jork

nicolai.jork@tuwien.ac.at Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria

1. Introduction

In this talk, we are concerned with error estimates for the numerical approximation of affine optimal control problems subject to semilinear elliptic PDEs. For the error estimates, we focus on local minimizers that satisfy certain local growth conditions. The local growth conditions we consider appeared recently in the context of solution stability and encompass the joint growth of the first and second variations of the objective functional. These types of growth conditions are especially meaningful for affine control constrained optimal control problems because the first variation can satisfy a growth condition, which is not the case for unconstrained problems. The main results of this talk are the achievement of error estimates for the numerical approximations generated by a finite element scheme with piecewise constant controls or a variational discretization scheme. Even though the considered growth conditions are weaker than the ones appearing in the recent literature on finite element error estimates for affine problems, we can substantially improve the existing error estimates for both the optimal controls and the states.

Acknowledgements

The second author was partially supported by the FWF grants P-31400-N32, I4571-N, and P36344N.

CONTROL, INVERSE PROBLEMS AND MACHINE LEARNING

Algae for bioenergy: an optimal control aproach

Lino J. Alvarez-Vázquez[1], **Aurea Martínez**[2], **Carmen Rodríguez**[3], **Miguel E. Vázquez-Méndez**[4,5]

1. lino@dma.uvigo.es *Universidade de Vigo, Spain*
2. aurea@dma.uvigo.es *Universidade de Vigo, Spain.* Corresponding author
3. carmen.rodriguez@usc.es *Universidade de Santiago de Compostela, Spain*
4. *CITMAga, Spain*
5. miguelernesto.vazquez@usc.es *Universidade de Santiago de Compostela, Spain*

We propose a novel methodology to accomplish the optimal performance of a raceway: an open-channel pond with a rotating paddlewheel, where circulating wastewater is used for the cultivation of algae that will be employed as source for bioenergy production [1, 4].

Algal productivity maximization is addressed here by means of optimal control techniques for partial differential equations, within a simulation-based optimization framework. So, we introduce a rigorously detailed mathematical formulation of the optimal control problem (where the state system couples Navier-Stokes equations on a free surface domain with a large system of nonlinear convection-diffusion-reaction equations for the evolution of algae, nutrients and oxygen, where the design variables are the initial height of water and the velocity of the paddlewheel, and where the objective function corresponds to the maximization of algal concentration at the process final time -under some geometric and technological constraints) [2, 3].

Then, we suggest a numerical algorithm for its computational resolution where, within an ALE strategy, we use finite element techniques for the simulation step and a gradient-free algorithm for the maximization one.

Finally, we show some preliminary results related to the numerical optimization of the problem.

Acknowledgements

This research was funded by Ministerio de Ciencia e Innovación (Spain) and NextGenerationEU (European Union) grant number TED2021-129324B-I00.

References

[1] S.K. Bhatia, S. Mehariya et al. Wastewater based microalgal biorefinery for bioenergy production: Progress and challenges. *Sci. Total Environ.*, 751:141599, 2021.

[2] A. Martínez, L.J. Alvarez-Vázquez, C. Rodríguez and M.E. Vázquez-Méndez. Algal cultivation for bioenergy production: First mathematical modeling results in raceways. *ADMOS 2023*, Gothenburg (Sweden), 2023.

[3] A. Martínez, L.J. Alvarez-Vázquez, C. Rodríguez and M.E. Vázquez-Méndez. Preliminary numerical results in the optimization of bioenergy-intended raceway ponds. *ENUMATH 2023*, Lisboa (Portugal), 2023.

[4] F. Rayen, B. Taidi and D. Pareau. Optimization of a raceway pond system for wastewater treatment: A review. *Critical Rev. Biotechnol.* , 39:422–435, 2019.

CONTROL, INVERSE PROBLEMS AND MACHINE LEARNING

Shape-programming in hyperelasticity through differential growth

Rogelio Ortigosa-Martínez[1], **Jesús Martínez-Frutos**[1], **Carlos Mora-Corral**[2], **Pablo Pedregal**[3], **Francisco Periago**[1]

1. https://multisimo.com/ Multiphysics Simulation and Optimization Lab, Technical University of Cartagena, Spain
2. carlos.mora@uam.es Universidad Autónoma de Madrid and Instituto de Ciencias Matemáticas, Spain
3. pablo.pedregal@uclm.es Universidad de Castilla-La Mancha, Spain

1. Introduction

The presentation wil be concerned with the growth-driven shape-programming problem, which involves determining a growth tensor that can produce a deformation on a hyperelastic body reaching a given target shape. We consider the two cases of globally compatible growth, where the growth tensor is *a deformation gradient* over the undeformed domain, and the incompatible one, which discards such hypothesis. We formulate the problem within the framework of optimal control theory in hyperelasticity. The Hausdorff distance is used to quantify dissimilarities between shapes; the complexity of the actuation is incorporated in the cost functional as well. Boundary conditions and external loads are allowed in the state law, thus extending previous works where the stress-free hypothesis turns out to be essential. A rigorous mathematical analysis is then carried out to prove the well-posedness of the problem. The numerical approximation is performed using gradient-based optimisation algorithms. Our main goal in this part is to show the possibility to apply inverse techniques for the numerical approximation of this problem, which *allows us to address* more generic situations than those covered by analytical approaches. Several numerical experiments for beam-like and shell-type geometries illustrate the performance of the proposed numerical scheme.

2. Some details

Growth is biological process susceptible to being mimicked by artificial soft materials. However, the topic of mathematical analysis and numerical simulation of growth control is in its infancy, insofar as the mathematical analysis of soft materials actuated by growth is missing in the literature. This study sets up the problem of optimal growth within the framework of optimal control theory in nonlinear elasticity. The control variable is a growth tensor. The state variable is the deformation of the actuated soft continuum. As usual in hyperelasticity theory, that deformation is a minimiser of a polyconvex energy functional. The cost function uses the Hausdorff distance to account for dissimilarities between the desired shape and the final configuration. It also includes a term to deal with the complexity of the activation. More details may be found in [1].

Acknowledgements

R. Ortigosa-Martínez, J. Martínez-Frutos and F. Periago have been supported by grant PID2022-141957OA-C22 funded by MICIU/AEI/10.13039/501100011033, by RDF A way of making Europe, and by the Autonomous Community of the Región of Murcia, Spain, through the programme for the development of scientific and technical research by competitive groups (21996/PI/22), included in the Regional Program for the Promotion of Scientific and Technical Research of Fundación Séneca – Agencia de Ciencia y Tecnología de la Región de Murcia.

C. Mora-Corral has been supported by the Agencia Estatal de Investigación of the Spanish Ministry of Research and Innovation, through project PID2021-124195NB-C32 and the Severo Ochoa Programme for Centres of Excellence in R&D CEX2019-000904-S, by the Madrid Government (Comunidad de Madrid, Spain) under the multiannual Agreement with UAM in the line for the Excellence of the University Research Staff in the context of the V PRICIT (Regional Programme of Research and Technological Innovation), and by the ERC Advanced Grant 834728.

P. Pedregal has been supported by grants PID2020-116207GB-I00 and SBPLY/19/180501/000110.

References

[1] R. Ortigosa-Martínez, J. Martínez-Frutos, C. Mora-Corral, P. Pedregal and F. Periago *Shape-programming in hyperelasticity through differential growth*. Applied Mathematics and Optimization. Accepted for publication (2024).

Control, inverse problems and machine learning

Matching ensembles of measures with Transformers

Domènec Ruiz i Balet
d.ruiz-i-balet@imperial.ac.uk Imperial College, UK

Transformers have been one of the keys for the success of large language models such as ChatGPT (among others). In this talk, we will first introduce Transformers together with the attention mechanism. Later, we will consider the problem of matching an ensemble of N input probability measures with N output probability measures via a mean-field continuity equation direcly derived from Transformers.

Control, inverse problems and machine learning

Mathematical programming for the geometric design of linear transport infrastructures

Pedro Blanco-Valcárcel[1], Gerado Casal[2], Alberte Castro[3], Duarte Santamarina[4], Miguel E. Vázquez-Méndez[5,6]

1. pedro.blanco.valcarcel@rai.usc.es *Universidade de Santiago de Compostela, Spain*
2. xerardo.casal@usc.es *Universidade de Santiago de Compostela, Spain*
3. alberte.castro@usc.es *Universidade de Santiago de Compostela, Spain*
4. duarte.santamarina@usc.es *Universidade de Santiago de Compostela, Spain*
5. *CITMAga, Spain*
6. miguelernesto.vazquez@usc.es *Universidade de Santiago de Compostela, Spain.* Corresponding author

The geometric design process of a linear transport infrastructure, such as a highway or railway line, is a very time-consuming task. Usually, this process is done in two different phases that are carried out consecutively. In the first stage, a corridor is determined considering economical aspects ([3]), environmental aspects ([1]) and so on. In the second stage, the final layout is designed within this corridor, also considering aspects of functionality that must include all safety and comfort conditions required by the legislation of each country. In recent years, many scientific works have proposed different models and optimization methods to automate this second stage (see, for instance, [2] and [5]), but today it is still carried out manually by the civil engineer, who only uses some commercial road design software and relies mainly on his own experience.

Considering economical, enviromental, social, comfort, security, etc. aspects, and prioritizing infrastructure costs, in this work we present a new mathematical approach ([4]) to simultaneously address the two stages discussed above. In the framework of Mixed Integer Non Liner Programming (MINLP), a model is proposed to minimize the main infrastructure costs under all comfort and safety constraints included in the current Spanish legislation. This model is solved with an algorithm based on an exhaustive search into the integer variables, combined with a random multi-start of a sequential quadratic programming (SQP) method to deal with the continuous ones. The local minima provide by this algorithm is used to determine all possible corridors by means of a clustering technique. Finally, a decision-making aid module is also proposed to choose (i) a layout alternative within each corridor and (ii) the best of this alternatives resulting in the final layout. The usefulness of this tool is shown in a real life study: the design of a bypass on the Spanish road N-640, circumventing the urban area of Meira (Lugo) and avoiding other small villages, isolated buildings and environmentally protected areas.

Acknowledgements

This work was supported by funding from project TED2021-129324B-I00 of the Ministerio de Ciencia e Innovación (Spain) and NextGenerationEU (European Union).

References

[1] Néstor García-Chan, Lino J. Alvarez-Vázquez, Aurea Martínez, and Miguel E. Vázquez-Méndez. Designing an ecologically optimized road corridor surrounding restricted urban areas: A mathematical methodology. *Math. Comput. Simul.*, 190:745–759, 2021.

[2] Sukanto Mondal, Yves Lucet, and Warren Hare. Optimizing horizontal alignment of roads in a specified corridor. *Comput. Oper. Res.*, 64:130–138, 2015.

[3] Yasha Pushak, Warren Hare, and Yves Lucet. Multiple-path selection for new highway alignments using discrete algorithms. *Eur. J. Oper. Res.*, 248(2):415–427, 2016.

[4] Miguel E. Vázquez-Méndez, Pedro Blanco-Valcarcel, Gerardo Casal, Alberte Castro, and Duarte Santamarina. A mathematical tool for road design projects. *Eng. Optimiz.*, 0(0):1–23, 2024.

[5] Miguel E. Vázquez-Méndez, Gerardo Casal, Alberte Castro, and Duarte Santamarina. Optimization of an urban railway bypass. A case study in A Coruña-Lugo line, Northwest of Spain. *Comput. Ind. Eng.*, 151:106935, 2021.

Control, inverse problems and machine learning

Optimizing the design of an ecological road corridor

Lino J. Alvarez-Vázquez[1], **Néstor García-Chan**[2], **Aurea Martínez**[3], **Miguel E. Vázquez-Méndez**[4]

1. lino@dma.uvigo.es Universidade de Vigo, Spain
2. nestor.gchan@academicos.udg.mx Universidad de Guadalajara, Mexico
3. aurea@dma.uvigo.es Universidade de Vigo, Spain
4. miguelernesto.vazquez@usc.es CITMAga & Universidade de Santiago, Spain

The use of optimization techniques for the optimal design of roads and railways has increased in recent years. The environmental impact of a layout is usually given in terms of the land use where it runs (avoiding some ecologically protected areas), without considering air pollution (in these or other sensitive areas) due to vehicular traffic on the road [1–3].

This work addresses this issue and proposes an automatic method for obtaining a specific corridor (optimal in terms of air pollution), where the economically optimized road must be designed in a later stage. In particular, we introduce a novel methodology in order to obtain the optimal design (length and location) for a road corridor to be connected to an already existing urban road network, avoiding a set of restricted zones, and reducing the undesirable effects of air pollution at some sensitive areas that need to be protected. The usefulness of this approach is shown in a real case study posed in a region that suffers from serious episodes of environmental pollution: the Guadalajara Metropolitan Area (Mexico).

Combining a 1D traffic simulation model with a 2D air pollution model and using classical techniques for optimal control of partial differential equations, the problem is formulated and solved in the framework of Mixed Integer Nonlinear Programming. So, after a rigorous mathematical formulation of the environmental problem, we propose a full algorithm for computing the ecologically optimized design of the road corridor.

The efficiency of our methodology has been assessed through several computational experiences for a real-world case study. By a direct analysis of the numerical results obtained there, we can deduce that some issues in the model present a low sensitivity when computing the optimal design (as could be, for instance, the volume of traffic expected for the road corridor), but other ones show a much higher sensitivity (mainly, the total length of the corridor).

Finally, the achieved results also indicate a simple (expected) fact: if the decision-maker does not have enough resources to maintain a sufficiently large road corridor, it will be very hard to hold the objective of dropping down the pollution levels in those sensitive areas.

Last but not least, we should also mention that, although our study focuses on the optimal design of an urban road corridor, the novel methodology presented here can be used in many other different scenarios, for example, the construction of a road through protected forest reserves (not necessarily within a city), or even the design of a trainway.

Acknowledgements

This research was funded by Ministerio de Ciencia e Innovación (Spain) and NextGenerationEU (European Union) grant number TED2021-129324B-I00. The second author also thanks the support from Sistema Nacional de Investigadores (Mexico) under Grant SNI-52768, from Programa para el Desarrollo Profesional Docente (Mexico) under Grant PRODEP/103.5/16/8066, and from CONACyT/Ciencia de Frontera (Mexico) under Grant 217556.

References

[1] L.J. Alvarez-Vázquez, N. García-Chan, A. Martínez and M.E. Vázquez-Méndez. Optimal control of urban air pollution related to traffic flow in road networks. *Math. Control Rel. Fields*, 8:177–193, 2018.

[2] L.J. Alvarez-Vázquez, N. García-Chan, A. Martínez and M.E. Vázquez-Méndez. Designing an ecologically optimized road corridor surrounding restricted urban areas: A mathematical methodology. *Math. Comput. Simul.*, 190:745–759, 2021.

[3] M.E. Vázquez-Méndez, L.J. Alvarez-Vázquez, N. García-Chan and A. Martínez. Optimal management of an urban road network with an environmental perspective. *Comput. Math. Appl.*, 77:1786–1797, 2019.

Numerical analysis in PDE constrained optimization

Superlinear convergence of a semismooth Newton method for some optimization problems with applications to control theory

Eduardo Casas

eduardo.casas@unican.es Universidad de Cantabria, Spain

Let $(X, \mathcal{S}, \mu)$ be a measure space with $\mu(X) < \infty$. In this paper, we prove the superlinear convergence of a semismooth Newton method to solve the following abstract optimization problem:

$$\text{(P)} \qquad \min_{\alpha \le u(x) \le \beta \text{ a.e.}[\mu]} \mathcal{J}(u) + \frac{\kappa}{2}\|u\|^2_{L^2(X)}$$

where $\kappa > 0$, $-\infty \le \alpha < \beta \le +\infty$, and $\mathcal{J} : L^p(X) \longrightarrow \mathbb{R}$ is a function of class C^2 for some $p \in [2, +\infty)$. Many control problems fit this abstract formulation. In particular, we apply this abstract result to distributed control problems of a semilinear elliptic equation, to boundary bilinear control problems associated with a semilinear elliptic equation, and to distributed control of a semilinear parabolic equation. The superlinear convergence to a local minimizer $\bar{u}$ is proved assuming that the no-gap second order sufficient optimality condition and the strict complementarity condition are fulfilled at $\bar{u}$.

NUMERICAL ANALYSIS IN PDE CONSTRAINED OPTIMIZATION

On the identification and optimization of nonsmooth superposition operators in semilinear elliptic PDEs

Constantin Christof[1], **Julia Kowalczyk**[2]

1. christof@tum.de, Technische Universität München, CIT, Department of Mathematics, Germany
2. julia.kowalczyk@tum.de, Technische Universität München, CIT, Department of Mathematics, Germany

We study infinite-dimensional optimization problems of the form

$$\left.\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|y - y_D\|^2_{L^2(\Omega)} + \nu_1 \|g'\|_{L^1(\mathbb{R})} + \frac{\nu_2}{2}\|g' - u_D\|^2_{L^2(\mathbb{R})} \\ \text{w.r.t.} \quad & y \in H^1(\Omega), \quad g \in H^1_{loc}(\mathbb{R}), \\ \text{s.t.} \quad & -\Delta y + g(y) = f \text{ in } \Omega, \quad y = 0 \text{ on } \partial\Omega, \\ \text{and} \quad & g' \geq 0 \text{ a.e. in } \mathbb{R}, \\ \text{and} \quad & g(0) = 0, \end{aligned}\right\} \tag{P}$$

that aim to identify the Nemytskii operator g in the nonlinear part of a prototypical semilinear elliptic partial differential equation which minimizes the $L^2(\Omega)$-distance between the PDE-solution y and a given desired state y_D. Here, $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, is a bounded domain which is convex or possesses a $C^{1,1}$-boundary; $y_D \in L^2(\Omega)$, $f \in L^2(\Omega) \setminus \{0\}$, and $u_D \in L^2(\mathbb{R})$ are given; g' denotes the weak derivative of g; $\nu_1 \geq 0$ and $\nu_2 > 0$ are (small) regularization parameters; Δ denotes the Laplace operator; and the objective function is understood as an extended real-valued function on $H^1(\Omega) \times H^1_{loc}(\mathbb{R})$ with values in $[0, \infty]$. The salient feature of (P), that distinguishes our analysis from previous works, is that the function g inducing the Nemytskii operator is a-priori only assumed to be an element of $H^1_{loc}(\mathbb{R})$. This low regularity of g makes the problem (P) a suitable point of departure for the rigorous analysis of training problems for learning-informed PDEs in which an unknown superposition operator is approximated by means of a neural network with nonsmooth activation functions (ReLU, leaky-ReLU, etc.). We establish that, despite the low regularity of the controls g, it is possible to derive a classical stationarity system for local minimizers of (P) and to identify the solutions of (P) by means of a gradient projection method. It is also shown that the established first-order necessary optimality conditions imply that locally optimal superposition operators of (P) share various characteristic properties with commonly used activation functions: They are always sigmoidal, continuously differentiable away from the origin, and typically possess a distinct kink at zero. The talk concludes with numerical experiments that confirm the theoretical findings.

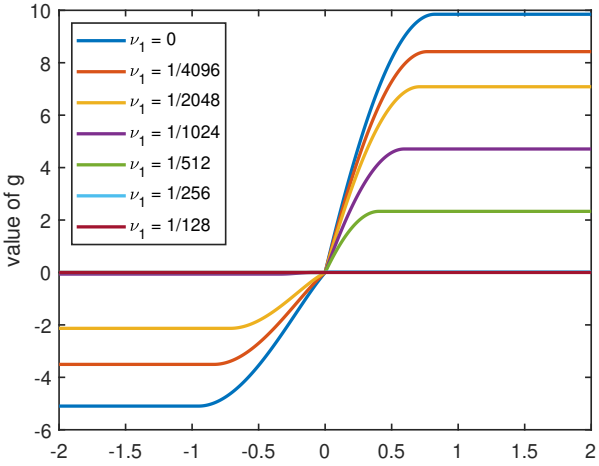


Fig. 1 Optimal superposition operators for an instance of (P) with fixed ν_2 and varying ν_1.

References

[1] Constantin Christof and Julia Kowalczyk. On the Identification and Optimization of Nonsmooth Superposition Operators in Semilinear Elliptic PDEs. *ESAIM COCV*, to appear, 2024.

Numerical analysis in PDE constrained optimization

Error estimates for the discretization of the velocity tracking problem with pointwise-integral control constraints in time-space

Eduardo Casas[1], **Konstantinos Chrysafinos**[2]

1. eduardo.casas@unican.es Universidad de Cantabria, Spain
2. chrysafinos@math.ntua.gr National Technical University of Athens, Greece

The velocity tracking control problem associated to the evolutionary Navier-Stokes equations for two-dimensional flows is studied. The Navier-Stokes equations are written as:

$$\begin{cases} \mathbf{y}_t - \nu\Delta\mathbf{y} + (\mathbf{y}\cdot\nabla)\mathbf{y} + \nabla p = \mathbf{f} + \mathbf{u}\chi_\omega \ \text{ in } Q = \Omega\times(0,T), \\ \operatorname{div}\mathbf{y} = 0 \ \text{ in } Q,\ \mathbf{y}(0) = \mathbf{y}_0 \ \text{ in } \Omega,\ \mathbf{y} = 0 \ \text{ on } \Sigma = \Gamma\times(0,T), \end{cases} \tag{0.1}$$

where, $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ is the velocity field of the fluid, p is the pressure, $\nu > 0$ is the viscosity, $\mathbf{f}$ and $\mathbf{u}$ represent the body forces, and $\mathbf{y}_0$ denotes the initial velocity. The goal of this work is to control the system through the forces $\mathbf{u}$ acting on $Q_\omega = \omega\times(0,T)$ with $0 < T < \infty$ where ω a subdomain of Ω of positive Lebesgue measure. The optimal control problem under consideration is formulated as follows

$$(\mathbf{P}) \quad \begin{cases} \min J(\mathbf{u}) := \dfrac{1}{2}\displaystyle\int_Q |\mathbf{y}_\mathbf{u} - \mathbf{y}_d|^2 dxdt \\ \mathbf{u}\in\mathcal{U}_{ad} \end{cases}$$

where $\mathcal{U}_{ad} = \{\mathbf{u}\in L^\infty(0,T;\mathbf{L}(\omega)) : \|\mathbf{u}(t)\|^2_{\mathbf{L}(\omega)} \le \gamma \text{ a.e. in } (0,T)\}$ with $0 < \gamma < \infty$.

Our aim is to study the minimization of the tracking functional J in $\mathcal{U}_{ad}$, without using a regularizing $L^2(0,T;\mathbf{L}^2(\omega))$ term in the functional, often called Tiknonov regularizing term, and under the physical constraints imposed by the structure of the admissible set $\mathcal{U}_{ad}$ (see, e.g., [1]). Here, $(\mathbf{u},\mathbf{y})$ satisfies (0.1), and the tracking profile $\mathbf{y}_d$ is chosen in a suitable Banach space.

The regularizing Tikhonov term is not involved in the cost functional and pointwise-integral control constraints in time-space are considered. First and second order optimality conditions are established using a suitable extended cone for the sufficient ones. The emphasis of the presentation will be given in the numerical approximation of the control problem, proving the convergence of the discrete problems and establish error estimates between the optimal discrete and continuous states (see, e.g., [1]).

Acknowledgements

The first author was supported by MCIN/ AEI/10.13039/501100011033/ under research project PID2020-114837GB-I00. The second author was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment" grant 3270.

References

[1] E. Casas, and K. Chrysafinos The velocity tracking problem for Navier-Stokes equations with pointwise-integral control constraints in time-space. *submitted.*

Numerical analysis in PDE constrained optimization

Dirichlet control of the Stokes problem towards treatment of non-convex domains

Thomas Apel [1], **Katharina Lorenz**[2]

1. thomas.apel@unibw.de Universität der Bundeswehr München, Germany
2. katharina.lorenz@unibw.de Universität der Bundeswehr München, Germany

For the Dirichlet Boundary Control of the Stokes equation, it must first be discussed how to understand the solution of the Dirichlet problem for the Stokes equations when the Dirichlet data are not smooth, i.e. when they are only in $L^2(\Gamma)$.

A weak solution $(y,p) \in H^1(\Omega)^2 \times L^2(\Omega)$ cannot be expected. Instead, the very weak formulation is considered and a solution is seeked in $L^2(\Omega)^2 \times H^{-1}(\Omega)$. Previous results on that topic are restricted to convex domains where the dual problem has a solution in $H^2(\Omega)^2 \times H^1(\Omega)$ which is not true for non-convex domains. Existence and uniqueness results for the very weak solution are provided.

To obtain a little more regularity than $L^2(\Omega)^2 \times H^{-1}(\Omega)$ the corner singularities are studied such that the approximation error decreases when the mesh size tends to zero.

Numerical analysis in PDE constrained optimization

Regularization and outer approximation for optimal control problems with controls in BV

Christian Meyer[1], **Annika Schiemann**[2]

1. christian2.meyer@tu-dortmund.de Technische Universität Dortmund, Germany
2. annika.schiemann@tu-dortmund.de Technische Universität Dortmund, Germany

We consider a convex elliptic optimal control problem with a constraint on the TV-seminorm of the control. We apply a dual regularization of the TV seminorm and solve the resulting optimization problems with an outer approximation algorithm. We prove the convergence of the algorithm to the global optimal solutions, which in turn converge to the optimal solution to the original problem as the regularization parameter tends to zero. The theoretical findings are confirmed by numerical experiments.

Numerical analysis in PDE constrained optimization

On second order sufficient conditions of a regularized-phase-field fracture control problem

Andreas Hehl[1], **Ira Neitzel**[2]

1. hehl@ins.uni-bonn.de Universität Bonn, Germany
2. neitzel@ins.uni-bonn.de Universität Bonn, Germany

This talk is concerned with an optimal control problem governed by the quasilinear Euler-Lagrange equation of a time-discrete regularized phase-field fracture or damage energy minimization problem, where the control enters as a Neumann boundary force. The problem formulation includes in particular a penalization term for the fracture irreversibility and a viscous regularization. While the latter guarantees unique solvability of the Euler-Lagrange equations under certain conditions since the associated energy minimization problem is then striclty convex, the overall control problem remains nonconvex. We will focus on second order sufficient optimality conditions for the control problem, see [2] for details and further reference. Convergence of the sequential quadratic programming method will be briefly addressed, cf. [1], based on arguments as in e.g. [4] as well as [3].

References

[1] A. Hehl , I. Neitzel, Local quadratic convergence of the SQP method for an optimal control problem governed by a regularized fracture propagation model, *SPP1962 preprint, SPP1962-203*, (2023).

[2] A. Hehl , I. Neitzel, Second-order optimality conditions for an optimal control problem governed by a regularized phase-field fracture propagation model, *Opimization*, 72(6), 1665-1689, (2023).

[3] F. Hoppe , I. Neitzel, Convergence of the SQP method for quasilinear parabolic optimal control problems, *Optimization and Engineering*, 22, 2039-2085, (2020).

[4] F. Tröltzsch. On the Lagrange-Newton-SQP method for the optimal control of semilinear parabolic equations, *SIAM J. Control. Optim.*, **38**(1) (1999), 267–288.

Numerical analysis in PDE constrained optimization

Numerical analysis for Dirichlet optimal control problems on convex polyhedral domains

Johannes Pfefferer[1], **Boris Vexler**[2]

1. johannes.pfefferer@unibw.de *University of the Bundeswehr Munich, Germany*
2. vexler@cit.tum.de *Technical University of Munich, Germany*

This presentation is concerned with the error analysis for finite element discretizations of Dirichlet boundary control problems. In contrast to most of the publications from the literature the underlying domain is assumed to be convex and polyhedral but not only polygonal. For the first time, optimal discretization error estimates are established in this case using the concept of variational discretization or using the approach of full discretzation each based on standard linear finite elements. The convergence rates, which are proven, solely depend on the size of the largest interior edge angle. To be more precise, below the critical angle of $2\pi/3$, a convergence rate of one (times a log-factor) can be achieved for the discrete controls in the L^2-norm on the boundary. For larger interior edge angles the convergence rates are reduced depending on the size of this angle, which is due the impact of singular (domain dependent) terms in the solution. The results are comparable to those for the two dimensional case. However, the theoretical approaches from the two dimensional setting seem not to be directly extendable such that new techniques have to be used. At the end of the talk, the theoretical results are confirmed by numerical experiments. More details can also be found in [1].

References

[1] Johannes Pfefferer, Boris Vexler. Numerical Analysis for Dirichlet Optimal Control Problems on Convex Polyhedral Domains. arXiv:2401.02399 [math.NA], 2024.

Numerical analysis in PDE constrained optimization

Non-coercive Neumann boundary control problems

Thomas Apel [1], **Mariano Mateos** [2], **Arnd Rösch**[3]

1. thomas.apel@unibw.de *Universität der Bundeswehr München, Germany*
2. mmateos@uniovi.es *Universidad de Oviedo, Spain*
3. arnd.roesch@uni-due.de *Universität Duisburg Essen*

We examine a linear-quadratic Neumann control problem that is governed by a non-coercive elliptic equation. Due to the non-self-adjoint nature of the linear control-to-state operator, it is necessary to independently study both the state and adjoint state equations. We establishe the existence and uniqueness of solutions for both equations, with minimal assumptions made about the problem's data. Next, the regularity of these solutions is studied in different types of spaces. These regularity results enable a numerical analysis of the finite element approximation of both the state and adjoint state equations. The results cover both convex and non-convex domains and quasi-uniform and graded meshes. Finally, the optimal control problem is analyzed and discretized. Existence and uniqueness of the solution, first-order optimality conditions, and error estimates for the finite element approximation of the control are obtained. A significant highlight is that the discretization error estimates known from the literature, are improved even for the coercive case.

NUMERICAL ANALYSIS IN PDE CONSTRAINED OPTIMIZATION

Convergence analysis for coefficient control of obstacle problems

Nicolai Simon[1], **Winnifried Wollner**[2]

1. nicolai.simon@uni-hamburg.de Universität Hamburg, Germany
2. winnifried.wollner@uni-hamburg.de Universität Hamburg, Germany

1. Introduction

We consider convergence analysis for a coefficient controlled optimal control problem on a bounded domain $\Omega \subset \mathbb{R}^2$ that is regular in the sense of Gröger and subject to variational inequality constraints. The constraint is an obstacle problem, such that the optimal control problem is given by

$$\begin{aligned} \min_{q\in Q^{\mathrm{ad}},u\in K} J(q,u) &= \frac{1}{2}\|u-u_d\|^2 + \frac{\alpha}{2}\|q\|^2 \\ \text{s.t. } (q\nabla u, \nabla(v-u)) &\geq (f, v-u) \qquad \forall v \in K \end{aligned} \tag{1.1}$$

with $u_d, f \in L^2(\Omega)$. The state is subject to an obstacle $\psi \in \mathbb{R}$ with $\psi < 0$ and the control is subject to bounds $0 < q_{\min} < q_{\max} \in \mathbb{R}$ with

$$\begin{aligned} K &= \left\{v \in H_0^1(\Omega) |\ v \geq \psi \text{ q.e.}\right\} \\ \text{and } Q^{\mathrm{ad}} &= \left\{q \in L^2\left(\Omega, \mathbb{R}^{2\times 2}_{\mathrm{sym}}\right) \,\middle|\, 0 \prec q_{\min} I \preccurlyeq q(x) \preccurlyeq q_{\max} I \text{ a.e.}\right\}. \end{aligned} \tag{1.2}$$

Note that $q_{\min} I \preccurlyeq q$ describes a semidefinite ordering in the sense that $q - q_{\min} I$ is a positive semidefinite matrix.

2. Convergence Analysis

In this talk, we will discuss an approach to convergence analysis where we combine convergence results for a regularization of Problem (1.1) with convergence analysis for the regularized problem, see, e.g., [4] for such an approach. Introducing a regularization function $r(\gamma, u_\gamma)$ into the constraints of (1.1), we get the regularized constraint

$$-\nabla\cdot\left(q_\gamma \nabla u_\gamma\right) + r(\gamma, u_\gamma) = f \qquad \text{in } H^{-1}(\Omega), \tag{2.1}$$

which removes the obstacle and results in a coefficient controlled PDE-constraint. By utilizing existing results for the convergence analysis of such PDE-constrained problems, see [1], in conjunction with convergence results for the regularization of the original problem, see [2], we will discuss convergence results for the obstacle problem with control in the coefficients.
During this talk we will focus on the utilization of H-convergence techniques, see, e.g., [1, 3], to handle coefficient control. The novelty of this talk will be the introduction of a coefficient control in the variational inequality setting.

Acknowledgements

Funded by the Deutsche Forschungsgesellschaft (DFG) - Projektnummer 314067056 within SPP 1962.

References

[1] K. Deckelnick and M. Hinze. Identification of matrix parameters in elliptic PDEs. *Control Cybernet.*, 40(4):957–969, 2011.

[2] A. Hehl, D. Khimin, I. Neitzel, N. Simon, T. Wick, and W. Wollner. Coefficient control of variational inequalities. Preprint 2307.00869, arXiv, 2023. URL: https://doi.org/10.48550/arXiv.2307.00869.

[3] F. Murat and L. Tartar. H-convergence. In A. V. Cherkaev and R. Kohn, editors, *Topics in the mathematical modelling of composite materials*, Modern Birkhäuser Classics, chapter 3. Birkhäuser Cham, 2018.

[4] R. Nochetto. Sharp L^∞-error estimates for semilinear elliptic problems with free boundaries. *Numerische Mathematik*, 54, 1988.

Numerical analysis in PDE constrained optimization

On the bang-bang principle for parabolic optimal control problems

Fredi Tröltzsch
troeltzsch@math.tu-berlin.de Technische Universität Berlin, Germany.

We discuss the construction of parabolic optimal control problems with time-dependent control, such that the optimal control has a desired switching structure. Our main result is that, for each given set of real numbers $0 < s_1 < \ldots < s_k < T$, there is a target state such that the optimal control of the boundary control problem below has the switching points $s_1, \ldots, s_k$. Even though the method is constructive, its numerical application is limited to a small number k of switching points. We present numerical examples for $k = 2, 3$.

In the main part of the talk, the following spatially one-dimensional optimal boundary control problem is considered:

$$\min \int_0^1 |y(x,T) - y_\Omega(x)|^2 dx$$

subject to

$$\begin{aligned} \partial_t y(x,t) - \partial_{xx} y(x,t) &= 0 && \text{in } (0,1)\times(0,T) \\ \partial_x y(0,t) &= 0 && \text{in } (0,T) \\ \partial_x y(1,t) + \lambda\, y(1,t) &= u(t) && \text{in } (0,T) \\ y(x,0) &= 0 && \text{in } (0,1), \end{aligned}$$

and

$$|u(t)| \le 1 \quad \text{a.e. in } (0,T).$$

Here, $T > 0$ and $\lambda > 0$ are given constants, while $y_\Omega \in L^2(0,1)$ is the given target state. Let us set for convenience $\Omega := (0,1)$.

We will also briefly sketch the case of distributed and boundary control in higher dimension. For details, we refer to [1].

References

[1] F. Tröltzsch. On the bang-bang principle for parabolic optimal control problems. *Ann. Acad. Rom. Sci. Ser. Math. Appl.*, 15(1-2):286–307, 2023.

Numerical analysis in PDE constrained optimization

A numerical solution approach for non-smooth optimal control problems based on the Pontryagin maximum principle

Daniel Wachsmuth
daniel.wachsmuth@uni-wuerzburg.de University of Wurzburg, Germany

We consider optimal control problems of the following type:

$$\min \int_\Omega g(x, y(x)) + h(x, u(x))dx$$

subject to the semilinear elliptic partial differential equation

$$-\Delta y + d(y) = u$$

with homogeneous Dirichlet boundary conditions. We assume that d and g give rise to Fréchet differentiable operators, whereas $h : \Omega \times \mathbb{R} \to \bar{\mathbb{R}}$ is merely assumed to be lower semicontinuous. In particular, we do not assume convexity of h.

It is well-known that local solutions satisfy the celebrated Pontryagin maximum principle. In this talk, we will investigate an optimization method that is based on the maximum principle. The discrepancy in the maximum principle vanishes along the resulting sequence of iterates. Numerical experiments confirm the theoretical findings.

Numerical analysis in PDE constrained optimization

SQP method for hyperbolic PDE-constrained optimization in acoustic full waveform inversion

Luis Ammann[1], **Irwin Yousept**[2]

1. luis.ammann@uni-due.de Universität Duisburg-Essen, Fakultät für Mathematik, Thea-Leymann-Str. 9, D-45127 Essen, Germany
2. irwin.yousept@uni-due.de Universität Duisburg-Essen, Fakultät für Mathematik, Thea-Leymann-Str. 9, D-45127 Essen, Germany

This talk presents recent results on the SQP method for hyperbolic PDE-constrained optimization in acoustic full waveform inversion. The analysis of the SQP method is mainly challenging due to the involved hyperbolicity and second-order bilinear structure. This notorious character leads to undesired effects of regularity loss in the SQP iteration calling for a substantial extension of developed parabolic techniques. We propose and explore a novel strategy for the well-posedness and convergence analysis of the SQP method based on the use of a smooth-in-time initial condition, a tailored self-mapping operator, and a two-step estimation process along with Stampacchia's method. Our final theoretical result is the R-superlinear convergence of the SQP method.

Numerical methods and software for solving optimal control problems

Control-toolbox: solving control problems within Julia

Jean-Baptiste Caillau[1], **Olivier Cots**[2], **Joseph Gergaud**[3], **Pierre Martinon**[4]

1. jean-baptiste.caillau@univ-cotedazur.fr Université Côte d'Azur, CNRS, Inria, LJAD
2. olivier.cots@univ-toulouse.fr Université de Toulouse, CNRS, ENSEEIHT-IRIT
3. joseph.gergaud@univ-toulouse.fr Université de Toulouse, CNRS, ENSEEIHT-IRIT
4. pierre.martinon@inria.fr Inria Paris center

There is a strong trend to use Julia in scientific computing, so as to take advantage not only of the performance but also of the high level traits of the language. These features allow to cast problems and algorithms in a form close to their mathematical definitions. As powerful Julia libraries to solve ODE's and optimisation problems are now available, it is possible to attack efficiently optimal control problems. Several methods will be presented, including direct transcription and shooting. While the first approach consists in a brutal approximation of the infinite dimensional control problem by a nonlinear program with sparse constraints, the second one leverages Pontrjagin maximum principle to ensure a very precise computation of optimal controls. Rather than competing methods, these two approaches must be made to collaborate: direct codes capture the structure of the solution (typically made of bang and singular arcs), which then allows to devise and initialise a tailored shooting function. These points will be illustrated in the framework of ongoing developments of Julia packages from the ct: `control-toolbox` suite.

Acknowledgements

The first author was partially supported by grant no. ANR-23-PEIA-0004.

Numerical methods and software for solving optimal control problems

Solving mixed-integer optimal control problems with a discretize-then-optimize approach using CasADi with an application to preparative chromatography

Dominik H. Cebulla[1], **Christian Kirches**[2], **Andreas Potschka**[3]

1. d.cebulla@tu-bs.de *Institute for Mathematical Optimization, Technische Universität Braunschweig, Germany*
2. c.kirches@tu-bs.de *Institute for Mathematical Optimization, Technische Universität Braunschweig, Germany*
3. andreas.potschka@tu-clausthal.de *Institute of Mathematics, Clausthal University of Technology, Germany*

In this talk, we provide a live demonstration on how mixed-integer optimal control problems (MIOCPs) can be solved with a discretize-then-optimize approach using CasADi—a tool for nonlinear optimization and algorithmic differentiation.

Mathematically, we first transform the MIOCP in a continuous optimal control problem (OCP) by means of partial outer convexification and relaxation. The obtained OCP is then transformed into a nonlinear program (NLP) with, e.g., direct shooting or collocation methods. With a solution of the NLP at hand, we apply the so-called sum-up rounding strategy to obtain δ-feasible and δ-optimal controls with respect to the original mixed-integer formulation.

Using an illustrative example, we demonstrate how the above mentioned steps can be easily implemented using CasADi. To also underline the real-world applicability of the presented procedure, we briefly show results from the optimization of a real-world chromatographic separation process, where the underlying dynamics are given by a parabolic partial differential equation with high nonlinearities. Even in this challenging setting, the presented methodology works satisfactorily.

References

[1] H. G. Bock, D. H. Cebulla, C. Kirches, and A. Potschka. Mixed-integer optimal control for multimodal chromatography. *Computers & Chemical Engineering*, 153:107435, 2021.

[2] D. H. Cebulla, C. Kirches, N. Kümmerer, and A. Potschka. Model-based optimization of an ion exchange chromatography process for the separation of von Willebrand factor fragments and human serum albumin. In *Proceedings of the 93rd Annual Meeting of the International Association of Applied Mathematics and Mechanics*, page e202300027, 2023.

NUMERICAL METHODS AND SOFTWARE FOR SOLVING OPTIMAL CONTROL PROBLEMS

WORHP Lab: teaching and showcasing numerical methods of optimization and optimal control

Christof Büskens[1], **Matthias Knauer**[2]

1. bueskens@uni-bremen.de Universität Bremen, Germany
2. knauer@uni-bremen.de Universität Bremen, Germany

In a wide range of industrial applications, modeling and simulation are used to better understand processes. In order for the industry to take the next step towards optimization and recognize their benefits, a lot of educational work is still needed. With the graphical user interface WORHP Lab, we develop an easy-to-use software tool to explain concepts of optimization and optimal control to industrial partners. In our MINT activities with schools we learned, that also students can work with the software to learn about useful applications of mathematics.

1. Optimization

The ESA NLP solver WORHP [1] was designed to solve nonlinear programming problems using an SQP method or an interior point method. WORHP exploits the sparsity patterns of the occurring derivative matrices of the user-defined functions (for objective and constraints). Typically for structured problems each function only depends on few variables, and hence large numbers of variables and constraints can be considered.

After a successful optimization with WORHP, a parametric sensitivity analysis can be performed using the module WORHP Zen [4] to find out, how the solution would change if the problem is changed slightly.

2. Optimal Control

The transcription method TransWORHP [2] was developed to solve general optimal control problems using direct methods. After a full discretization (with Trapezoidal method, Hermite-Simpson, etc.) or application of a multiple shooting method the problem is solved using WORHP, taking into account the problem specific derivative structures and values, if available. In this way we can also receive the valuable sensitivity information for the discretized optimal control problem.

3. Graphical User Interface

The graphical user interface WORHP Lab allows an easy access to formulate problems for the solvers WORHP and TransWORHP. Optimization variables can be declared and ranges can be provided easily. The functions for the objective and the constraints can be provided in the interface as C++ code, which will be converted to executable code during run-time for faster evaluations.

The optimization output and the results (including sensitivity derivatives) can be analyzed and visualized easily. In the context of TransWORHP, techniques for an adaptive grid refinement or model predictive control [3] can be tried out.

The software (including various versions of training material) was already used successfully in various lectures, industrial workshops, and STEM activities.

References

[1] Christof Büskens and Dennis Wassel. The ESA NLP Solver WORHP. In Giorgio Fasano and János D. Pintér, editors, *Modeling and Optimization in Space Engineering*, volume 73 of *Springer Optimization and Its Applications*, pages 85–110. Springer, 2012. doi:10.1007/978-1-4614-4469-5_4.

[2] Matthias Knauer and Christof Büskens. Real-Time Optimal Control Using TransWORHP and WORHP Zen. In Giorgio Fasano and János D. Pintér, editors, *Modeling and Optimization in Space Engineering*, volume 144 of *Springer Optimization and Its Applications*, pages 211–232. Springer, 2019. doi:10.1007/978-3-030-10501-3_9.

[3] Matthias Knauer and Christof Büskens. Processing User Input in Tracking Problems using Model Predictive Control. *IFAC-PapersOnLine*, 50(1):9846–9851, 2017. 20th IFAC World Congress. doi:10.1016/j.ifacol.2017.08.905.

[4] Renke Kuhlmann, Sören Geffken, and Christof Büskens. WORHP Zen: Parametric Sensitivity Analysis for the Nonlinear Programming Solver WORHP. In Natalia Kliewer, Jan Fabian Ehmke, and Ralf Borndörfer, editors, *Operations Research Proceedings 2017*, pages 649–654. Springer, 2018. doi:10.1007/978-3-319-89920-6_86.

NUMERICAL METHODS AND SOFTWARE FOR SOLVING OPTIMAL CONTROL PROBLEMS

Optimize-then-discretize interior-point methods in optimal control: convergence results and primal-dual implementation

Paul Malisani

paul.malisani@ifpen.fr IFP Energies nouvelles, Applied Mathematics Dpt, Rueil-Malmaison, France

1. Introduction

Interior point methods and their primal-dual implementation are widely and successfully used to solve optimal control problems in a first-discretize-then-optimize fashion. However, their adaptation to the first-optimize-then-discretize framework is not straightforward. In recent contributions [2, 3], we established a proof of convergence of function space interior-point methods for a general class of optimal control problems. This convergence proof allows for adapting primal-dual methods in the first-optimize-then-discretize framework. The primal-dual optimal control solving algorithm solves differential and algebraic equations with initial and final conditions. The proposed method is fast, accurate, and easy to use.

2. Some details

The problem we are interested in consists in finding a solution (x, u) of the following Constrained Optimal Control Problem (COCP)

$$\min_{(u,x)} J(x,u) := \varphi(x(T)) + \int_0^T \ell(x(t),u(t))\mathrm{d}t \tag{2.1a}$$
$$\dot{x}(t) = f(x(t),u(t)) \tag{2.1b}$$
$$0 = h(x(0),x(T)) \tag{2.1c}$$
$$0 \geq g(x(t)) \tag{2.1d}$$
$$0 \geq c(x(t),u(t)) \tag{2.1e}$$

The main convergence result for primal-dual methods in the first-optimize-then-discretize framework is as follows.

Theorem 2.1 *Let $(\epsilon_n)_n$ be a sequence of decreasing positive parameters with $\epsilon_n \to 0$ and let $(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}, \bar{p}_{\epsilon_n}, \bar{\lambda}^g_{\epsilon_n}, \bar{\lambda}^c_{\epsilon_n}, \bar{\lambda}_{\epsilon_n})_n$ be a solution of*

$$\dot{\bar{x}}_{\epsilon_n}(t) = f(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)) \tag{2.2a}$$
$$\dot{\bar{p}}_{\epsilon_n} = -\frac{\partial H}{\partial x}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t), \bar{p}_{\epsilon_n}(t)) - \frac{\partial g}{\partial x}(\bar{x}_{\epsilon_n}(t))^\top.\bar{\lambda}^g_{\epsilon_n}(t) - \frac{\partial c}{\partial x}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t))^\top.\bar{\lambda}^c_{\epsilon_n}(t) \tag{2.2b}$$
$$0 = \frac{\partial H}{\partial u}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t), \bar{p}_{\epsilon_n}(t)) + \frac{\partial c}{\partial u}(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t))^\top.\bar{\lambda}^c_{\epsilon_n}(t) \tag{2.2c}$$
$$0 = \mathrm{FB}(\bar{\lambda}^g_{\epsilon_n}(t), g(\bar{x}_{\epsilon_n}(t)), \epsilon_n) \tag{2.2d}$$
$$0 = \mathrm{FB}(\bar{\lambda}^c_{\epsilon_n}(t), c(\bar{x}_{\epsilon_n}(t), \bar{u}_{\epsilon_n}(t)), \epsilon_n) \tag{2.2e}$$
$$0 = h(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T)) \tag{2.2f}$$
$$0 = \bar{p}_{\epsilon_n}(0) + \frac{\partial h}{\partial x(0)}(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T))^\top.\bar{\lambda}_{\epsilon_n} \tag{2.2g}$$
$$0 = \bar{p}_{\epsilon_n}(T) - \nabla\varphi(\bar{x}_{\epsilon_n}(T)) - \frac{\partial h}{\partial x(T)}(\bar{x}_{\epsilon_n}(0), \bar{x}_{\epsilon_n}(T))^\top.\bar{\lambda}_{\epsilon_n} \tag{2.2h}$$

where $\mathrm{FB} : \mathbb{R}\times\mathbb{R}, \times\mathbb{R}_+ \mapsto \mathbb{R}_+$ is the Fisher-Burmeister complementarity function defined as follows

$$\mathrm{FB}(x,y,\epsilon) := x - y - \sqrt{x^2+y^2+2\epsilon} \tag{2.3}$$

Then $(\bar{x}_{\epsilon_n}, \bar{u}_{\epsilon_n}, \bar{p}_{\epsilon_n}, \bar{\lambda}^g_{\epsilon_n}, \bar{\lambda}^c_{\epsilon_n}, \bar{\lambda}_{\epsilon_n})_n$ contains a subsequence converging to a point satisfying the first-order optimality conditions [1] for problem (2.1)*.*

References

[1] J. F. Bonnans, X. Dupuis, and L. Pfeiffer. Second-order necessary conditions in pontryagin form for optimal control problems. *SIAM Journal on Control and Optimization*, 52(6):3887–3916, 2014. URL: `https://hal.inria.fr/hal-00825273/document`.

[2] P. Malisani. Interior point methods in optimal control problems for affine systems: Convergence results and solving algorithms. *SIAM Journal on Control and Optimization*, 61(6), 2023.

[3] Paul Malisani. Interior point methods in optimal control, 2023. `arXiv:2309.01425`.

Optimal control and machine learning: recent advances

Neural networks for differential games

Olivier Bokanowski[1], **Xavier Warin**[2]

1. olivier.bokanowski@u-paris.fr Université Paris Cité, France
2. xavier.warin@edf.fr EDF R&D, France

We study deterministic optimal control problems for differential games with finite horizon [2]. We propose new approximations of the strategies in feedback form and show error estimates and a convergence result of the value in some weak sense for one of the formulations. This result applies in particular to neural network approximations. This work follows some ideas introduced in Bokanowski, Prost and Warin [1] for deterministic optimal control problems, yet with a simplified approach for the error estimates, which allows to consider a global optimization scheme instead of a time-marching scheme. We also give a new approximation result between the continuous and the semi-discrete optimal control value in the game setting, improving the classical convergence order $O(\Delta t^{1/2})$ to $O(\Delta t)$ (where Δt is the time step) under some assumptions on the dynamical system. In order to validate the approach, numerical examples are shown on some academic two-player game problems related to backward reachability, in presence of state constraints, using the level approach to represent the regions of interests. Stochastic gradient-type algorithms are used to deal with the minimax problem.

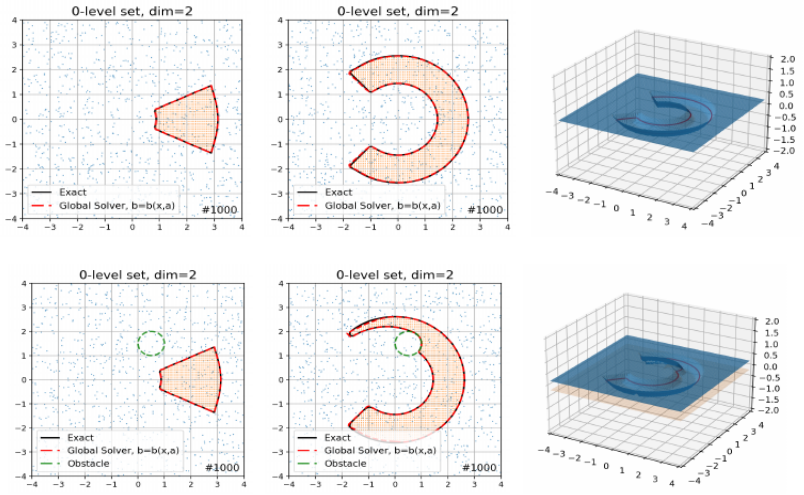


Fig. 1 A very elementary example of a controlled front propagation with an adverse control (left: initial data, center/right: reachable region after some time evolution) Up: no obstacle. Down: with disk obstacle.

Acknowledgements

This research benefited from the support of the FMJH Program PGMO and from the support of EDF.

References

[1] O. Bokanowski, A. Prost, and X. Warin. Neural networks for first order HJB equations and application to front propagation with obstacle terms. *Partial Differ. Equ. Appl.*, 4:45, 2023.

[2] O. Bokanowski and X. Warin. Neural networks for differential games, 2024. arXiv:2402.02792.

OPTIMAL CONTROL AND MACHINE LEARNING: RECENT ADVANCES

Stabilization of semilinear parabolic equations: explicit and optimal control based feedbacks

Karl Kunisch[1], **Sérgio S. Rodrigues**[2*], **D. Walter**[3]

1. karl.kunisch@uni-graz.at Karl-Franzens University of Graz, Austria, and RICAM of Austrian Academy of Sciences, Linz, Austria
2. sergio.rodrigues@oeaw.ac.at RICAM of Austrian Academy of Sciences, Linz, Austria
3. daniel.walter@hu-berlin.de Humboldt University, Berlin, Germany

1. Introduction

The construction of stabilizing feedback controls for nonautonomous parabolic-like equations is addressed. The discussion includes stabilizing feedbacks either given explicitly, based on suitable oblique projections, or given by operators constructed by using optimization tools.

2. Controlled nonautonomous parabolic-like systems

We are given a control system as

$$\dot{y} = -Ay - N(y) - Bu, \qquad y(0) = y_0 \in H,$$

with state $y(t) \in H$ (where H is a Hilbert space) and control input $u(t) \in \mathbb{R}^M$, at time $t \geq 0$. The operator A is a linear diffusion-like time-independent operator, and $N(y) = N(t, y(t))$ is a time-dependent, possibly nonlinear, reaction–convection-like operator.

The control force $Bu(t) = \sum_{j=1}^{M} u_j(t)\Phi_j$ is a linear combination of a finite number M of apriori given linearly independent actuators $\Phi_j \in H$, $1 \leq j \leq M$.

We assume that the free dynamics (i.e., with $u = 0$) is not stable. The goal is to find the control input in feedback form $u(t) = K(t, y(t))$, so that the resulting system

$$\dot{y} = -Ay - N(y) - BK(y), \qquad y(0) = y_0 \in H,$$

is stable. In other words we are looking for a stabilizing feedback control operatot K.

3. Strategy

For a general class of nonlinearities, we show how the feedback control input can be taken in the form $u(t) = K(P_M y(t))$, depending on an M-dimensional component $z(t) = P_M y(t)$ of the state, for example, the orthogonal projection of $y(t)$ onto the linear span $\mathcal{U}_M$ of the set of actuators.

In this case we have that $K := \mathcal{U}_M \to \mathbb{R}^M$ can be identified with a mapping $K := \mathbb{R}^M \to \mathbb{R}^M$.

We show firstly that K can be found in explicit form. Then, we address the problem of finding an optimal $\overline{K}$ among (nonlinear) functions $K := \mathbb{R}^M \to \mathbb{R}^M$, where we use techniques inspired by machine learning. Results of simulations are presented showing the stabilizing performance of the proposed feedback controls [1].

Appendix: detectability, observer design

We show how this type of explicit feedback operators can be used as an output injection operator in the context of observer design (i.e., state estimation; continuous data assimilation), and present results of corresponding simulations [2].

Acknowledgements

S. Rodrigues thanks the partial support from Austrian Science Fund (FWF) P 33432-NBL.

References

[1] K. Kunisch, S. S. Rodrigues, and D. Walter. Learning an optimal feedback operator semiglobally stabilizing semilinear parabolic equations. *Appl. Math. Optim.*, 84(1):277–318, 2021. doi:10.1007/s00245-021-09769-5.

[2] S. S. Rodrigues. Oblique projection output-based feedback exponential stabilization of nonautonomous parabolic equations. *Automatica J. IFAC*, 129:109621, 2021. doi:10.1016/j.automatica.2021.109621.

Optimal control and machine learning: recent advances

Neural network-based approximation of optimal value functions: exploring decaying sensitivity

Lars Grüne[1], **Mario Sperl**[2]

1. lars.gruene@uni-bayreuth.de Universität Bayreuth, Germany
2. mario.sperl@uni-bayreuth.de Universität Bayreuth, Germany

We consider an infinite horizon optimal control problem of the form

$$\begin{aligned} &\text{minimize} \quad J(x_0,u) = \int_0^\infty e^{-\delta t}\ell(x(t),u(t))dt, \\ &\text{subject to} \quad \dot{x}(t) = f(x(t),u(t)), \quad x(0) = x_0. \end{aligned} \tag{1}$$

For such a problem, an (approximately) optimal feedback control $u^*(t) = F(x^*(t))$ can be computed from (an approximation of) the optimal value function $V(x_0) := \inf_u J(x_0,u)$. However, conventional grid- or mesh-based numerical methods encounter the curse of dimensionality for the computation of V as solution of the Hamilton-Jacobi-Bellman equation. This results in an exponential growth of the computational effort in the dimension of the state x, making the methods impractical for higher dimensions. In this presentation, we show that neural networks are able to mitigate the curse of dimensionality for interconnected optimal control problems exhibiting a decaying sensitivity property among their subsystems.

Initially, we introduce a feedforward neural network architecture, demonstrated in [1], which provably only requires a polynomial growth of neurons in the dimension n to approximate so-called d-separable functions. A function $F\colon \mathbb{R}^n \to \mathbb{R}$ is called d-separable if it can be written as $F(x) = \sum_{j=1}^s F_j(z_j)$, where $z_j \in \mathbb{R}^{d_j}$ represent lower-dimensional components of the whole state x with $d_j \leq d$ for all $j = 1,\dots,s$. We then consider an optimal control problem (1) that can be decomposed into s subsystems whose interconnection is expressed via a directed graph. For such a problem we propose a decaying sensitivity assumption stating that the sensitivity between two subsystems on the optimal value function is decreasing with an increasing distance of the subsystems in the graph. In the context of linear quadratic problems with $V(x) = x^TPx$ for some $P \in \mathbb{R}^{n\times n}$, this sensitivity assumption corresponds to a decrease of the entries in P with an increasing graph distance between the respective subsystems and has been utilized in [2] for approximating the optimal feedback matrix. In this presentation, we introduce a decaying sensitivity formulation tailored for nonlinear problems and demonstrate how to leverage overlapping neighborhoods in the graph to construct a separable approximation of the optimal value function, as discussed in [3]. Furthermore, we provide error bounds based on the type of decay and the structure of the underlying graph. Altogether, this establishes a condition for interconnected optimal control problems that enables avoiding the curse of dimensionality using neural networks. To provide empirical validation of our theoretical framework, we present results from a numerical test case.

Acknowledgements

This research has been supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project number 463912816 within the priority program 441826958.

References

[1] Lars Grüne. Computing Lyapunov functions using deep neural networks. *Journal of Computational Dynamics*, 8(2):131–152, 2021.

[2] Sungho Shin, Yiheng Lin, Guannan Qu, Adam Wierman, and Mihai Anitescu. Near-optimal distributed linear-quadratic regulator for networked systems. *SIAM Journal on Control and Optimization*, 61(3):1113–1135, 2023.

[3] Mario Sperl, Luca Saluzzi, Lars Grüne, and Dante Kalise. Separable approximations of optimal value functions under a decaying sensitivity assumption. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 259–264, 2023.

OPTIMAL CONTROL AND MACHINE LEARNING: RECENT ADVANCES

Controllability of neural ODEs for data classification

Antonio Álvarez-López

antonio.alvarezl@uam.es Universidad Autónoma de Madrid, Spain

1. Introduction

Neural ordinary differential equations (ODEs) have emerged as a natural tool for data-driven supervised learning, particularly for modeling dynamical processes. They represent the continuous-time limit of neural networks:

$$\dot{x} = \sum_{i=1}^{p} w_i(t)\sigma(a_i(t) \cdot x + b_i(t)), \qquad t \in (0,T), \tag{1.1}$$

where $x \in \mathbb{R}^d$, $\theta_i := (w_i, a_i, b_i) \in L^\infty((0,T); \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R})$ are piecewise constant controls, $\sigma : \mathbb{R}^p \to \mathbb{R}^p$ is the component-wise ReLU function, defined for each component by $\sigma(z) = \max\{z, 0\}$ for $z \in \mathbb{R}$, $p \geq 1$ is the *width*, and the number of discontinuities $L \geq 1$ of every θ_i is the *depth*. One of the main advantages that neural ODEs offer is the possibility to reinterpret several machine learning paradigms from a control perspective using the tools from differential equations. For example, the problem of data classification can be understood in terms of the property of simultaneous controllability that the system (1.1) exhibits to interpolate N points in $\mathbb{R}^d$.

2. Results

In this work, I present two results in this direction. They focus on identifying the optimal architecture of the model described by (1.1), determined by its depth and width, and on measuring its finite-sample expressivity. First, we estimate the required number of neurons for efficient cluster-based classification, especially in the worst-case scenario where points are independently and uniformly distributed in $[0,1]^d$. It has been shown in [3] that this task can be accomplished using $O(N)$ neurons. In [1], we propose an algorithm that classifies clusters of d points from any initial configuration, provided they are in general position, resulting in a complexity of $O(N/d)$ neurons.

Second, we explore in [2] the interplay between the width p and depth L in interpolating a dataset of N pairs of points. Our findings reveal a balancing trade-off, with L scaling as $O(1 + N/p)$. In the autonomous case, where $L = 0$, a separate study is required. We address the relaxed problem of ε-approximate controllability of N pairs of points and establish an error decay of $\varepsilon \sim O(\log(p)p^{-1/d})$. This decay rate results from applying a universal approximation theorem to a custom-built Lipschitz vector field that interpolates $\mathcal{D}$. In the high-dimensional setting, we further demonstrate that $p = O(N)$ neurons are likely sufficient to achieve exact control. To conclude, we consider the natural extension of the problem to ε-approximate control of two measures in the Wasserstein-1 space, obtaining $L = O\left(1 + (p\varepsilon^d)^{-1}\right)$ discontinuities.

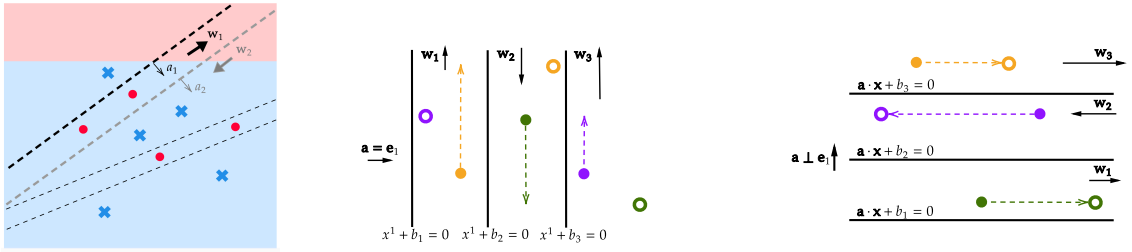


Fig. 1 Left: Algorithm for cluster-based classification. Right: Algorithm for N-point interpolation (in two steps).

Acknowledgements

The work presented was supported by the Madrid Government (Comunidad de Madrid – Spain) under the multiannual Agreement with UAM in the line for the Excellence of the University Research Staff in the context of the V PRICIT (Regional Programme of Research and Technological Innovation). The first author has been also funded by a contract FPU21/05673 from the Spanish Ministry of Universities.

References

[1] A. A-L, A. Hadj-Slimane, and E. Zuazua. Depth vs width for interpolation in neural odes. 2024. arXiv:arXiv:2401.09902.

[2] A. A-L, R. Orive-Illera, and E. Zuazua. Optimized classification with neural odes via separability. 2023. arXiv:2312.13807.

[3] D. Ruiz-Balet and E. Zuazua. Neural ODE control for classification, approximation, and transport. *SIAM Rev.*, 65(3):735–773, 2023.

Optimal control of ODEs: theory and applications

Existence theorem for relaxed control problems on infinite time horizon utilizing weight functions

Ilya Dikariev[1], **Valeriya Lykina**[2], **Sabine Pickenhain**[3]

1. dikarill@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany
2. Valeriya.Lykina@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany
3. sabine.pickenhain@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany

In this paper, we consider an optimal control problem in Lagrange form on an unbounded time interval without convexity assumptions. A relaxation method of Gamkrelidze [2] is adapted to the infinite time horizon and weighted Sobolev spaces are used as state spaces (compare [1]):

$$\begin{aligned} J(x,\mu) &= \int_\Omega \langle \mu_t, r(t,x(t),\mathrm{v})\rangle d\zeta \longrightarrow Min \\ x &\in W_p^{1,n}(\Omega,\nu) \\ \mu &\in \mathcal{M}_U,\ U \subseteq \mathbb{R}^m \\ \dot{x}(t) &= \langle \mu_t, f(t,x(t),\mathrm{v})\rangle \text{ a.e. on } \Omega, \\ x(t_0) &= x_0,\ t_0 \in \overline{\Omega} \end{aligned}$$

The product topology consisting of the weak topology on Sobolev space and the tight topology on a measure space is used. To obtain the closedness of the admissible set, one type of growth condition is used, in which a continuous weight function ψ acts as a decay factor for the right-hand sides of the state equations:

$$f(t,\xi,\mathrm{v})\psi^{-1}(\mathrm{v}) \in C_b(K \times \mathbb{R}^n \times U),\ \forall K \in \mathrm{comp}(\Omega).$$

The compactness of the admissible set is achieved by a variant of the Prokhorov condition.

Acknowledgements

The authors are partially supported by the German Research Foundation.

References

[1] V. Lykina and S. Pickenhain Weighted functional spaces approach to infinite horizon optimal control problems: a systematic analysis of hidden opportunities and advantages. *Journal of Mathematical Analysis and Applications,* 454(1), pp. 195 – 218, 2017

[2] R.V. Gamkrelidze Principles of Optimal Control Theory. Plenum Press, New York and London, 1978

Optimal control of ODEs: theory and applications

On competition for spatially distributed resources in networks

Giorgio Fabbri[1], **Silvia Faggian**[2], **Giuseppe Freni**[3]

1. giorgio.fabbri@univ-grenoble-alpes.fr Univ. Grenoble Alpes, CNRS, INRIA, Grenoble INP, GAEL, Grenoble, France
2. faggian@unive.it Università "Ca' Foscari" Venezia, Italy
3. giuseppe.freni@uniparthenope.it Department of Business and Economics, University of Naples Parthenope, Naples, Italy

We study the dynamics of the exploitation of a natural resource, distributed in space and mobile, where spatial diversification is introduced by a via a weighted, directed network. The network represents both the locations and the interactions of the resource nodes. A regulator decides to designate some of the nodes as natural reserves where no exploitation is allowed. The remaining nodes are assigned (one-to-one) to players, who will exploit the resource at the node, maximizing their own utility.

We show existence of a Nash equilibrium, as well as its uniqueness among linear Markovian equilibria, providing an explicit formula for the solution.

We show how the Nash equilibrium and the resource stocks depend on the productivity of the resource sites, on the structure of the connections between the sites, and on the number and the preferences of the agents. The best locations to host nature reserves are identified according to the model's parameters, and we find that they correspond to the most central (in the sense of eigenvector centrality) nodes of a suitably redefined network that considers the nodes' productivity.

A variety of possible generalizations and relative applications are also analyzed.

References

[1] Giorgio Fabbri, Silvia Faggian, Giuseppe Freni. On competition for spatially distributed resources on networks. *Theor. Econ.*, in press.

[2] Giorgio Fabbri, Silvia Faggian, Giuseppe Freni. *Growth Models with Externalities on Networks*. Working Papers from Department of Economics, University of Venice "Ca' Foscari", vol. 23/WP/2023, pp. 1-18 (ISSN 1827-3580)

[3] Giorgio Fabbri, Silvia Faggian, Giuseppe Freni. Policy effectiveness in spatial resource wars: a two-region model. *J Econ. Dyn. Control*, vol. 111, pp. 1-30 (ISSN 0165-1889).

Optimal control of ODEs: theory and applications

On the use of Henstock-Kurzweil integral in optimal control problems

Valeriya Lykina

valeriya.lykina@b-tu.de *Brandenburg University of Technology Cottbus-Senftenberg, Cottbus, Germany*

1. Introduction

In this talk we consider a class of infinite horizon optimal control problems involving an integral functional and ordinary differential equations describing the dynamics of the process. The integral is interpreted in the sense of Henstock-Kurzweil, see [1], [2]. The chosen interpretation of the integral results from unsatisfactory points in using Lebesgue or improper Riemann integrals for handling control problems with infinite horizon, cf. [3]. Due to the fact that Henstock-Kurzweil integrals generalize both Lebesgue and Riemann integrals, it is possible to formulate optimality results for more general class of optimal control problems. Here, we discuss the advantages and challenges of using the HK integral in the problem setting.

2. Some details

It is considered the infinite horizon control problem of minimizing the integral objective

$$J_{\infty}^{HK}(x,u) = HK\text{-}\int_0^{\infty} r(\cdot,x(\cdot),u(\cdot))\,\mathfrak{v}(\cdot) \tag{2.1}$$

with respect to all pairs

$$(x,u) \in HK^{1,n}(\mathbb{R}^+,\nu) \times HK^{m}(\mathbb{R}^+,\nu)\,; \tag{2.2}$$

governed by differential equation of the form

$$\dot{x}(t) = f(t,x(t),u(t)) \quad \text{a.e. on } \mathbb{R}^+\,; \tag{2.3}$$

$$x(0) = x_0\,; \tag{2.4}$$

and satisfying the control constraints

$$u(t) \in U \quad \text{a.e. on } \mathbb{R}^+ \tag{2.5}$$

Hereby U denotes a compact convex subset of $\mathbb{R}^m$, ν and $\mathfrak{v}$ are weight functions.

A first result concerning sufficient optimality conditions for the new class of optimal control problems is obtained. Relations between admissible sets and optimal solutions of the new control problem and the problems involving Lebesgue or improper Riemann integrals are discussed by means of an example. The applicability of sufficient optimality conditions is also shown.

References

[1] Kurtz, D. S.; Swartz, C. W. Theories of Integration. The Integrals of Riemann, Lebesgue, Henstock-Kurzweil, and McShane. *World Scientific; New Jersey-London-Singapore-etc, 2005.*

[2] Lee Peng-Yee. Langzhou Lectures on Henstock Integration *World Scientific, 1989.*

[3] Lykina, V.; Pickenhain, S.; Wagner, M. Different interpretations of the improper integral objective in an infinite horizon control problem. *Mathematical Analysis and Applications: Vol 340, pp. 498 – 510, 2008.*

Optimal control of ODEs: theory and applications

TV regularization for integer optimal control problems with vector valued controls

Gerd Wachsmuth[1], **Jonas Marko**[2]

1. wachsmuth@b-tu.de BTU Cottbus-Senftenberg, Germany
2. Jonas.Marko@b-tu.de BTU Cottbus-Senftenberg, Germany

1. Overwiev

Integer optimal control problems with vector-valued controls pose challenges in ensuring solution existence and preventing undesirable chattering behavior. A common tool to address these issues effectively is the total variation functional. Our approach involves adapting the total variation functional to accommodate any p-vector norm, providing greater flexibility in shaping the solution properties.

We observe optimality conditions of first and second order via a switching time optimization problem. Furthermore, we present a trust region algorithm grounded in Bellman's optimality principle, which offers a systematic framework for tackling such problems. The efficacy of our method is illustrated by two ODE-constrained examples from the benchmark library `mintoc.de`.

2. Problem Formulation and details

We investigate problems of the form

$$\begin{aligned} &\text{Minimize} && F(u) + \beta\,\mathrm{TV}_p(u) \\ &\text{such that} && u(t) \in \mathcal{V} \subset \mathbb{Z}^M \quad \text{for a.a. } t \in (0,T). \end{aligned} \tag{P}$$

Here, $u \in \mathrm{BV}(0,T)^M$, $\beta > 0$ and the set $\mathcal{V}$ is assumed to be finite. The term TV_p is defined by

$$\mathrm{TV}_p(u,J) := \sup\left\{ \int_J u^\top \varphi' \,\mathrm{d}t \,|\, \varphi \in C_c^1(J)^M,\ \|\varphi(t)\|_{p'} \le 1 \ \forall t \in J \right\}$$

where p' is the Hölder conjugate of p and $J \subset (0,T)$ is open. Also, we write $\mathrm{TV}_p(u) := \mathrm{TV}_p(u,(0,T))$.

Finally, F is assumed to be lower semicontinuos, bounded from below and Gâteaux-differentiable. It might contain the solution operator of a differential equation, i.e. $F(u) = \hat{F}(S(u),u)$, where S is a control-to-state operator, e.g. mapping u to a state satisfying some ODE or PDE.

Optimal control of ODEs: theory and applications

Pareto solutions in optimal control of a global model of climate change

Helmut Maurer
helmut.maurer@uni-muenster.de Universität Münster, Germany

The Paris 2015 agreement on climate change is aiming at reducing the temperature increase to below $2^o\,C$. This implies that effective mitigation policies need to be pursued that not only prevent the CO_2 emission from rising further but reduce the annual emission substantially. The modeling strategy presented in Atolia, Loungani, Maurer, Semmler [1] attempts to answer three questions: First, what are the best strategies to keep the CO_2 emission bounded by a predefined upper bound. Second, what resources should be allocated to the adaptation effort when climate risk, due to a lack of emission reduction, is rising and future economic, social, and ecological damages can be expected. A third issue is of how the efforts of mitigation and adaptation are funded and how the funds should efficiently be allocated between traditional infrastructure investment, mitigation and adaptation efforts.

The control model [1] has 5 state and 8 control variables that allows to consider the specific policies of infrastructure investment, mitigation and adaptation. A numerical challenge arises from the fact that the optimal control model involves a nonlinear mixed control-state constraint. Optimal control policies for various initial conditions and terminal constraints are obtained via discretization and nonlinear programming methods. The difficulty of determining appropriate weights in the cost functional measuring, eg., the negative externality of the use of brown energy leads us to consider two different sets of parameters in the welfare functionals. We use the numerical methods developed in Kaya, Maurer [2,3] to determine the Pareto front of optimal solutions.

References

[1] Manoj Atolia, Prakash Loungani, Helmut Maurer, Willi Semmler. Optimal control of a global model of climate change with adaptation and mitigation. *Mathematical Control and Related Fields*, 13(2): 583–604, 2023. Doi: 10.3934/mcrf.2022009 .

[2] Yalcin C. Kaya, Helmut Maurer. Optimization over the Pareto front of nonconvex multi-objective optimal control problems. *Computational Optimization and Applications*, 2024, doi: 10.1007/s10589-023-00535-7

[3] Yalcin C. Kaya, Helmut Maurer. A numerical method for nonconvex multi-objective optimal control problems, *Computational Optimization and Applications*, 57:685–702, 2014. DOI 10.1007/s10589-013-9603-2, 2013.

Optimal control of ODEs: theory and applications

Duality for infinite horizon relaxed control problems

Sabine Pickenhain[1], **Ilya Dikariev**[2], **Valeriya Lykina**[3]

1. sabine.pickenhain@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany
2. dikarill@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany
3. Valeriya.Lykina@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany

We consider relaxed control problems whose objective functional is in the economic context a discounted utility functional or an energy functional in mechanical or quantum mechanical systems. The objective functional can also be chosen in such a way that the asymptotic controllability of the system is guaranteed.

The following problem $(\bar{P})$ is considered:

$$\begin{aligned} J(x,\mu) &= \int_0^\infty \int_U r(t,x(t),\mathrm{v})\, d\mu_t(\mathrm{v}) e^{-\varrho t}\, dt \longrightarrow Min, \quad \varrho \geq 0 \\ x &\in W_2^{1,n}((0,\infty),\nu), \quad \mu \in \mathcal{M}_U,\ U \in comp(\mathbb{R}^m) \\ \dot{x}(t) &= \int_U f(t,x(t),\mathrm{v}) d\mu_t(\mathrm{v}) \text{ a.e. on } (0,\infty),\ x(t_0) = x_0. \end{aligned}$$

All integrals are to be understood in the Lebesgue sense. The relaxed controls μ are taken from a regular family of probability measures $\mathcal{M}_U$. A weighted Sobolev space $W_2^{1,n}((0,\infty),\nu)$ with a suitable weight function ν is chosen as the state space.

In comparison to the literature, see [1], [2], where overtaking or weakly overtaking optimality is mainly used as optimality criterion, the classical comparison of Lebesque integrals in the objective of $(\bar{P})$ is used here.
Under conditions that ensure the existence of the solution, cf. also the contribution by I. Dikariev, entitled

Existence Theorem for Relaxed Control Problems on Infinite Time Horizon Utilizing Weight Functions,

we treat the problem with dual methods. Here, we mainly refer to the ideas of Caratheodory and Klötzler for the construction of a dual problem. This dual based approach has already been used for special optimal control problems with infinite horizon in [3], [4].

Acknowledgements

The authors are partially supported by the German Research Foundation.

References

[1] D. A. Carlson. A Caratheodory-Hamilton-Jacobi Theory for Infinite-Horizon Optimal Control Problems *JOTA: VoL 48, No. 2, 1986.*

[2] A. Haurie. Existence and Global Asymptotic Stability of Optimal Trajectories for a Class of Infinite-Horizon, Nonconvex Systems *JOTA: Vol, 31, No, 4, 1980.*

[3] V. Lykina. Beiträge zur Theorie der Optimalsteuerungsprobleme mit unendlichem Zeithorizont. Dissertation, BTU Cottbus, German, 2010.

[4] S. Pickenhain, V. Lykina and K.Kolo. Exponential Stabilization of linear control systems with control constraints - a dual based approach *submitted to Optimization, 2023.*

Optimal control of ODEs: theory and applications

Optimal control for control affine systems with average cost

Michele Palladino[1], **Teresa Scarinci**[2]

1. michele.palladino@univaq.it University of L'Aquila, Italy
2. teresa.scarinci@unicas.it University of Cassino and Southern Lazio, Italy

1. Introduction

In this talk, we present some results pursuing the investigation in [2, 3] and consider some affine optimal control problems where a term, $f(x)$, is partially unknown. We assume that the admissible functions f vary in a space of functions $\mathcal{X}$ for which we know a probability distribution π. We prove some convergence properties for the optimal policies and the multipliers of the problem as a family of measures π_n weakly converges to π. The proofs strongly rely on Γ-convergence techniques (see the book [1] for details on Γ-convergence).

2. Some details

We consider an affine optimal control which does not require the exact knowledge of the drift term f, but merely a probability distribution defined on a space of functions containing f. The space of such distribution is denoted by $\mathcal{M}(\mathcal{X})$, while the set $\mathcal{X}$ satisfies certain conditions which will be later specified. Let $T > 0$ be given. For each $s \in [0,T]$, $x_0 \in \mathbb{R}^n$ and $\pi \in \mathcal{M}(\mathcal{X})$, consider the following optimal control problem, which we will refer to as *Problem* (B_π):

$$\begin{cases} \text{minimize } J_{s,\pi}[u] \\ \text{over } \{(x_f,u)(\cdot) : f \in \mathcal{X}\} \text{ such that } u \in \mathcal{U} \text{ and} \\ \dot{x}_f(t) = f(x_f(t)) + g(x_f(t))u(t), \quad f \in \mathcal{X}, \quad t \in [s,T], \\ x_f(s) = x_0, \quad f \in \mathcal{X}, \end{cases}$$

where $\mathcal{U} := \{u : [s,T] \to \mathbb{R}^m \text{ Lebesgue measurable}\}$ and

$$\begin{aligned} J_{s,\pi}[u] &= \mathbb{E}_\pi\left[\int_s^T \left(G\left(x_f(t)\right) + \frac{1}{2}u(t)^T R u(t)\right)dt + \Phi\left(x_f(T)\right)\right] \\ &= \int_{\mathcal{X}}\left[\int_s^T \left(G\left(x_f(t)\right) + \frac{1}{2}u(t)^T R u(t)\right)dt + \Phi\left(x_f(T)\right)\right]d\pi(f)\,. \end{aligned}$$

We address the following research question. Let us suppose that we have a family of measures $\{\pi_n\}_n$ weakly convergent to π as $n \to \infty$. If we are able to solve Problem (B_{π_n}) and compute optimal controls u^n, we wonder if a cluster point of u^n, u^∞, is a minimizer of Problem (B_π) and if there is convergence also of the optimal values of the cost functions.

$$\begin{array}{ccc} \text{Problem } (B_{\pi_n}) & & \text{Problem } (B_\pi) \\ u_n \text{ minimizers} & \overset{?}{\to} & u \text{ minimizers} \\ & \text{as } \pi_n \rightharpoonup \pi, n \to \infty. & \end{array}$$

The results presented in this talk are contained in a manuscript currently under revision.

References

[1] G. Dal Maso. An introduction to Γ-convergence. Birkhäuser Boston Inc., Boston, 1993.

[2] A. Pesare, M. Palladino, M. Falcone. Convergence of the value function in optimal control problems with unknown dinamics, 2021 European Control Conference (ECC), IEEE, 2021, pp. 2426-2431.

[3] A. Pesare, M. Palladino, M. Falcone. A convergent approximation of the Linear Quadratic Optimal Control problem for Reinforcement Learning, Mathematics of Control. Signals and Systems, vol. 33, Issue 3, September 2021, pgs. 379-411.

Optimal control of ODEs: theory and applications

Optimal control synchronization of a complex network of Lotka-Volterra systems

Cristiana J. Silva[1, 2], **Guillaume Cantin**[3]

1. cristiana.joao.silva@iscte-iul.pt *Department of Mathematics, ISTA, Iscte - Instituto Universitário de Lisboa, Av. das Forças Armadas, 1649-026 Lisboa, Portugal*
2. *Center for Research and Development in Mathematics and Applications (CIDMA), Department of Mathematics, University of Aveiro, 3810-193 Aveiro, Portugal*
3. guillaume.cantin@univ-nantes.fr *Laboratoire des Sciences du Numérique de Nantes, Nantes Université, CNRS UMR 6004, France*

In this work, we consider a controlled complex network of Lotka-Volterra systems, first proposed in [1], where the strength of the migrations of biological individuals is replaced by control functions, reproducing the implementation of ecological corridors We prove that a solution of the controlled complex network can reach a near-synchronization state, under sufficient conditions which highlight the importance to consider a positive lower on the controls functions. After, we study optimal control problems where the main goal is the minimization of the default of synchronization in the complex network. We consider different cost functionals taking into account that the dynamics of the controlled complex network ensure the conservation of both species, namely, our goal is to impose synchronization or synchronization of limit cycles. Therefore, the solutions of the optimal control problems lead to a restoration of the biodiversity of life species in a heterogeneous habitat by reaching at least a global coexistence equilibrium, or in a better scenario, a global limit cycle which would guarantee biological oscillations, which means rich life dynamics [2].

Acknowledgements

The second author was partially supported by Grant 12345678 of the Government.

References

[1] Cantin G. and Silva C. J. *Complex network near-synchronization for non-identical predator-prey systems*. AIMS Mathematics, 7 (11), 19975–19997, 2022

[2] Silva C. J. and Cantin G. *Optimal Control Synchronization of a Complex Network of Predator-Prey Systems*, accepted for publication.

Optimal control of ODEs: theory and applications

Metric regularity in optimal control

Vladimir M. Veliov
vladimir.veliov@tuwien.ac.at TU Wien, Austria

The talk is devoted to certain extensions of the properties of strong metric regularity and sub-regularity of mappings, focusing on mappings associated with optimal control problems. Namely, for control constrained optimal control problems, the first order necessary optimality conditions can be recast as an inclusion $0 \in F(y)$ (often in the form of a variational inequality), where y includes the state, co-state and control variables, and F is a mapping between subsets of Banach spaces (the so-called *optimality mapping*). Various types of regularity properties of the optimality mapping are of interest, but often the standard definitions of regularity have to be modified in order to capture in a relevant way the mappings into question.

The regularity properties we discuss in the talk are the so called *Strong Metric sub-Regularity* (SMsR) and *Strong Metric Regularity* (SMR), where, however, two metrics are used in each of the domain and image spaces. Validity of these two regularity properties of the optimality mapping have been established for various optimal control problems. In the talk we focus on problems that satisfy the Legendre-Clebsch condition and on affine optimal control problems, which will be considered separately.

Two applications will be briefly presented:
(i) existence of Lipschitz continuous optimal feedback control;
(ii) convergence with error estimates of discretization and Newton-like methods for optimal control.

The talk is partly based on results obtained in the papers given in the reference list below.

References

[1] N.P. Osmolovskii, V.M. Veliov. Metric sub-regularity in optimal control of affine problems with free end state. *ESAIM: Control, Optimisation and Calculus of Variations*, 26(47), 2020. DOI: https://doi.org/10.1051/cocv/2019046 .

[2] N.P. Osmolovskii, V.M. Veliov. On the strong subregularity of the optimality mapping in an optimal control problem with pointwise inequality control constraints. *Applied Mathematics and Optimization*, 87(3), June, 2023, https://doi.org/10.1007/s00245-022-09959-9.

[3] M. Quincampoix, T. Scarinci, V.M. Veliov. On the metric regularity of affine optimal control problems. *Journal of Convex Analysis*, 27(2), 2020.

Optimal control theory and applications

Optimal control problems with parameters and applications

Abdallah Abdel Wahab[1], **Piernicola Bettiol**[2]

1. abed901@hotmail.com *Université de Bretagne Occidentale, France*
2. Piernicola.Bettiol@univ-brest.fr *Université de Bretagne Occidentale, France*

Optimal control problems involving parameters appear a natural framework for some models arising in applications such as aerospace engineering, biology, among many others. These comprise minimax optimization problems, and bi-level problems in which an optimal control problem is coupled with a nonlinear programming problem. We shall provide necessary optimality conditions for a class of optimal problems which involve parameters, and show how these conditions can be applied to derive necessary optimality conditions for problems having a bi-level structure.

References

[1] F. H. Clarke and R. B. Vinter, *Applications of optimal multiprocesses*, SIAM J. Control Optim. 27, No. 5, 1048–1071 (1989)

[2] F. H. Clarke and R. B. Vinter, *Optimal multiprocesses*, SIAM J. Control Optim. 27, No. 5, 1072–1091 (1989)

[3] R. B. Vinter, *Minimax optimal control*, SIAM J. Control Optim. 44, No. 3, 939–968 (2005)

Optimal control theory and applications

Loss control regions in optimal control problems

Térence Bayen[1], **Anas Bouali**[2], **Loïc Bourdin**[3], **Olivier Cots**[4]

1. terence.bayen@univ-avignon.fr Avignon University, France
2. anas.bouali@univ-avignon.fr Avignon University, France
3. loic.bourdin@unilim.fr Limoges University, France
4. olivier.cots@toulouse-inp.fr INP-ENSEEIHT-IRIT, France

In this presentation, we address optimal control problems involving loss control regions. In this context, the state space is partitioned into disjoint sets referred to as regions:

$$\mathbb{R}^n = \bigcup_{k=1}^{N} \overline{X}_k,$$

where X_k are non-empty, open, disjoint open sets. These regions are classified into two types: control regions and loss control regions. We introduce an indexation $q_k \in \{0, 1\}$ allowing us to separate *control regions* and *loss control regions* as follows

$$q_k := \begin{cases} 1 \text{ if } X_k \text{ is a control region,} \\ 0 \text{ if } X_k \text{ is a loss control region,} \end{cases}$$

for all $k = 1, \dots, N$. When the state belongs to a control region, the control is permanent (i.e., the control value can be modified at any time t). In this case, the control system is given by

$$\dot{x}(t) = f(x(t), u(t)), \text{ if } q_k = 1.$$

On the other hand, when the state belongs to a loss control region, the control must remain constant, equal to the last assigned value u_k before the state enters the loss control region, and this value is kept until the state exits this region. In this case the control system is given by

$$\dot{x}(t) = f(x(t), u_k), \text{ when } q_k = 0.$$

The goal of this presentation is twofold. First, we derive a corresponding Pontryagin maximum principle based on an augmentation technique (allowing to reduce an optimal control problem with loss control regions to a (classical) optimal control problem). Second, we propose a two-step numerical scheme to solve optimal control problems with loss control regions. The approach is based on a direct numerical method applied to a regularized problem, which initializes an indirect numerical method based on the previously mentioned optimality conditions and applied to the original problem. Lastly, we apply this numerical approach to several illustrative examples.

References

[1] T. Bayen, A. Bouali, and L. Bourdin, "Optimal control problems with non-control regions: necessary optimality conditions," in *Proceedings of the IFAC workshop CAO*, 2022. https://doi.org/10.1016/j.ifacol.2022.09.002.

[2] T. Bayen, A. Bouali, L. Bourdin, and O. Cots, "Loss control regions in optimal control problems," 2023. *(submitted)*. https://hal.archives-ouvertes.fr/hal-04137550.

[3] T. Bayen, A. Bouali, L. Bourdin, and O. Cots, "On the reduction of a spatially hybrid optimal control problem into a temporally hybrid optimal control problem," *To appear in volume AIMS dedicated to Ivan Kupka*, 2023. https://hal.archives-ouvertes.fr/hal-04319045.

Optimal control theory and applications

Geometric preconditioner for indirect shooting and application to hybrid vehicle

Olivier Cots[1], **Rémy Dutto**[1,2,3], **Sophie Jan**[3], **Serge Laporte**[3]

1. *Institut de Recherche en Informatique de Toulouse, France*
2. *remy.dutto@orange.fr Vitesco Technologies, France*
3. *Institut de Mathématiques de Toulouse, France*

We are interested in the hybrid electric vehicle torque split and gear shift problem, which can be formulated as a classical Lagrange optimal control problem with fixed initial condition. The Pontryagin's maximum principle gives necessary optimality conditions adjoining to the state a covector called costate. Thus, the optimal state trajectory has to be found among the projections of the lifted trajectories, called Pontryagin extremals, given by the maximum principle. The indirect simple shooting method aims to compute Pontryagin extremals reducing the resolution to the research of the initial costate.

Classically, a Newton-like solver is used to compute zeros of the so-called shooting equations. The main drawback of this method is its sensitivity to the initial guess. Therefore, a good initial guess need to be given to make the Newton solver converge, which is not an easy task in practice.

We propose a preconditioning method [2] of the shooting function based on two main results. The first one is a geometrical interpretation of the costate that connects the final costate of the augmented system to the normal cone of the reachable set of the augmented system. The second result is well known as the Mathieu transformation, which gives the lifted canonical diffeomorphism on state-costate space from a diffeomorphism on the state space, and is related to the underlying symplectic structure.

For the considered application, we construct a preconditioner based on the affine transformation of an ellipse into the unit circle. We numerically show that the proposed preconditioning method reduces the number of iterations of our solver. Remarkably, in our experiments, it is better to use the preconditioner than to provide a good initial guess [1] for the shooting function.

References

[1] Olivier Cots, Rémy Dutto, Sophie Jan, and Serge Laporte. A bilevel optimal control method and application to the hybrid electric vehicle, 2023. Submitted to Optim. Control Appl. Methods.

[2] Olivier Cots, Rémy Dutto, Sophie Jan, and Serge Laporte. Geometric preconditioner for indirect shooting and application to hybrid vehicle, 2024. Proceeding submitted to the IFAC MICNON 2024 conference.

Optimal control theory and applications

Low-thrust satellite collision avoidance as a minimum-time problem for optimal control

Lamberto Dell'Elce[1], **Jean-Baptiste Pomet**[2], **Frank de Veld**[3]

1. lamberto.dell-elce@inria.fr Université Côte d'Azur, Nice, France, LJAD, Nice, France
2. jean-baptiste.pomet@inria.fr Université Côte d'Azur, Nice, France, LJAD, Nice, France
3. frank.develd@inria.fr Université Côte d'Azur, Nice, France

1. Introduction

Near-Earth space is becoming increasingly congested with man-made space debris, resulting in the need for satellites to perform a collision avoidance manoeuvre (CAM) whenever a space debris object is predicted to come dangerously close to their orbit. Through this CAM, the satellite reaches a safe separation distance with the secondary object, assumed uncooperative. While high-thrust ($\approx 1-100N$) CAM's on large, conventional satellites are relatively straightforward to approach, low-thrust ($\approx 0.1-10mN$) CAM's are more challenging to optimise due to the long time scales involved. Typically, they are optimised for fuel consumption of orbital energy loss [2]. This work instead focuses on optimising low-thrust CAM's for minimum time taken to perform the CAM, an objective which is pertinent for late conjunction alerts and when satellite operation time is valuable. As the encounter is assumed to be fast, an analysis is required only at one time instance; moment of closest approach t_{CA}.

The research is comparable to the reachable set analysis of Evans et al., with the addition of a constraint characterising a closest approach [3]. Like their work, the CAM optimisation is investigated as a two-point boundary value problem (TPBVP) and optimal control is applied in first-order approximation. Their assumption of optimal steering direction is proven in this work in the same first-order approximation.

2. Methodology

Given the satellite state at closest approach y, a relative position Δr and a relative velocity Δv, there are two constraints characterising the problem which are valid at the final time t_f:

$$g(y_f, t_f) = \langle \Delta r(y_f, t_f), \Delta v(y_f, t_f) \rangle = 0, \tag{2.1}$$

fixing the characterising the conjunction as a closest approach and

$$s(y_f, t_f) = |\Delta r(y_f, t_f)|^2 - r_{\text{safe}}^2 = 0, \tag{2.2}$$

ensuring a safe outcome of the CAM. As the thrust duration t_0 is to be minimised, the initial conditions of the TPBVP are not fully characterised. Developing purely Keplerian dynamics with t_{CA} as a state variable, linearising the TPBVP and employing the necessary conditions of Pontryagin's Maximum Principle [1] lead to the result that the co-state λ is constant for this problem. This finding allows for a formulation of the linearised TPBVP involving only a single integral, hereby coupling the minimised t_0 with the small thrust magnitude and safe separation distance. Through backward integration starting from a safe outcome of the encounter, minimum-time trajectories can be found for any initial condition up until a perfect collision with zero initial separation distance. In first-order approximation, it is shown that these trajectories are confined to a two-dimensional plane called the conjunction plane or B-plane.

Special attention in results analysis is given for cut points for varying t_0 in minimum-time contours found through backward integration, as these points represent different minimum-time trajectories originating from the same initial condition. The control found through solving this linearised TPBVP can be used further serving as initial guesses in a true optimal control problem.

References

[1] Andrei A. Agrachev and Yuri L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II. doi:10.1007/978-3-662-06404-7.

[2] Andrea De Vittori, Maria Francesca Palermo, Pierluigi Di Lizia, and Roberto Armellin. Low-thrust collision avoidance maneuver optimization. *Journal of Guidance, Control, and Dynamics*, 45(10):1815–1829, 2022.

[3] Adam Evans, Claudio Bombardelli, Roberto Armellin, and Laura Pirovano. Analytical state transition matrix approach to low-thrust reachability computation. *Number AAS*, pages 23–206, 2023.

Optimization and stabilization of infinite-dimensional dynamical systems

On the stabilization of a kinetic model by feedback-like control fields in a Monte Carlo framework

Jan Bartsch[1], **Alfio Borzì**[2]

1. jan.bartsch@uni-konstanz.de University of Konstanz, Germany
2. alfio.borzi@uni-wuerzburg.de Julius-Maximilians Universität Würzburg, Germany

The construction of feedback-like control fields for a kinetic model in phase space is investigated. The purpose of these controls is to drive an initial density of particles in the phase space to reach a desired cyclic trajectory and follow it in a stable way. For this purpose, an ensemble optimal control problem governed by the kinetic model is formulated in a way that is amenable to a Monte Carlo approach. The proposed formulation allows to define a one-shot solution procedure consisting in a backward solve of an augmented adjoint kinetic model. Results of numerical experiments demonstrate the effectiveness of the proposed control strategy.

1. Setting

Our kinetic model consists of a Liouville-type non-homogeneous streaming operator and a linear collision term $C[f]$ as follows:

$$\partial_t f(x,v,t) + \nabla_x \cdot \big(v\, f(x,v,t)\big) + \nabla_v \cdot \big(F(x,v,t;u)\, f(x,v,t)\big) = C[f](x,v,t), \tag{1.1}$$

with a control function $u = u(x,v,t)$ within the force term F.

We consider an initial- and boundary-value problem with this model in the phase space $\Omega \times \mathbb{R}^d$, where $x \in \Omega \subset \mathbb{R}^d$ represents the position space coordinate and $v \in \mathbb{R}^d$ represents the velocity. On the inflow part of the boundary $\partial\Omega$, we require (partial) specular reflection boundary conditions. Our objective function is given by

$$J(u,f) = \int_0^T \int_{\Omega\times\mathbb{R}^d} \Big(\theta(x,v,t) + \frac{\nu}{2}|u(x,v,t)|^2\Big) f(x,v,t)\, dx\, dv\, dt + \int_{\Omega\times\mathbb{R}^d} \varphi(x,v) f(x,v,T)\, dx\, dv, \tag{1.2}$$

where θ encodes the task of the control to drive the particles along a desired trajectory in phase space; similarly, in φ the desired configuration at final time is encoded.

2. Optimization

Introducing an adjoint variable $q = q(x,v,t)$ and exploiting the Lagrange framework, we derive the following first-order optimality condition

$$f(x,v,t)\Big(\nu\, u(x,v,t) - \partial_u F(x,v,t;u)\nabla_v q(x,v,t)\Big) = 0. \tag{2.1}$$

Assuming that the density f is positive everywhere, a necessary and sufficient condition for (2.1) is to set

$$u(x,v,t) = \frac{1}{\nu}\partial_u F(x,v,t;u)\nabla_v q(x,v,t). \tag{2.2}$$

This is an essential step in our development because the u given by (2.2) does not depend on f but solely on the optimization functions θ and φ that define the control tasks. These are the characterizing features of a feedback control. It is now sufficient to solve an augmented adjoint model once in order to derive a control, that performs given tasks for any initial condition. In Fig. 1, we present results showing the ability of our control field to drive an initial density of particles distributed in the phase-space to reach and maintain a

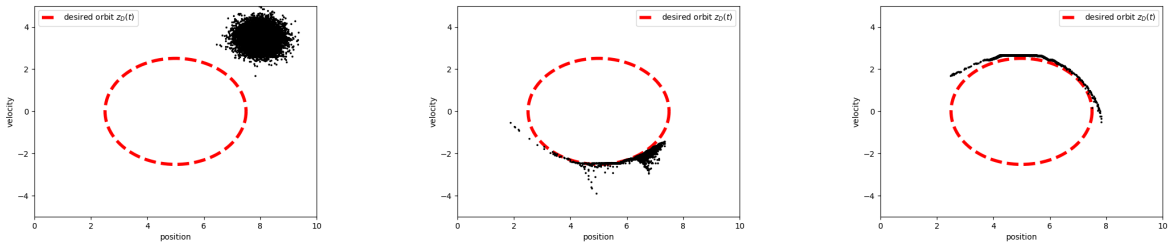


Fig. 1 Time-Snapshots of the evolution of particles in phase space (ord.: velocity, abs.: position, dashed red: z_D)

Acknowledgements

The first author was partially supported by the Deutsche Forschungsgemeinschaft within the project SFB1432.

Optimization and stabilization of infinite-dimensional dynamical systems

Sensitivity analysis and optimal control for the Chemical Master Equation

Frédéric Bonnans[1], **Arthur Bottois**[2], **Laurent Pfeiffer**[3]

1. frederic.bonnans@inria.fr L2S CentraleSupélec & INRIA Saclay-Île-de-France, France
2. arthur.a.bottois@inria.fr L2S CentraleSupélec & INRIA Saclay-Île-de-France, France
3. laurent.pfeiffer@inria.fr L2S CentraleSupélec & INRIA Saclay-Île-de-France, France

We consider a chemical system composed of d species $\mathsf{X}_1, \dots, \mathsf{X}_d$ interacting via R reactions,

$$\text{for } 1 \leq r \leq R, \quad \mathsf{R}_r : \ \alpha_r \cdot \mathsf{X} \xrightarrow{a_r(t,x)} \beta_r \cdot \mathsf{X}, \quad \text{with } \mathsf{X} = (\mathsf{X}_1, \dots, \mathsf{X}_d).$$

Each reaction R_r is characterized by its stoichiometric coefficients $\alpha_r, \beta_r \in \mathbb{N}^d$ and its reaction rate $a_r(t,x)$. This reaction rate may depend on time $t \geq 0$ and on the state $x \in \mathbb{N}^d$ of the system. Being in state $x = (x_1, \dots, x_d)$ means that there are x_i molecules of species X_i, for all $1 \leq i \leq d$. Such a chemical system may be described by a continuous-time Markov chain $X(t)$, with law $p(t,x) = \mathbb{P}[X(t) = x]$, the probability to be in state x at time t. This law satisfies a Kolmogorov equation, called the Chemical Master Equation (CME) in this context,

$$\frac{\partial p}{\partial t}(t,x) = \sum_{r=1}^{R} \Big(a_r(t, x - \nu_r) p(t, x - \nu_r) - a_r(t,x) p(t,x) \Big), \quad \text{with } \nu_r = \beta_r - \alpha_r. \tag{CME}$$

This equation is used in biology (cf. [1]) to model chemical reactions taking place inside cells, in order to take into account the randomness of these reactions, in a context where the number of molecules involved is too low to make a continuous approximation of the state space $\mathcal{S} = \mathbb{N}^d$. Thus, the CME is a countable collection of ordinary differential equations.

In the simple case of bounded reaction rates a_r, the Cauchy-Lipschitz theorem ensures the existence and uniqueness of the solution, while the implicit function theorem allows to conduct a sensitivity analysis of this solution with respect to the data (initial condition, reaction rates, ...). However, practical applications in biology lead to consider reaction rates that are proportional to the number of reactants and therefore unbounded. In this framework, the uniqueness of the solution is no longer guaranteed. One may nevertheless use the appropriate (in a physical sense) notion of minimal solution (cf. [2]).

Using stability estimates for the minimal solution, we establish la differentiability of this solution with respect to the data. We then give an expression for the directional derivative involving an adjoint equation, for which we also define a concept of minimal solution. This allows us to derive optimality conditions (in the form of a Pontryagin's principle) for a class of optimization problems for which we control the reaction rates.

References

[1] S. Winkelmann & C. Schütte. *Stochastic dynamics in computational biology.* Cham: Springer, 2020.

[2] L. Gauckler & H. Yserentant. *Regularity and approximability of the solutions to the chemical master equation.* ESAIM: M2AN, 48(6), 1757–1775, 2014.

Optimization and stabilization of infinite-dimensional dynamical systems

Optimal control of third grade fluids with multiplicative noise

Fernanda Cipriano[1], **Yassine Tahraoui**[2]

1. cipriano@fct.unl.pt *New University of Lisbon, Portugal*
2. tahraouiyacine@yahoo.fr *Scuola Normale Superiore di Pisa, Italy*

This work aims to control the dynamics of certain non-Newtonian fluids in a bounded domain of $\mathbb{R}^d$, $d = 2, 3$ perturbed by a multiplicative Wiener noise, the control acts as a predictable distributed random force, and the goal is to achieve a predefined velocity profile under a minimal cost. Due to the strong nonlinearity of the stochastic state equations, strong solutions are available just locally in time (cf. [1]), and the cost functional includes an appropriate stopping time. First, we show the existence of an optimal pair. Then, we show that the solution of the stochastic forward linearized equation coincides with the Gâteaux derivative of the control-to-state mapping, after establishing some stability results. Next, we analyse the backward stochastic adjoint equation; where the uniqueness of solution holds only when $d = 2$. Finally, we establish a duality relation and deduce the necessary optimality conditions. For further information, we address the reader to [2].

Acknowledgements

This work is funded by national funds through the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 and UIDP/00297/2020 (Center for Mathematics and Applications).

References

[1] Y. Tahraoui, F. Cipriano. Local strong solutions to the stochastic third grade fluid equations with Navier boundary conditions. Stoch PDE: Anal Comp (2023). https://doi.org/10.1007/s40072-023-00314-9.

[2] Y. Tahraoui, F. Cipriano. Optimal control of third grade fluids with multiplicative noise https://doi.org/10.48550/arXiv.2306.13231.

Optimization and stabilization of infinite-dimensional dynamical systems

Designing the monodomain model with artificial neural networks

Sébastien Court[1], **Karl Kunisch**[2]

1. sebastien.court@uibk.ac.at University of Innsbruck, Austria
2. karl.kunisch@uni-graz.at University of Graz & RICAM Linz, Austria

1. Introduction

We propose a data-driven method in order to identify the nonlinearity in the monodomain model. The monodomain model is a system coupling a semilinear parabolic PDE with an ODE, describing the time evolution of an electric potential. Our approach provides a general answer to the problem of selecting the model when studying phenomena related to cardiac electrophysiology: From measurements, instead of determining coefficients of a prescribed nonlinearity (like the FitzHugh-Nagumo model for instance), we design the nonlinearity itself, in the form of an artificial neural network (ANN), more specifically a feedforward residual network. The relevance of this approach relies on the approximation capacities of neural networks. Training the ANN corresponds to solving an identification problem constrained by the monodomain model so parameterized. We formulate this inverse problem as an optimal control problem, and provide mathematical analysis and derivation of optimality conditions for identifying the weights of the ANN. One of the difficulties comes from the lack of smoothness of activation functions which are classically used for training deep neural networks. We will also present numerical results that demonstrate the feasibility of the strategy proposed in this work. This is joint work with Prof. Karl Kunisch (RICAM & University of Graz), published in [1].

2. State equations

Consider a bounded domain of $\mathbb{R}^d$ ($d = 2$ or 3). The monodomain model writes as follows

$$\begin{aligned} \frac{\partial v}{\partial t} - \nu\Delta v + \phi_v(v,w) &= f_v && \text{in } \Omega\times(0,T),\\ \frac{\partial w}{\partial t} + \delta w + \phi_w(v,w) &= f_w && \text{in } \Omega\times(0,T),\\ \frac{\partial v}{\partial n} &= 0 && \text{on } \partial\Omega\times(0,T),\\ (v,w)(\cdot,0) &= (v_0,w_0) && \text{in } \Omega, \end{aligned} \tag{$*$}$$

where $\mathrm{z} = (v,w)$ is the state variable, (f_v, f_w) are given right-hand-sides, (v_0, w_0) are given initial conditions, and the nonlinearities (ϕ_v, ϕ_w) are designed from data in the form of a feedforward neural network with L layers parameterized with affine functions (weights) $W_1, \dots, W_L$:

$$\Phi(\mathrm{z}) = \begin{pmatrix} \phi_v(\mathrm{z}) \\ \phi_w(\mathrm{z}) \end{pmatrix} = \begin{cases} W_2(\rho(W_1(\mathrm{z}))) & \text{if } L = 2,\\ W_L(\rho(W_{L-1}(\rho(\dots\rho(W_1(\mathrm{z})))))) & \text{if } L \geq 3. \end{cases}$$

3. Illustrations

After training the ANN represented by mapping Φ with measurements based on the Aliev-Panvilov model, we illustrate the time evolution of the corresponding state variables in Figure 1.

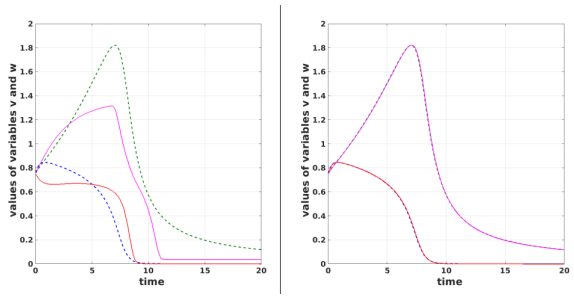


Fig. 1 Values of the state variables $(v_{\mathrm{AP}}, w_{\mathrm{AP}})$ of the Aliev-Panvilov model in dashed blue and green respectively. In red and magenta respectively the values of the solution (v,w) of the monodomain model $(*)$ with the trained ANN as nonlinearity. Left: with Architecture 1. Right: with Architecture 2.

References

[1] Sébastien Court and Karl Kunisch. Design of the monodomain model by artificial neural networks. *Discrete Contin. Dyn. Syst.*, 42(12):6031–6061, 2022. URL: https://doi.org/10.3934/dcds.2022137, doi:10.3934/dcds.2022137.

Optimization and stabilization of infinite-dimensional dynamical systems

A robust tracking feedback control of linear dynamical systems with uncertain parameters

Philipp A. Guth[1], **Karl Kunisch**[2], **Sérgio S. Rodrigues**[3]

1. philipp.guth@ricam.oeaw.ac.at RICAM, Austrian Academy of Sciences, Austria
2. karl.kunisch@uni-graz.at Karl-Franzens University of Graz, Austria, and RICAM, Austrian Academy of Sciences, Austria
3. sergio.rodrigues@ricam.oeaw.ac.at RICAM, Austrian Academy of Sciences, Austria

1. Introduction

Stabilization of linear control systems with uncertain parameters is investigated. A Riccati based feedback mechanism is analyzed, constructed using an ensemble of candidate parameters from an *a-priori* chosen training set. This feedback stabilizes all systems of the training set and also systems in its vicinity. Moreover, for optimal control problems of tracking type the described strategy results in an affine tracking feedback control input. The suboptimality with respect to optimal feedback for each single parameter from the training set can be quantified. Numerical results are presented to demonstrate the robust performance of the proposed feedback mechanism for trained as well as untrained parameters.

2. Some details on the construction of the feedback

Optimal control problems subject to linear dynamical systems with uncertain parameters are considered, where the uncertain parameter σ enters the system via $\mathcal{A}_\sigma$ as

$$\dot{x}_\sigma(t) = \mathcal{A}_\sigma x_\sigma(t) + Bu(t), \tag{2.1}$$

for $t > 0$ and given $x_\sigma(0) = x_\circ$. Based on an *a-priori* chosen training ensemble $\Sigma \coloneqq \{\sigma_1, \dots, \sigma_N\}$ of potential realizations of the unknown parameter σ, an auxiliary system is construct as

$$\dot{\mathbf{x}}_\Sigma(t) = \mathbf{A}_\Sigma \mathbf{x}_\Sigma(t) + \mathbf{B}u(t), \tag{2.2}$$

for $t > 0$ and $\mathbf{x}_\Sigma(0) = \mathcal{E}x_\circ$, where $\mathcal{E} : x \mapsto (x^\top, \dots, x^\top)^\top$, $\mathbf{B} \coloneqq \mathcal{E}B$, and where $\mathbf{A}_\Sigma \coloneqq \operatorname{diag}(\mathcal{A}_{\sigma_1}, \dots, \mathcal{A}_{\sigma_N})$ is block diagonal. Given a quadratic cost functional, the optimal feedback law $\mathbf{K}_\Sigma$ for (2.2) is obtained by solving an associated Riccati equation. The corresponding optimal feedback control coincides with the optimal open-loop control for a robust formulation of the optimal control problem based on the sample average over the training ensemble. This motivates the use of the feedback control

$$u(t) = -B^\top \mathcal{E}^\top \mathbf{K}_\Sigma \mathcal{E} x_\sigma(t)$$

in (2.1). This presentation is based on [1–3].

Acknowledgements

The third author was partially supported by the State of Upper Austria and Austrian Science Fund (FWF): P 33432-NBL.

References

[1] P. A. Guth, K. Kunisch, and S. S. Rodrigues. Ensemble feedback stabilization of linear systems. *arXiv* preprint, 2023. doi:10.48550/arXiv.2306.01079.

[2] P. A. Guth, K. Kunisch, and S. S. Rodrigues. Stabilization of uncertain linear dynamics: an offline-online strategy. *arXiv* preprint, 2023. doi:10.48550/arXiv.2307.14090.

[3] P. A. Guth, K. Kunisch, and S. S. Rodrigues. Tracking optimal feedback control under uncertain parameters. *arXiv* preprint, 2024. doi:10.48550/arXiv.2402.00526.

Optimization and stabilization of infinite-dimensional dynamical systems

Deflation and stabilization for McKean-Vlasov PDEs

Sara Bicego[1], **Dante Kalise**[2], **Grigorios A. Pavliotis**[3]

1. s.bicego21@imperial.ac.uk *Department of Mathematics, Imperial College London, United Kingdom*
2. dkaliseb,g.pavliotis@imperial.ac.uk *Department of Mathematics, Imperial College London, United Kingdom*
3. g.pavliotis@imperial.ac.uk *Department of Mathematics, Imperial College London, United Kingdom*

Interacting particle systems have emerged as a powerful tool for modeling collective behavior across various disciplines, from physics and biology to social dynamics and machine learning. As the number of particles grows large, the system can be described by a probability density distribution that evolves according to a McKean-Vlasov partial differential equation (PDE). This mean-field limit often exhibits complex phenomena such as phase transitions and multistability, which are linked to the underlying microscopic interactions, the noise amplitude, and the interaction strength.

In this talk, we present a numerical framework for studying the steady states of the McKean-Vlasov PDE and designing control strategies to stabilize the system around desired configurations. We first discuss a spectral Galerkin approximation of the PDE, which leads to a finite-dimensional root-finding problem. To systematically identify distinct steady states, we employ a deflated Newton's method [1] that iteratively eliminates known solutions from consideration. The stability of the obtained stationary solutions is then analyzed using free energy arguments and self-consistency conditions, enabling a characterization of the system's phase transitions [2].

Building upon the knowledge of the steady states, we formulate an optimal control problem to stabilize the particle ensemble around unstable configurations of interest. The control enters as an additional drift term in the McKean-Vlasov dynamics, and a feedback law is computed by solving a PDE-constrained optimization problem [3]. We derive optimality conditions using the Pontryagin's Maximum Principle and propose a gradient-based algorithm for numerical solution. To achieve stabilization in the presence of instabilities, we employ a model predictive control strategy that solves the finite-horizon problem in a receding horizon fashion.

Throughout the talk, we highlight the interplay between the particle dynamics and their mean-field limit, illustrating how the micro-scale interactions shape the emergent macro-scale behavior. We present numerical examples that demonstrate the effectiveness of the proposed deflation technique in capturing phase transitions and the ability of the optimal control framework to steer the system towards desired states. The computational tools developed here offer valuable insights into the multiscale nature of interacting particle systems and open up new avenues for their analysis and control.

Acknowledgements

This research was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/T024429/1

References

[1] P. E. Farrell, Á. Birkisson, and S. W. Funke. Deflation techniques for finding distinct solutions of nonlinear partial differential equations. *SIAM Journal on Scientific Computing*, 37(4):A2026–A2045, 2015. doi:10.1137/140984798.

[2] G.A. Pavliotis M.G. Delgadino, R.S. Gvalani. On the diffusive-mean field limit for weakly interacting diffusions exhibiting phase transitions. *Archive for Rational Mechanics and Analysis*, pages 91––148, 2021. doi:10.1007/s00205-021-01648-1.

[3] Karl Kunisch Tobias Breiten and Laurent Pfeiffer. Control strategies for the Fokker–Planck equation. *ESAIM: COCV*, 24(2):741 – 763, 2018. doi:10.1051/cocv/2017046.

Optimization and stabilization of infinite-dimensional dynamical systems

Including gradients in optimal control of quasilinear parabolic equations

Fabian Hoppe[1], **Hannes Meinlschmidt**[2], **Ira Neitzel**[3]

1. fabian.hoppe@dlr.de Deutsches Luft- und Raumfahrtzentrum (DLR), Germany
2. meinlschmidt@math.fau.de FAU Erlangen-Nürnberg, Germany
3. neitzel@ins.uni-bonn.de Rhenish Friedrich Wilhelm University of Bonn, Germany

1. Setting

In this contribution [1] we consider optimal control problems subject to control– and state constraints $u \in U_{\mathrm{ad}}$ and $y \in Y_{\mathrm{ad}}$, with the governing equation being of abstract quasilinear parabolic type

$$\partial_t y + \mathcal{A}(y)y = Bu + \mathcal{F}(y) \quad \text{in } X, \qquad y(0) = y_0, \tag{QL}$$

on a finite time interval $(0,T)$, with some Banach space X. The underlying actual PDE formulation is posed on a bounded domain $\Omega \subset \mathbb{R}^d$ with minimal smoothness assumptions, essentially a Lipschitz manifold. It includes mixed boundary conditions on boundary parts $D \dot\cup N = \partial\Omega$ where it is allowed that $D \cap \overline{N} \neq \emptyset$, so that there is an actual interface between boundary conditions, for which we have further minimal assumptions.

The quasilinear operator $\mathcal{A}(y)y$ is of the form $-\operatorname{div}(\xi(y)\mu\nabla y)$, so with a "zero-order" nonlinearity induced by a real Lipschitz function ξ and a merely bounded and measurable coefficient matrix function μ; in contrast, we allow that the nonlinear function $\mathcal{F}$ also depends on first derivatives of y, with the prototype $\mathcal{F}(y) = |\nabla y|^2$. Such nonlinearities are interesting in several practical applications, but they are also notoriously hard to deal with in the analysis of (QL), in particular in a nonsmooth setup as we consider where one cannot expect, say, $H^2(\Omega)$ regularity for the solution. The control operator B incorporates the control into the equation in a linear and continuous fashion.

We further consider also the particular case where Y_{ad} is not merely a classical state constraint but in fact a constraint on the *gradient* of the state.

2. Results

In principle, the general way to do optimal control theory for the problems considered is well known whenever the equation admits the usual wellposedness and regularity results for its solutions. However, the regularity requirements are nontrivial to obtain in the nonsmooth setting. In particular, the established arguments for the weak setting $X = W_D^{-1,q}(\Omega)$—under the assumption that there is a $q > d$ such that there is optimal elliptic $W_D^{1,q}(\Omega)$ regularity for the operator $-\nabla\cdot\mu\nabla$ on X—are not sufficient for the present case with gradient nonlinearities and/or gradient constraints, whereas the strong setting $X = L^q(\Omega)$ is generally not tractable in the given setting.

We thus derive a most flexible theory for (QL) in the scale $X = X_\theta = [W_D^{-1,q}(\Omega), L^q(\Omega)]_\theta$ *between* the weak and strong setting, based nonautonomous maximal parabolic regularity, using on a bilinear interpolation technique to determine the needed multiplier regularity [2]. The upper limit for θ depends on the degree of Hölder regularity κ for the elliptic $L^q(\Omega)$ problem for $-\nabla\cdot\mu\nabla$. Based on the $W_D^{-1,q}(\Omega)$-setting, this allows to deal with the prototype nonlinearity $\mathcal{F}(y) = |\nabla y|^2$ and a spatially integrated gradient constraint $y \in Y_{\mathrm{ad}}$ without further assumptions, generalizing earlier results [3]. In the optimal case, $\kappa > 1$—which is to be expected under stronger assumptions only—, and we get that $\nabla y \in C(\overline{Q})$ which in fact allows to consider even pointwise bounds on the gradients in the state constraints.

The theory developed thus gives a comprehensive and flexible framework to deal with the quasilinear equation (QL) and associated optimal control problems, which falls back to the expected "classical" results whenever the data admits it.

Acknowledgements

F.H. was partially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–Projektnummer 211504053–SFB 1060. I.N. gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–Projektnummer 211504053–SFB 1060.

References

[1] Fabian Hoppe, Hannes Meinlschmidt and Ira Neitzel. Optimal control of quasilinear parabolic PDE with gradient terms and constraints on the gradient of the state. In preparation.

[2] Fabian Hoppe, Hannes Meinlschmidt and Ira Neitzel. Global-in-time solutions and Hölder continuity for quasilinear parabolic PDEs with mixed boundary conditions in the Bessel dual scale. arXiv:2303.04659, 2023.

[3] Lucas Bonifacius and Ira Neitzel. Second order optimality conditions for optimal control of quasilinear parabolic equations. *Math. Control Relat. Fields*, 8(1):1–34, 2018.

Optimization and stabilization of infinite-dimensional dynamical systems

A mesh-independent method for second-order potential mean field games

Kang Liu[1], **Laurent Pfeiffer**[2]

1. kang.l.liu@fau.de *Friedrich-Alexander-Universität, Erlangen-Nürnberg, Germany*
2. laurent.pfeiffer@inria.fr *Inria and CentraleSupélec, Paris-Saclay, University, France*

We consider a class of potential mean-field games (MFGs) with a potential structure: they are equivalent to certain convex optimal control problems of the Fokker-Planck equation. The Fokker-Planck equation modeling the evolution of the distribution of a stochastic differential equation, the problems that we will consider also encompass a large class of nonlinear stochastic optimal control problems.

It was recently shown that the Generalized Conditional Gradient (GCG) algorithm (also called Generalized Frank-Wolfe algorithm) is a method of choice for the resolution of convex and potential second-order MFGs as it exhibits a linear rate of convergence, when applied to the continuous model [2].

We will discuss in the talk the application of the GCG method to a discretized MFG model and show that the initial convergence properties of the algorithm are preserved at the discrete level, a property referred to as mesh-independence [3]. The discrete model under consideration is obtained with the theta-scheme, a finite-difference method which we introduced in [1]. Numerical results will be presented.

References

[1] J.F. Bonnans, K. Liu, and L. Pfeiffer. Error estimates of a theta-scheme for second-order mean field games. *ESAIM: Mathematical Modelling and Numerical Analysis*, 57(4):2493–2528, 2023.

[2] P. Lavigne and L. Pfeiffer. Generalized conditional gradient and learning in potential mean field games. *Applied Mathematics & Optimization*, 88(3):89, 2023.

[3] K. Liu and L. Pfeiffer. A mesh-independent method for second-order potential mean field games. *arXiv preprint arXiv:2310.20041*, 2023.

Optimization and stabilization of infinite-dimensional dynamical systems

Maximal L^p-regularity for an abstract evolution equation with applications to closed-loop boundary feedback control problems

Irena Laseicka[1], **Buddhika Priyasad**[2], **Roberto Triggiani**[3]

1. lasiecka@memphis.edu University of Memphis, USA
2. priyasad@uni-konstanz.de University of Konstanz, Germany
3. rtrggani@memphis.edu University of Memphis, USA

1. Abstract

In this talk, we present an abstract maximal L^p-regularity result for linear partial differential equations of parabolic type defined on a bounded domain and subject to finite-dimensional stabilizing feedback controls applied to a portion of the boundary, extending the analysis to $T = \infty$. Our aim is to provide a comprehensive framework that accommodates various scenarios, including both classical and more contemporary examples.

Firstly, we examine the application of our framework to the classical boundary parabolic setting, illustrating how maximal L^p-regularity can be effectively utilized in this context.

Secondly, we delve into more recent developments by considering specific systems. One such system is the 3D Navier-Stokes Equations, where we explore the integration of finite-dimensional, localized, boundary tangential feedback stabilizing controls. This investigation is crucial for understanding how boundary controls can influence the stability and behavior of fluid flows governed by the Navier-Stokes equations.

Additionally, we investigate Boussinesq Systems, focusing on the integration of finite-dimensional, localized feedback for stabilizing Dirichlet boundary control within the thermal equation component of these systems. This exploration sheds light on the interplay between boundary control strategies and the dynamics of thermally driven flows.

Our study not only contributes theoretical insights into the abstract theory of maximal L^p-regularity for parabolic PDEs but also offers practical implications for the design and implementation of control strategies in fluid dynamics and related fields. By extending our analysis to infinite time horizon ($T = \infty$), we emphasize the long-term stability and effectiveness of the proposed control frameworks.

Optimization and stabilization of infinite-dimensional dynamical systems

MOR based RHC for stabilization of time-varying linear parabolic PDEs

Behzad Azmi[1], **Jan Rohleff**[2], **Stefan Volkwein**[3]

1. behzad.azmi@uni-konstanz.de *University of Konstanz, Germany*
2. jan.rohleff@uni-konstanz.de *University of Konstanz, Germany*
3. stefa.volkwein@uni-konstanz.de *University of Konstanz, Germany*

This talk deals with the stabilization of a class of time-varying parabolic partial differential equations using Receding Horizon Control (RHC) based on Model Order Reduction (MOR). Here the control is finite-dimensional, i.e. it appears as a time-dependent linear combination of finitely many indicator functions whose total support covers only a small part of the spatial domain. We also include the squared ℓ_1 norm the control cost. This leads to an infinite-horizon nonsmooth problem, which allows for stabilizing optimal control with a low number of active actuators over time. First, the stabilizability of RHC is discussed. Then, to speed up the numerical computation, the data-driven MOR approaches are appropriately integrated into the RHC framework. Numerical experiments illustrating the advantages of our MOR approaches are also reported.

References

[1] Behzad Azmi, Jan Rohleff and Stefan Volkwein. *Finite-Dimensional RHC Control of Linear Time-Varying Parabolic PDEs: Stability Analysis and Model-Order Reduction.* URL: https://arxiv.org/abs/2401.09111, 2024.

Optimization and stabilization of infinite-dimensional dynamical systems

On the well-posedness of an energy-optimal state observer for a defocusing cubic wave equation

Jesper Schröder

j.schroeder@tu-berlin.de Technische Universität Berlin, Germany

In this presentation the analytical background of nonlinear observers based on *minimal energy estimation* is discussed. Originally, such strategies were proposed for the reconstruction of the state of finite dimensional nonlinear dynamical systems by means of a measured output where both the dynamics and the output are subject to white noise. For a realization of the output given by real life measurements the energy of disturbances is minimized over a deterministic interpretation of the system disturbance. This perspective allows the application of well developed techniques from optimal control theory.

The resulting observer is also referred to as the *Mortensen observer* and offers an alternative to the widely used *extended Kalman filter*. While in practice the latter yields good results at reasonable computational effort for many systems its derivation is purely heuristic and no optimality properties can be expected.

Our work aims at lifting this concept to a class of partial differential equations using the example of a wave equation with a cubic defocusing term in three space dimensions. In this setting infinite dimensional state spaces introduce new challenges. Even the well-posedness of the *Kalman filter* for linear systems is not immediately clear and showing well-posedness of the extended Kalman filter for nonlinear systems is an open field of research for many PDEs. In a pursuit of the rigorous derivation of the Mortensen observer for the aforementioned wave equation we discuss local regularity of the corresponding value function which is defined on an infinite dimensional Hilbert space. Furthermore, operator Riccati equations are considered in order to characterize the second order spatial derivative of said value function.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 504768428.

OPTIMIZATION AND STABILIZATION OF INFINITE-DIMENSIONAL DYNAMICAL SYSTEMS

Policy gradient algorithms for robust MDPs with non-rectangular uncertainty sets

Mengmeng Li[1], **Daniel Kuhn**[1], **Tobias Sutter**[2]

1. mengmeng.li@epfl.ch, daniel.kuhn@epfl.ch *École Polytechnique Fédérale de Lausanne, Switzerland*
2. tobias.sutter@uni-konstanz.de *University of Konstanz, Germany*

This is an extended abstract summarizing the paper [1].

1. Introduction

Markov decision processes (MDPs) form the backbone of reinforcement learning and dynamic decision-making. Classical MDPs operate in a time-invariant stochastic environment represented by a known constant transition kernel. In most applications, however, the transition kernel is only indirectly observable through a state-action trajectory generated under a fixed policy. In these setting it is expedient to work with *robust* MDPs [3], which assume that the unknown true transition kernel falls within a known uncertainty set and aim to identify a policy that exhibits the best performance under the worst-case transition kernel in this uncertainty set.

The literature on robust MDPs distinguishes rectangular and non-rectangular uncertainty sets. An uncertainty set is called (s)-rectangular (or (s,a)-rectangular) if it is representable as a Cartesian product of separate uncertainty sets for the transition probabilities associated with the different current states s (or current state-action pairs (s,a)). Otherwise, the uncertainty set is called non-rectangular. Rectangularity is intimately related to computational tractability. Indeed, robust MDPs with rectangular polyhedral uncertainty sets can be solved in polynomial time, whereas robust MDPs with non-rectangular polyhedral uncertainty sets are NP-hard [3]. Most existing papers on robust MDPs focus on rectangular uncertainty sets. However, statistically optimal uncertainty sets often fail to be rectangular. Robust MDPs with rectangular uncertainty sets are usually addressed with value iteration, policy iteration, convex reformulation, or policy gradient methods. These solution methods, however, become inefficient or converge to strictly suboptimal solutions of the robust MDP if the uncertainty set fails to be rectangular.

2. Contributions

We propose policy gradient algorithms for robust infinite-horizon Markov decision processes (MDPs) with non-rectangular uncertainty sets, thereby addressing an open challenge in the robust MDP literature. The main contributions of our paper can be summarized as follows.

1. We show that robust policy evaluation problems with non-rectangular uncertainty sets can be solved to global optimality with a projected Langevin dynamics algorithm. Numerical results suggest that if the uncertainty set happens to be rectangular, then this randomized algorithm is competitive with state-of-the-art deterministic first-order methods in terms of runtime.
2. We present a conservative policy iteration algorithm that solves robust policy evaluation problems approximately. The approximation error is shown to scale with a new measure of non-rectangularity of the uncertainty set. We prove that the same method solves robust policy evaluation problems with rectangular uncertainty sets to any accuracy $\epsilon > 0$ in $O(S^2/\epsilon^2)$ iterations, where S denotes the number of states. In contrast, the iteration complexity of the state-of-the-art policy gradient method for this problem class developed in [2] includes an extra factor S^3A, where A denotes the number of actions.
3. We present an actor-critic method that solves robust policy improvement problems with non-rectangular uncertainty sets to any accuracy $\epsilon > 0$ in $O(1/\epsilon^4)$ iterations. This is the first complete solution scheme for robust MDPs with non-rectangular uncertainty sets offering global optimality guarantees. A similar projected gradient descent algorithm with access to an abstract approximate robust policy evaluation oracle is described in [2]. However, the policy evaluation oracle is is not made explicit for general non-rectangular uncertainty sets. In addition, the convergence proof in [2] relies on the implicit assumption that the set of worst-case transition kernels for any given policy is finite, which would be difficult to certify in practice.

References

[1] Mengmeng Li, Daniel Kuhn, and Tobias Sutter. Policy gradient algorithms for robust mdps with non-rectangular uncertainty sets. *arXiv preprint, arXiv:2305.19004*, 2024.

[2] Qiuhao Wang, Chin Pang Ho, and Marek Petrik. Policy gradient in robust MDPs with global convergence guarantee. In *International Conference on Machine Learning*, 2023.

[3] Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. Robust Markov decision processes. *Mathematics of Operations Research*, 38(1):153–183, 2013.

Optimization and stabilization of infinite-dimensional dynamical systems

Some advances in variational approaches for data assimilation in hemodynamics

Jorge Tiago

jorge.tiago@tecnico.ulisboa.pt CEMAT and Department of Mathematics, Instituto Superior Técnico, ULisboa, Portugal

Numerical simulations have proved in the past decades to be an important tool to provide additional clinical insight in several cardiovascular pathologies [1, 2]. Some of the remaining challenges are related to the persecution of precise and personalized descriptions of blood flow. Integrating measured data into the simulations represents a means to achieve such a goal. Among several techniques, variational approaches have been tested with partial success [3,4]. These approaches consist of trying to obtain, as a control variable, meaningful boundary conditions at artificial boundaries. Here we are going to present several new advances in this direction. In particular, we address the cases of Dirichlet control, mixed with different stress-type boundary conditions. For the time-dependent case, we will present some numerical results for a semi-discrete approach which may substantially reduce the computational cost while still achieving good results.

Acknowledgements

The author was partially supported by Project HemoControl DOI 10.54499/PTDC/MAT-APL/7076/2020 from the FCT - Portuguese national funding agency for science, research and technology.

References

[1] Adam Updegrove et al., SimVascular - An open source pipeline for cardiovascular simulation, *Ann Biomed Eng 45*, 45: 525–541, 2017.

[2] Sara Gaur et al., Rationale and design of the HeartFlowNXT (HeartFlow analysis of coronary blood flow using CT angiography: NeXt sTeps) study. *J Cardiovasc Comput Tomogr*, 7(5):279–288, 2013.

[3] Sigmund Funke et al., Variational data assimilation for transient blood flow simulations: cerebral aneurysms as an illustrative example. *Int J Numer Method Biomed Eng*, 35: 3152, 2019.

[4] Telma Guerra et al., A data assimilation approach for non-newtonian blood flow simulations in 3D geometries. *Applied Mathematics and Computation*, 321: 176–194, 2018.

Optimization and stabilization of infinite-dimensional dynamical systems

Approximately optimal smooth feedback laws for non-differentiable value function using machine learning techniques

Karl Kunisch[1], **Donato Vásquez-Varas**[2]

1. karl.kunisch@uni-graz.at *University of Graz, Austria*
2. donato.vasquez-varas@ricam.oeaw.ac.at *Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austria*

The usage of machine learning methods for synthesizing feedback laws in the context of infinite horizon control problems has been a topic of increased interest, replacing, in part, the use of classical methods. This is due to the fact that the later suffer from the *curse of dimensionality*. Although numerical examples provide promising results, theoretical guarantees are still needed. In this regard, there are two types of methods which will be addressed in this talk.

Regarding the first type of methods, they consist in finding a feedback law in a finite dimensional functional space (polynomials, neural networks, RKHS) by minimizing the averaged cost functional of the control problem with respect to a set of initial conditions. In the case of the second type, the minimization is performed by minimizing the L^2 distance in the control space. Thanks to the results in [1] we are able to provide a convergence result which relies on the Hölder continuity of the value function and the existence of a Lyapunov type function. In the case of regression methods we are able to prove convergence, provided that the value function is smooth enough.

In addition, we present a family of infinite horizon optimal control problem for which the degree of smoothness of the value function depends on a penalty parameter. This dependence is such that the value function is C^2 if the penalty parameter is close to 0 and it is non-differentiable but Lipschitz if it is large. This allows us to study and compare the behavior of the methods, depending on the degree of smoothness of the value function by performing numerical realizations for both approaches for different penalty parameters.

References

[1] Karl Kunisch and Donato Vásquez-Varas. Smooth approximation of feedback laws for infinite horizon control problems with non-smooth value functions, 2023. arXiv:2312.11981.

Optimization and stabilization of infinite-dimensional dynamical systems

Conditional gradients for total variation regularization with PDE constraints: a graph cuts approach

Giacomo Cristinelli[1], **José A. Iglesias** [2], **Daniel Walter**[3]

1. g.cristinelli@utwente.nl *University of Twente, The Netherlands*
2. jose.iglesias@utwente.nl *University of Twente, The Netherlands*
3. daniel.walter@hu-berlin.de *Humboldt-Universität zu Berlin, Germany*

1. Introduction

Total variation regularization has proven to be a valuable tool in the context of optimal control of differential equations. This is particularly attributed to the observation that TV-penalties often favor piecewise constant minimizers with well-behaved jumpsets. On the downside, their intricate properties significantly complicate every aspect of their analysis, from the derivation of first-order optimality conditions to their discrete approximation and the choice of a suitable solution algorithm. In this talk, we discuss these topics for a general class of minimization problems of the form

$$\min_{u\in V} [F(Ku) + \mathrm{TV}(u,\Omega)]$$

where $V \subset L^q(\Omega)$ is a weakly closed linear subspace of $L^q(\Omega)$, $q = d/(d-1)$, F is a smooth fidelity term and K denotes a linear continuous but potentially costly to evaluate forward operator, e.g. stemming back to a PDE.

A possible approach to alleviating some of the difficulties caused by the TV-term can be found in its interpretation as the Minkowski functional of the set

$$B_{\hat{V}} = \left\{ u \in V \;\middle|\; \mathrm{TV}(u,\Omega) \le 1,\ \int_\Omega u(x)\,\mathrm{d}x = 0 \right\}$$

and studying the convex geometry of the latter, in particular its set of extremal points $\mathrm{Ext}(B_{\hat{V}})$. This leads to a variety of novel theoretical insights on minimization problems with total variation regularization as well as tools for their practical realization.

2. Contribution

We characterize the extremal points of the respective total variation unit balls for different choices of V, e.g. the "continuous" case, $V = L^q(\Omega)$, as well as piecewise linear continuous, $V = P_1(\mathcal{T}_h)$, and piecewise constant, $V = P_0(\mathcal{T}_h)$, discretizations on triangulations $\mathcal{T}_h$ of Ω. This enables the efficient solution of Problem by geometry exploiting fully-corrective generalized conditional gradient methods. Loosely speaking, this type of algorithms operates on two variables, an "active set" $\mathcal{A}_k$ comprising a finite number of characteristic functions as well as a sparse iterate u_k in the cone spanned by $\mathcal{A}_k$. Each iteration then consists of two subproblems, the solution of a linear minimization problem over a superset of $\mathrm{Ext}(B_{\hat{V}})$ in order to update $\mathcal{A}_k$ and a finite dimensional ℓ_1-type problem to improve the iterate.

A detailed account on the practical realization of such a method is given for piecewise constant finite element approximations. Second, in the same setting and for suitable sequences of uniformly refined meshes, it is shown that minimizers to discretized PDE-constrained optimal control problems approximate solutions to a continuous limit problem involving an anisotropic total variation reflecting the fine-scale geometry of the mesh.

References

[1] G. Cristinelli, J. A. Iglesias, D. Walter. Conditional gradients for total variation regularization with PDE constraints: a graph cuts approach https://arxiv.org/abs/2310.19777

Optimization methods for inverse problems and beyond

Exact sparse representation recovery in convex optimization

Marcello Carioni[1], **Leonardo Del Grande**[2]

1. m.c.carioni@utwente.nl Department of Applied Mathematics, University of Twente, The Netherlands
2. l.delgrande@utwente.nl Department of Applied Mathematics, University of Twente, The Netherlands

1. Introduction

In this talk, we present the results obtained in [2], where we developed a general theory for recovering the sparse representation of solutions to infinite-dimensional inverse problems regularized via convex functionals. We consider the minimization of a functional made of a fidelity term, penalizing the difference between the unknown and a noisy measurement, and a general convex, positively 1−homogeneous, coercive regularizer, that is responsible for enforcing sparsity on the solution. We will show that under a suitable non-degeneracy condition on the problem, for small regularization parameters and low noise levels, the minimizer is unique and sparse, namely it is uniquely represented as a linear combination of n extreme points of the ball of the regularizer, denoted by $\mathrm{Ext}(B)$ [1].

2. A metric non-degenerate source condition (MNDSC)

The key non-degeneracy condition that allows for exact sparse representation recovery extends the classical non-degeneracy source condition (NDSC) for total variation regularized inverse problems in the space of measures introduced in [3]. More precisely, our non-degeneracy condition is connected to the behaviour of the solution of the dual problem when tested against elements in $\overline{\mathrm{Ext}(B)}^{w^*}$, seen as a metric space when metrized by the weak* topology. This justifies the name Metric Non-Degenerate Source Condition (MNDSC). In particular, we require that for any pair of extreme points that are close enough, there exists a curve $\gamma : [0,1] \to \overline{\mathrm{Ext}(B)}^{w^*}$ that is connecting them and such that

$$\frac{d^2}{dt^2}\langle \eta_0, \gamma \rangle < 0$$

for every $t \in [0,1]$ and uniformly in the space of curves.

3. Applications

To showcase the generality of approach, in [2], we obtain explicit formulations of the MNDSC, which lead us to specific results of sparse recovery for three inverse problems of interest:

i) Total variation regularized deconvolution problems, where we show that the classical NDSC implies our MNDSC;

ii) 1-dimensional BV functions regularized with their BV-seminorm;

iii) Pairs of measures regularized with their mutual 1-Wasserstein distance.

References

[1] K. Bredies and M. Carioni. Sparsity of solutions for variational inverse problems with finite-dimensional data. *Calculus of Variations and Partial Differential Equations*, 59(1):14, 2020.

[2] M. Carioni and L. Del Grande. A General Theory for Exact Sparse Representation Recovery in Convex Optimization. *arXiv preprint arXiv:2311.08072*, 2023.

[3] V. Duval and G. Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5):1315–1355, 2014.

Optimization methods for inverse problems and beyond

Online optimization for dynamic electrical impedance tomography

Neil Dizon[1], **Jyrki Jauhiainen**[2], **Tuomo Valkonen**[3]

1. neil.dizon@helsinki.fi *University of Helsinki, Finland*
2. jyrki.jauhiainen@uef.fi *University of Eastern Finland, Finland*
3. tuomo.valkonen@iki.fi *Escuela Politécnica Nacional, Ecuador & University of Helsinki, Finland*

Online optimization generally studies the convergence of optimization methods as more data is introduced into the problem; think of deep learning as more training samples become available. We adapt the idea to dynamic inverse problems that naturally evolve in time. We introduce an improved primal-dual online method specifically suited to these problems, and demonstrate its performance on dynamic monitoring of electrical impedance tomography.

Acknowledgement

This research has been supported by the Academy of Finland decision 338614.

References

[1] Valkonen, *Predictive online optimisation with applications to optical flow*, Journal of Mathematical Imaging and Vision (2021), DOI: doi:10.1007/s10851-020-01000-4, ARXIV: 2002.03053.

Optimization methods for inverse problems and beyond

Modeling, analysis and solution of parameter identification problems in thermo-piezoelectricity

Raphael Kuess[1], **Daniel Walter**[2], **Andrea Walther**[3]

1. raphael.kuess@hu-berlin.de Humboldt Universität zu Berlin, Germany
2. daniel.walter@hu-berlin.de Humboldt Universität zu Berlin, Germany
3. andrea.walther@math.hu-berlin.de Humboldt Universität zu Berlin, Germany

Piezoelectric components are essential in various electronic devices, ranging from everyday items such as electronic toothbrushes and headphones to advanced medical and industrial applications such as ultrasound imaging and power generation systems. Thoroughly understanding the behaviour of these materials is essential, especially given their temperature-dependent characteristics.

Simplistically, the piezoelectric material is described by a linearly coupled PDE system for the mechanical displacement and the electrical potential, which can then be extended by non-linear constitutive laws in order to take its temperature-dependent behavior into account. Since many applications require high precision and material data provided by the manufacturers often deviate significantly from real data, a consistent and reproducible characterization of the temperature-dependent material parameter functions is of decisive importance to properly determine the material behavior.

Therefore, we will focus on the temperature-dependent inverse parameter identification problem for the system of piezoelectric partial differential equations based on measured and simulated observation data. In this context, we will analyze the underlying PDE system and discuss a generalized existence and uniqueness result, which also applies on the adjoint PDE. Based on that, we will investigate the forward operator of the inverse problem using the classical reduced approach. Finally we employ adapted optimization and regularization techniques for solving our linearized parameter identification problem. Thus, modeling, analyzing and solving this linearized inverse problem will be the main focus including numerical results.

Acknowledgements

This research was funded by the DFG FOR 5208: Model-based determination of nonlinear properties of piezoceramics for high-power ultrasound applications (NEPTUN).

Optimization methods for inverse problems and beyond

Accelerating Griffin Lim: A fast and provably converging method for phase retrieval

Peter Balazs[1], **Radu Ioan Boţ**[2], **Rossen Nenov**[1], **Dang-Khoa Nguyen**[2]

1. *Acoustics Research Institute, Vienna, Austria*
2. *University of Vienna, Austria*

1. Introduction

The recovery of a signal from the magnitudes of its transformation, like the Fourier transform, is known as the phase retrieval problem and is of big relevance in various fields of engineering and applied physics. In this talk, we present a fast inertial/momentum based algorithm for the phase retrieval problem. Our method can be seen as an extended algorithm of the Griffin-Lim Algorithm [**?**] and Fast Griffin-Lim Algorithm [**?**], methods originally designed for phase retrieval in acoustics. The new numerical algorithm can be applied to a more general framework than acoustics, and as a main result, we prove a convergence guarantee of the new scheme. Consequently, we also provide an affirmative answer for the convergence of its ancestor Fast Griffin-Lim Algorithm, whose convergence remained unproven in the past decade.

2. Contribution

A linear and injective transformation from $\mathbb{C}^L \to \mathbb{C}^M$ with $M \geq L$ can be written as a transformation matrix T in $\mathbb{C}^{M\times L}$ with full column rank, for example, the discrete Fourier transform. The vector $s \in [0,+\infty)^M$ will denote the measured magnitudes of the coefficients of the transform. The phase retrieval problem can be expressed mathematically as finding the signal $x^* \in \mathbb{C}^L$, whose transform coefficients match the magnitudes s, that is $|Tx^*| = s$, where $|\cdot|$ is understood as the absolute value applied componentwise. In [**?**] it was proposed to consider this problem, as the task in finding a vector c^* in the set of coefficients admitted by the transformation matrix T, namely its range, which is as close as possible to the set of coefficients, whose magnitude match with s, which can be written as

$$\min_{c\in C_1} \operatorname{dist}(c, C_2)^2, \tag{2.1}$$

where C_1 is a linear subspace defined as $C_1 = \{c \in \mathbb{C}^M \mid \exists x \in \mathbb{C}^L : c = Tx\}$ and C_2 is the compact nonconvex set $C_2 = \{c \in \mathbb{C}^M \mid |c_i| = s_i \quad \forall i \in \{1,\dots,M\}\}$.

We propose the Accelerated Griffin-Lim Algorithm for this optimization problem and analyze its convergence.

Algorithm 1 Accelerated Griffin-Lim algorithm

Initialize $c_0 \in \mathbb{C}^L$, $t_0, d_0 \in C_1$ and $\alpha, \beta, \gamma > 0$

Iterate for $n = 1,\dots,N$

$t_n = (1-\gamma)d_{n-1} + \gamma P_{C_1}(P_{C_2}(c_{n-1}))$

$c_n = t_n + \alpha(t_n - t_{n-1})$,

$d_n = t_n + \beta(t_n - t_{n-1})$

Return $T^\dagger c_N$

For $\gamma = 1$ this algorithm coincides with the Fast Griffin-Lim algorithm and for $\gamma = 1$ and $\alpha = 0$ it reduced to the Griffin-Lim algorithm.

Acknowledgements

This work was supported by the Austrian Science Fund FWF-project NoMASP ("Nonsmooth nonconvex optimization methods for acoustic signal processing"; P 34922-N).

Optimization methods for inverse problems and beyond

Learning data-driven priors for image reconstruction: from bilevel optimisation to neural network-based unrolled schemes

Kostas Papafitsoros

k.papafitsoros@qmul.ac.uk Queen Mary University of London, UK

Combining classical model-based variational methods for image reconstruction with deep learning techniques has attracted a significant amount of attention during the last years. The aim is to combine the interpretability and the reconstruction guarantees of a model-based method with the flexibility and the state-of-the-art reconstruction performance that the deep neural networks are capable of achieving. We introduce a general novel image reconstruction approach that achieves such a combination which we motivate by recent developments in deeply learned algorithm unrolling and data-driven regularisation as well as by bilevel optimisation schemes for regularisation parameter estimation. We consider a network consisting of two parts: The first part uses a highly expressive deep convolutional neural network (CNN) to estimate a spatially varying (and temporally varying for dynamic problems) regularisation parameter for a classical variational problem (e.g. Total Variation). The resulting parameter is fed to the second sub-network which unrolls a finite number of iterations of a method which solves the variational problem (e.g. PDHG). The overall network is then trained end-to-end in a supervised fashion. This results to an entirely interpretable algorithm since the "black-box" nature of the CNN is placed entirely on the regularisation parameter and not to the image itself. We prove consistency of the unrolled scheme by showing that, as the number of unrolled iterations tends to infinity, the unrolled energy functional used for the supervised learning Γ-converges to the corresponding functional that incorporates the exact solution map of the TV-minimization problem. We also provide a series of numerical examples that show the applicability of our approach: dynamic MRI reconstruction, quantitative MRI reconstruction, low-dose CT and dynamic image denoising.

References

[1] Andreas Kofler, Fabian Altekrüger, Fatima Antarou Ba, Christoph Kolbitsch, Evangelos Papoutsellis, David Schote, Clemens Sirotenko, Felix Frederik Zimmermann, Kostas Papafitsoros. Learning Regularization Parameter-Maps for Variational Image Reconstruction using Deep Neural Networks and Algorithm Unrolling, *SIAM Journal on Imaging Sciences*, 16(4):2202–2246, 2023.

Optimization methods for inverse problems and beyond

The impact of adjoint mismatches in the primal-dual Douglas-Rachford method – Existence of stationary points and convergence

Felix Schneppe[1,2], **Dirk A. Lorenz**[1]

1. {schneppe,d.lorenz}@uni-bremen.de Universität Bremen, Germany
2. f.schneppe@tu-braunschweig.de Technische Universität Braunschweig, Germany

The primal-dual Douglas-Rachford method is a well known method to solve saddle-point problems of the form

$$\min_x \max_y G(x) + \langle Ax, y\rangle - F^*(y).$$

However, in practical applications like computed tomography the adjoint operator is often replaced by a computationally more efficient approximation. This leads to an adjoint mismatch in the algorithm, which translates to a replacement of A^* with some linear operator V^* in the algorithm and to it being changed to

$$\begin{aligned} x^{k+1} &= \operatorname{prox}_{\tau G}\left(p^k\right) \\ z^{k+1} &= \operatorname{prox}_{\sigma F^*}\left(q^k\right) \\ \begin{pmatrix} u^{k+1} \\ v^{k+1} \end{pmatrix} &= \begin{pmatrix} I & \tau V^* \\ -\sigma A & I \end{pmatrix}^{-1} \begin{pmatrix} 2x^{k+1} - p^k \\ 2z^{k+1} - q^k \end{pmatrix} \\ p^{k+1} &= p^k + \theta\left(u^{k+1} - x^{k+1}\right) \\ q^{k+1} &= q^k + \theta\left(v^{k+1} - z^{k+1}\right). \end{aligned}$$

In this talk, we analyze the convergence properties of the primal-dual Douglas-Rachford method with the adjoint mismatch and prove conditions, under which the existence of a solution can still be guaranteed. Additionally, we observe methods to calculate the step sizes of the algorithm. Furthermore we discuss, how the adjoint mismatch affects the implementation of the algorithm.

Optimization methods for inverse problems and beyond

Interweaved first-order methods for PDE-constrained and bilevel optimisation

Tuomo Valkonen[1], **Ensio Suonperä**[2], **Bjørn Jensen**[3]

1. tuomo.valkonen@iki.fi *Escuela Politécnica Nacional, Ecuador & University of Helsinki, Finland*
2. *University of Helsinki, Finland*
3. *University of Jyväskylä, Finland*

PDE-constrained optimisation problems, as well as bilevel optimisation problems—which cover nonsmooth PDE constraints, such as Bingham flow—have traditionally been solved with one of two approaches: (a) Newton-type methods applied to sufficiently smooth optimality conditions for a constrained problem formulation, or (b) through treating the inner problem/PDE through its solution mapping, always calculating the solution mapping and its derivative near-exactly. Recently in bilevel optimisation, single-loop approaches have been introduced, which only take a single step of a conventional optimisation method for the inner problem, bridging the gap between the two approaches. The same principle can be applied to PDE-constrained optimisation, where we have recently obtained significant performance improvements by interweaving the steps of a conventional iterative solver (Jacobi, Gauss–Seidel, conjugate gradients) for the PDE with the steps of the optimisation method. Moreover, in this talk, we demonstrate how the adjoint equation in bilevel problems can also benefit from such interweaving with conventional linear system solvers.

Acknowledgements

This research has been supported by the Academy of Finland decisions 314701, 320022, and 345389.

References

[1] Suonperä and Valkonen, *Linearly convergent bilevel optimization with single-step inner methods*, Computational Optimization and Applications (2023), DOI: doi:10.1007/s10589-023-00527-7, ARXIV: 2205.04862.

[2] Jensen and Valkonen, *A nonsmooth primal-dual method with simultaneous adaptive PDE constraint solver* (2022), ARXIV: 2205.04862.

Optimization methods with worst-case complexity guarantees

A derivative-free trust-region method based on finite-differences for composite nonsmooth optimization

Dânâ Davar
dana.davar@uclouvain.be UCLouvain, Belgium

We present a derivative-free trust-region method based on finite-difference gradient approximations for minimizing composite functions of the form $F(x) = h(c(x))$, where h is a convex Lipschitz function, possibly nonsmooth, while c is a blackbox function with Lipschitz continuous Jacobian. The proposed method approximates the Jacobian of c by forward-finite differences with stepsizes depending on an estimate of its Lipschitz constant. Such estimate is also used in the definition of the trust-region radius, allowing natural update rules to enforce sufficient decrease of the objective function. It is shown that the method needs at most $\mathcal{O}(n\epsilon^{-2})$ function evaluations to find a ϵ-approximate stationary point. In addition, if the components of c are convex and h is monotone, the worst-case evaluation complexity is improved to $\mathcal{O}(n\epsilon^{-1})$. Numerical results are also reported, showing the relative efficiency of the new method with respect existing derivative-free solvers.

Optimization methods with worst-case complexity guarantees

Polynomial preconditioning for Gradient Methods

Nikita Doikov
nikita.doikov@epfl.ch EPFL, Switzerland

We study first-order methods with preconditioning for solving structured nonlinear convex optimization problems. We propose a new family of preconditioners generated by symmetric polynomials. They provide first-order optimization methods with a provable improvement of the condition number, cutting the gaps between highest eigenvalues, without explicit knowledge of the actual spectrum. We give a stochastic interpretation of this preconditioning in terms of coordinate volume sampling and compare it with other classical approaches, including the Chebyshev polynomials. We show how to incorporate a polynomial preconditioning into the Gradient and Fast Gradient Methods and establish the corresponding global complexity bounds. Finally, we propose a simple adaptive search procedure that automatically chooses the best possible polynomial preconditioning for the Gradient Method, minimizing the objective along a low-dimensional Krylov subspace. Numerical experiments confirm the efficiency of our preconditioning strategies for solving various machine learning problems.

Optimization methods with worst-case complexity guarantees

Decentralized local stochastic extra-gradient for variational inequalities

Pavel Dvurechensky

Pavel.Dvurechensky@wias-berlin.de WIAS, Germany

We consider distributed stochastic variational inequalities (VIs) on unbounded domains with the problem data that is heterogeneous (non-IID) and distributed across many devices. We make a very general assumption on the computational network that, in particular, covers the settings of fully decentralized calculations with time-varying networks and centralized topologies commonly used in Federated Learning. Moreover, multiple local updates on the workers can be made to reduce the communication frequency between the workers. We extend the stochastic extragradient method to this very general setting and theoretically analyze its convergence rate in the strongly-monotone, monotone, and non-monotone (when a Minty solution exists) settings. The provided rates explicitly exhibit the dependence on network characteristics (e.g., mixing time), iteration counter, data heterogeneity, variance, number of devices, and other standard parameters. As a special case, our method and analysis apply to distributed stochastic saddle-point problems (SPP), e.g., to the training of Deep Generative Adversarial Networks (GANs) for which decentralized training has been reported to be extremely challenging. In experiments for the decentralized training of GANs we demonstrate the effectiveness of our proposed approach.

Optimization methods with worst-case complexity guarantees

Adaptive first-order methods with enhanced worst-case rates

Mihai Iulian Florea
mihai.florea@uclouvain.be UCLouvain, Belgium

The Optimized Gradient Method (OGM), its strongly convex extension, the Information Theoretical Exact Method (ITEM), as well as the related Triple Momentum Method (TMM) have superior convergence guarantees when compared to the Fast Gradient Method but lack adaptivity and their derivation is incompatible with composite problems. In this work we introduce a slightly modified version of the estimate sequence that can be used to simultaneously derive OGM, ITEM and TMM while adding memory along with the ability to dynamically adjust the convergence guarantees at runtime. Our framework can be extended to the composite setup and we use it to construct an Enhanced Accelerated Composite Gradient Method equipped with fully-adaptive line-search.

Optimization methods with worst-case complexity guarantees

First and zeroth-order implementations of the regularized Newton method with lazy approximated Hessians

Geovani Grapiglia

geovani.grapiglia@uclouvain.be UCLouvain, Belgium

We develop first-order (Hessian-free) and zeroth-order (derivative-free) implementations of the Cubically regularized Newton method for solving general non-convex optimization problems. For that, we employ finite difference approximations of the derivatives. We use a special adaptive search procedure in our algorithms, which simultaneously fits both the regularization constant and the parameters of the finite difference approximations. It makes our schemes free from the need to know the actual Lipschitz constants. Additionally, we equip our algorithms with the lazy Hessian update that reuse a previously computed Hessian approximation matrix for several iterations. Specifically, we prove the global complexity bound of $\mathcal{O}(n^{1/2}\epsilon^{-3/2})$ function and gradient evaluations for our new Hessian-free method, and a bound of $\mathcal{O}(n^{3/2}\epsilon^{-3/2})$ function evaluations for the derivative-free method, where n is the dimension of the problem and ϵ is the desired accuracy for the gradient norm. These complexity bounds significantly improve the previously known ones in terms of the joint dependence on n and ϵ, for the first-order and zeroth-order non-convex optimization.

Optimization methods with worst-case complexity guarantees

Universal gradient methods for Stochastic Convex Optimization

Anton Rodomanov
anton.rodomanov@cispa.de CISPA, Germany

We propose a new rule for adjusting the step size in the Stochastic Gradient method (SGD). This rule is related to that of the AdaGrad method but there are some significant differences. Most importantly, instead of using the norms of stochastic gradients, we use a stochastic approximation of the Bregman distance of the objective function. The resulting algorithm turns out to be the first universal method for Stochastic Convex Optimization in the sense that it automatically adjusts not only to the oracle's noise level but also to the level of smoothness of the objective function. More specifically, our method has state-of-the-art worst-case convergence rate guarantees for the entire Hölder class of convex functions including both nonsmooth functions and those with Lipschitz continuous gradient. We also show how to use our approach for constructing an accelerated version of the Universal SGD with even better efficiency estimates.

Optimization methods with worst-case complexity guarantees

Non-convex stochastic composite optimization with Polyak momentum

Sebastian U. Stich
stich@cispa.de CISPA, Germany

The stochastic proximal gradient method, a powerful generalization of the widely used stochastic gradient descent (SGD) method, has numerous applications in Machine Learning. However, this method can fail to converge in non-convex settings with high stochastic noise. Unlike in unconstrained optimization scenarios, these convergence issues cannot be resolved by simply reducing the step size. The practicable countermeasures are limited to using variance reduction techniques or increasing the batch size, leading to increased costs per iteration. In this talk, we focus on the stochastic proximal gradient method with Polyak momentum. We show that this method attains an optimal convergence rate for non-convex composite optimization problems, regardless of the batch size. Furthermore, we discuss extension with inexact proximal steps and applications in distributed optimization.

Optimization methods with worst-case complexity guarantees

Safe primal-dual optimization with a single smooth constraint

Ilnura Usmanova
ilnura94@gmail.com Paul Scherrer Institute, Switzerland

We address the problem of safe learning in a specific setting of a single smooth constraint. In particular, the goal is to solve a black-box minimization problem without constraint violation during the learning process. Unfortunately, most existing approaches addressing this problem are slow or noise-sensitive, even for simple settings. We propose a new primal-dual safe learning method to effectively minimize a smooth objective subject to a single smooth constraint. Despite its simplicity, this particular case covers a wide range of applications, such as robotics, adversarial learning, or empirical parameter tuning. We show that our primal-dual algorithm can achieve a convergence rate significantly better than current state-of-the-art approaches. We demonstrate its performance on simulations.

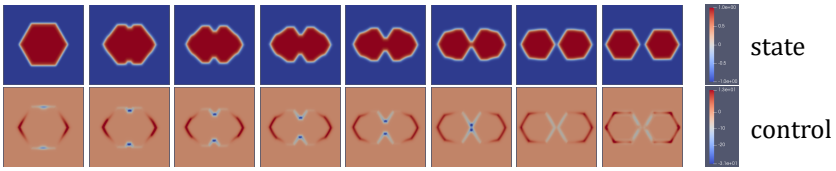
RECENT TRENDS IN NONSMOOTH OPTIMIZATION

Optimal control of anisotropic Allen-Cahn equations

Luise Blank[1], **Johannes Meisinger**[2]

1. luise.blank@ur.de, *Universität Regensburg, Germany*
2. johannes.meisinger@ur.de, *Universität Regensburg, Germany*

Anisotropic Allen-Cahn equations model for example the interface evolution in crystal growth. The diffuse interface approach allows for topology changes. The focus of the talk is the control of such an evolution, e.g. controlling the evolution of a hexagon such that it splits into two prescribed ones.



This leads to an optimization problem with a given end time goal function y_Ω

$$\min \frac{1}{2}\|y(T) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2\varepsilon}\|u\|^2_{L^2(Q)}$$

subject to the quasilinear nonsmooth parabolic equation

$$\int_Q \varepsilon \partial_t y \eta + \varepsilon A'(\nabla y)^T \nabla \eta + \frac{1}{\varepsilon}\psi'(y)\eta = \int_Q u\eta \quad \forall \eta \in L^2(0,T;H^1(\Omega)), \qquad y(0) = y_0 \text{ in } \Omega,$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain and $Q := [0,T] \times \Omega$.

Due to the anisotropy A nondifferentiable terms occur in the state equation which have to be treated appropriately to obtain an efficient optimization solver. Here we propose a particular regularization. The function ψ is nonconvex and e.g. a smooth double well potential. Moreover, the issue of differentiability also leads us to choose the implicit time discretization dG(0) where in addition energy stability is obtained. In this case 'first optimize then discretize' commutes.

First the existence of the control-to-state operator and its Lipschitz-continuity is shown for the time discretized as well as for the time continuous problem. The existence of a global minimizer of the original and of the regularized problem is provided in the time discretized as well as in the continuous setting. Also the convergence with respect to regularization and to discretization is considered. Furthermore the Fréchet differentiability of the regularized problem is studied and optimality conditions are obtained.

Subsequently the trust-region Newton Steihaug-cg method is applied to the time discretized problem which is then discretized in space. This provides us with iteration numbers bounded independently of the discretization level, where evidence is given numerically. Finally numerical examples with various anisotropies and configurations are presented.

Acknowledgements

The second author was supported by the RTG IntComSin of the German Research Foundation.

References

[1] L. Blank, J. Meisinger. Optimal control of anisotropic Allen-Cahn equations. *ESAIM: COCV 28, 71, 2022*

[2] L. Blank, J. Meisinger. Parabolic optimal control with strongly monotone quasilinearity and its time discretization. *Applied Mathematics and Optimization 86, 34, 2022.*

Recent trends in nonsmooth optimization

First- and second-order models for nonsmooth functions based on derivative sampling

Bennet Gebken
bgebken@math.upb.de Paderborn University, Germany

In optimization, many solution methods are based on iteratively building a local model of the objective function and then minimizing the model instead of the original function. In the smooth case, such models can be derived from the Taylor expansion based on derivatives of different orders, leading to the gradient descent method and Newton's method as examples. In the nonsmooth case on the other hand, models are more challenging to construct: The first issue is the lack of a Taylor expansion. While replacing the gradient in smooth first-order methods by the Clarke subdifferential from nonsmooth analysis yields a way to characterize descent directions, there is no simple way to derive higher-order models. The second issue is that generalized derivatives like the Clarke subdifferential are difficult to work with in practice, since they can be unstable and impossible to evaluate numerically in the general case.

In this talk, I will present two models for (unstructured) locally Lipschitz continuous functions. The first model is a simple first-order model, which is based on approximating the Clarke subdifferential by the Goldstein ε-subdifferential which, in turn, can be approximated in practice by a deterministic gradient sampling approach. The second model is based on the idea of sampling the Hessian matrix in addition to the gradient. More precisely, it is defined as the maximum of (existing) second-order Taylor expansions in a neighborhood of a given point. After introducing each model, I will present ways to generate them in practice and discuss the behavior of the resulting descent methods.

RECENT TRENDS IN NONSMOOTH OPTIMIZATION

Spatially sparse optimization problems in fractional order Sobolev spaces

Anna Lentz[1], **Daniel Wachsmuth**[2]

1. anna.lentz@uni-wuerzburg.de Universität Würzburg, Germany
2. daniel.wachsmuth@uni-wuerzburg.de Universität Würzburg, Germany

We consider time-dependent optimization problems in fractional Sobolev spaces H^s for $s \in (0,1)$ with spatially sparse solutions. For a time-space cylinder $I \times \Omega$, the optimization problem reads

$$\min_{u \in L^2(I\times\Omega), w \in H^s(\Omega)} f(u) + \frac{\alpha}{2}\|u\|^2_{L^2(I\times\Omega)} + \frac{\beta}{2}\|w\|^2_{H^s(\Omega)} + \gamma\|w\|^p_{L^p(\Omega)} \tag{1}$$

subject to

$$|u(t,x)| \le w(x) \qquad \text{for a.a. } (t,x) \in I \times \Omega \tag{2}$$

for some $\alpha, \beta, \gamma > 0$ and $p \in (0,1)$ [3]. Due to the compact embedding of $H^s(\Omega)$ into $L^2(\Omega)$, one can deduce existence of solutions by standard methods. This is not possible if the problem was set only in L^2 [2].
The formulation with two functions u and w coupled by the constraint (2) is used to avoid the computationally more expensive regularization in H^s over the whole time-space cylinder $I \times \Omega$. The non-smooth and non-convex L^p-pseudonorm leads to sparsity of solutions w which in turn implies spatially sparse solutions u.
To obtain a necessary optimality condition, a sequence of auxiliary problems is analysed. These auxiliary problems are unconstrained with differentiable objective as a consequence of penalizing the constraint (2) and replacing the L^p-pseudonorm by a smooth approximation [1]. Passing to the limit in the optimality condition of the auxiliary problem results in a necessary optimality condition for the original problem (1).
This optimality condition states that for a local solution $(\bar u, \bar w)$ of problem (1) there are $\bar\lambda \in H^s(\Omega)^*$ and $\bar\mu^1, \bar\mu^2 \in L^2(I\times\Omega)$ such that

$$f'(\bar u)v + \alpha(\bar u, v)_{L^2(I\times\Omega)} + \beta(\bar w, z)_{H^s(\Omega)} + \gamma\langle\bar\lambda, z\rangle_{H^s(\Omega)^*} + (\bar\mu^1, z - v)_{L^2(I\times\Omega)} + (\bar\mu^2, z + v)_{L^2(I\times\Omega)} = 0$$

for all $(v,z) \in L^2(I\times\Omega) \times H^s(\Omega)$ and it holds

$$\langle\bar\lambda, \bar w\rangle_{H^s(\Omega)^*} = p\int_\Omega |\bar w|^p \mathrm{d}x.$$

Furthermore, the following complementarity system is satisfied:

$$\begin{aligned} &\bar\mu^1 \le 0, \quad \bar w - \bar u \ge 0, \quad (\bar\mu^1, \bar w - \bar u)_{L^2(I\times\Omega)} = 0 \\ \text{and} \quad &\bar\mu^2 \le 0, \quad \bar w + \bar u \ge 0, \quad (\bar\mu^2, \bar w + \bar u)_{L^2(I\times\Omega)} = 0. \end{aligned}$$

Problem (1) is solved numerically with a DC-like algorithm. Weak accumulation points of the sequence of iterates of that algorithm satisfy a slightly weaker stationarity system than the one stated above.

Acknowledgements

This research was partially supported by the German Research Foundation DFG under project grant Wa 3626/5-1.

References

[1] H. Antil and D. Wachsmuth. Sparse optimization problems in fractional order Sobolev spaces. *Inverse Problems*, 39(4):Paper No. 044001, 17, 2023. URL: https://dx.doi.org/10.1088/1361-6420/acbe5e.

[2] K. Ito and K. Kunisch. Optimal control with $L^p(\Omega)$, $p \in [0,1)$, control cost. *SIAM J. Control Optim.*, 52(2):1251–1275, 2014. doi:10.1137/120896529.

[3] A. Lentz and D. Wachsmuth. Spatially sparse optimization problems in fractional order Sobolev spaces. *(Preprint)* 2024. arXiv:2402.14417.

Recent trends in nonsmooth optimization

On solving complementarity-constrained problems

Timo Kreimeier[1], **Marc Pfetsch**[2], **Adrian Schmidt**[3], **Andrea Walther**[4]

1. timo.kreimeier@hu-berlin.de Humboldt-Universität zu Berlin, Germany
2. pfetsch@mathematik.tu-darmstadt.de Technische Universität Darmstadt, Germany
3. adrian.schmidt@hu-berlin.de Humboldt-Universität zu Berlin, Germany
4. andrea.walther@math.hu-berlin.de Humboldt-Universität zu Berlin, Germany

In this talk, we consider finding global minima for constrained piecewise linear optimization problems. One particular type of problem that fits this setting is complementarity-constrained problems. Because they can be expressed in terms of the absolute value function, they can also be represented in the so-called abs-linear form - a matrix-vector representation [2, 3]. This allows us to handle them with the Constrained Active Signature Method, a solver for constrained piecewise linear optimization problems [4].

We present the incorporation of this algorithm into a global solver like SCIP to determine not only local minima, but also global minima, utilizing a branch-and-cut approach on so-called SOS1-constraints [1].

Acknowledgements

The authors gratefully acknowledge the support of Project B10 in the Collaborative Research Center/Transregio 154 Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks (project ID: 239904186) by the Deutsche Forschungsgemeinschaft.

References

[1] Tobias Fischer and Marc E. Pfetsch. Branch-and-cut for linear programs with overlapping sos1 constraints. *Mathematical Programming Computation*, 10(1):33–68, 2018. doi:10.1007/s12532-017-0122-5.

[2] Andreas Griewank. On stable piecewise linearization and generalized algorithmic differentiation. *Optimization Methods and Software*, 28(6):1139–1178, 2013. doi:10.1080/10556788.2013.796683.

[3] Andreas Griewank and Andrea Walther. Polyhedral dc decomposition and dca optimization of piecewise linear functions. *Algorithms*, 13(7), 2020. doi:10.3390/a13070166.

[4] Timo Kreimeier, Andrea Walther, and Andreas Griewank. An active signature method for constrained abs-linear minimization. 2023. URL: https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/474.

SHAPE OPTIMIZATION AND UNCERTAINTY

Eigenvalue optimization with respect to shape-variations in electromagnetic cavities

Christine Herter[1], **Sebastian Schöps**[2], **Winnifried Wollner**[3]

1*. christine.herter@uni-hamburg.de Department of Mathematics, University of Hamburg, Germany
2. sebastian.schoeps@tu-darmstadt.de Computational Electromagnetics Group, Technische Universität Darmstadt, Germany
3. winnifried.wollner@uni-hamburg.de Department of Mathematics, University of Hamburg, Germany

Keywords: Maxwell's Eigenvalue Problem, Eigenvalue Optimization, Shape Optimization.

1. Introduction

Particle accelerator cavities transfer energy to a charged particle beam, where the final performance of accurate simulation results is sensitive to small deformation changes. To find the optimal geometry, we consider a freeform optimization problem of eigenvalues in such a cavity by means of shape-variations with respect to small deformations. As constraint we utilize the mixed variational formulation by Kikuchi of the normalized Maxwell's time-harmonic eigenvalue problem, where we control the deformation of the domain by applying the method of mappings.

2. Shape Optimization Problem

We consider a simply connected Lipschitz domain $\mathfrak{Q} \subset \mathbb{R}^d$, $d \in \{2,3\}$, and we define the physical domain $\Omega_q = F_q(\mathfrak{Q}) \subset \mathbb{R}^d$, where $F_q \colon \mathfrak{Q} \to \mathbb{R}^d$ is given by $x \mapsto x + q(x)$ for a to be determined $q \in Q = W^{1,\infty}(\mathfrak{Q})$. By that, we define the following shape optimization problem constrained by the variational formulation by Kikuchi of Maxwell's eigenvalue problem: Find $\lambda \in \mathbb{R}$, $0 \neq u \in H_0(\mathrm{curl};\Omega_q)$ and $\psi \in H_0^1(\Omega_q)$, such that

$$\begin{aligned} \min_{(\lambda,q)} J(\lambda,q) &:= \frac{1}{2}|\lambda-\lambda_*|^2 + \frac{\alpha}{2}\left(\|q\|^2 + \|\nabla q\|^2\right) - \beta \int_{\Omega_q} \ln(\det(\mathrm{DF}_q) - \varepsilon) \\ \text{s.t. } (\nabla\times u, \nabla\times v)_{\Omega_q} + (\nabla\psi, v)_{\Omega_q} &= \lambda (u,v)_{\Omega_q}, \\ (u, \nabla\varphi)_{\Omega_q} &= 0, \\ \chi\left((u,u)_{\Omega_q} - 1\right) &= 0, \end{aligned} \tag{2.1}$$

for all $v \in H_0(\mathrm{curl};\Omega_q)$, $\varphi \in H_0^1(\Omega_q)$ and $\chi \in \mathbb{R}$. Further $(\cdot,\cdot)_{\Omega_q}$ denotes the scalar product depending on the domain Ω_q, $\|\cdot\|$ is the usual L^2 norm on $\mathfrak{Q}$, $\lambda_* \in \mathbb{R}$ is the target eigenvalue and $\alpha, \beta \in \mathbb{R}_+$ are the regularization parameters and DF_q is a deformation gradient.

3. Contribution and Outlook

We show results on the analysis of the time-harmonic Maxwell's eigenvalue problem in order to prove continuity and differentiability properties depending on a domain mapping. Further, to find an optimal solution, we compute the derivative of the reduced cost functional by using adjoint calculus. We solve the considered optimization problem by using a damped inverse BFGS method. We conclude with a numerical example, which show the functionality of the optimization method.

Open tasks on the application are e.g. an optimization with respect to the corresponding eigenvector of the particular eigenvalue which will be relevant for an optimization of the field flatness, see e.g. [1]. Further, the extension of the implementation of a three-dimensional Maxwell's eigenvalue problem is an area of interest, which is either possibly by extending the domain itself to 3D or by taking the advantage of rotation symmetry of the two-dimensional geometry.

References

[1] Jacopo Corno. *Numerical Methods for the Estimation of the Impact of Geometric Uncertainties on the Performance of Electromagnetic Devices*. PhD thesis, Technische Universität, Darmstadt, 2017. URL: http://tuprints.ulb.tu-darmstadt.de/7038/.

[2] Christine Herter, Sebastian Schöps, and Winnifried Wollner. Eigenvalue optimization with respect to shape-variations in electromagnetic cavities. *PAMM*, 22(1):e202200122, 2023. doi:10.1002/pamm.202200122.

[3] M. J. D. Powell. A fast algorithm for nonlinearly constrained optimization calculations. In *Numerical analysis (Proc. 7th Biennial Conf., Univ. Dundee, Dundee, 1977)*, volume 630 of *Lecture Notes in Math.*, 1978. doi:10.1007/BFb0067703.

SHAPE OPTIMIZATION AND UNCERTAINTY

Multi-level optimal control with neural surrogate models

Dante Kalise[1], **Estefania Loayza-Romero**[2], **Kirsten A. Morris**[3] , **Zhengang Zhong**[4]

1. d.kalise-balza@imperial.ac.uk *Department of Mathematics, Imperial College London, United Kingdom*
2. k.loayza-romero@imperial.ac.uk *Department of Mathematics, Imperial College London, United Kingdom*
3. kmorris@uwaterloo.ca *Department of Applied Mathematics, University of Waterloo, Canada*
4. z.zhong20@imperial.ac.uk *Centre for Process Systems Engineering, Imperial College London, United Kingdom*

1. Introduction

Optimal actuator and control design is studied as a multi-level optimization problem, where the actuator design is evaluated based on the performance of the associated optimal closed loop. The evaluation of the optimal closed loop for a given actuator realisation is a computationally demanding task, for which the use of a neural network surrogate is proposed. The use of neural network surrogates to replace the lower level of the optimization hierarchy enables the use of fast gradient-based and gradient-free consensus-based optimization methods to determine the optimal actuator design. The effectiveness of the proposed surrogate models and optimization methods is assessed in a test related to optimal actuator location for heat control.

2. Problem formulation

We consider parameter-dependent linear dynamical systems described by

$$\frac{\mathrm{d}z}{\mathrm{d}t} = Az(t) + B(r)u(t), \quad z(0) = z_0 \tag{2.1}$$

where $A \in \mathbb{R}^{n\times n}$, and $B(r) \in \mathbb{R}^{n\times m}$ is matrix-valued function depending on a parameter $r \in \mathbb{R}^m$. Such systems naturally arise after semi-discretization in space of systems governed by PDEs, where r parametrizes the location of m actuators. We assume that each parameter coordinate can be varied over some compact set $\Omega \subset \mathbb{R}$.

The linear quadratic controller design aims at finding a minimizing control $u(t) \in \mathbb{R}^m$ to the cost functional

$$J(u; z_0, r) = \int_0^\infty z(t)^\top Q z(t) + u(t)^\top R u(t), \tag{2.2}$$

where $Q \in \mathbb{R}^{n\times n}$, $Q \succeq 0$, $R \in \mathbb{R}^{m\times m}$, $R \succ 0$, and $z(t) \in \mathbb{R}^n$ is determined by the dynamics (2.1).

The optimal cost-to-go or value function for a given initial condition z_0 and parameter r is

$$V(z_0, r) \coloneqq \inf_{u\in\mathcal{U}} J(u; z_0, r) = z_0^\top \Pi(r) z_0, \tag{2.3}$$

with $\mathcal{U} = L^2((0,+\infty); \mathbb{R}^m)$, and where $\Pi(r)$ solves the parameter-dependent Algebraic Riccati Equation (ARE)

$$A^\top \Pi(r) + \Pi(r)A - \Pi(r)B(r)R^{-1}B(r)^\top \Pi(r) + Q = 0. \tag{2.4}$$

The simultaneous optimization of the control signal and the parameter, together with an optimality-based characterisation of the space of initial conditions of interest, induces a hierarchy of costs which are cast as a multi-level optimization problem

$$\max_{z_0\in\mathbb{R}^n} \min_{r\in\Omega^m} \min_{u\in\mathcal{U}} J(u; z_0, r) = \max_{z_0\in\mathbb{R}^n} \min_{r\in\Omega^m} V(z_0, r). \tag{2.5}$$

To alleviate the need for recurrently computing the solution of the Algebraic Riccati Equation (ARE) given by (2.4) at every evaluation of $V(z_0, r)$ within an algorithm for the outer max-min problem, we propose to build a closed-form surrogate for $V(z_0, r)$ using neural networks. We consider the construction of both unstructured and structured surrogates and assess their efficacy through a representative example. Subsequently, we aim to tackle the multi-level optimization problem by employing projected gradient descent ascent and a consensus-based approach for handling saddle point problems.

Acknowledgements

This research was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/T024429/1 and a Discovery Grant from NSERC (Canada).

Shape optimization and uncertainty

Distributionally robust shape optimization with Wasserstein distance and applications

Charles Dapogny , Julien Prando[1], **Boris Thibert**

julien.prando@etu.univ-grenoble-alpes.fr Université Grenoble Alpes, France

Shape optimization is about finding the best design of a mechanical structure according to a performance criterion. Those problems are described with the help of physical parameters such as material coefficients or viscosity which are difficult to estimate. It is an issue because the optimal design for a set of parameters may worsen under a slight perturbation of these parameters.

Several models have been developed to handle this type of instability starting from worst case approaches [1] where the uncertain parameters are known up to a certain amplitude. One of the issues of these approaches is the min-max structure which is theoretically difficult to handle. Also, the robust optimal design performance with respect to the reference parameters may be too pessimistic because worst-case situations will usually be unlikely. Hence, stochastic approaches have been considered [4] by assuming these parameters to be random. However, these stochastic models rely on the precise knowledge of the probability distribution of the uncertain parameters which is usually unavailable.

In the context of convex optimization, methods have been developed to overcome the lack of knowledge of this probability distribution. In the distributionally robust optimization approach, the optimization is usually done with respect to a worst case distribution close to the empirical law built from observations. The Wasserstein distance has been used to characterize the notion of proximity [5] leading to an efficient tractable formulation under reasonable assumptions [2].

In this talk, we will explore the use of distributionally robust optimization approaches in the context of shape and topology optimization [3]. Numerical examples will be discussed illustrating that optimal distributionally robust solutions indeed yield good performances with respect to the reference parameters while handling likely worst-case realisations.

References

[1] Grégoire Allaire and Charles Dapogny. A linearized approach to worst-case design in parametric and geometric shape optimization. *Mathematical Models and Methods in Applied Sciences*, 24(11):2199–2257, 2014.

[2] Waïss Azizian, Franck Iutzeler, and Jérôme Malick. Regularization for wasserstein distributionally robust optimization. *arXiv preprint arXiv:2205.08826*, 2022.

[3] Charles Dapogny, Franck Iutzeler, Andrea Meda, and Boris Thibert. Entropy-regularized wasserstein distributionally robust shape and topology optimization. *Structural and Multidisciplinary Optimization*, 66(3):42, 2023.

[4] Subhayan De, Jerrad Hampton, Kurt Maute, and Alireza Doostan. Topology optimization under uncertainty using a stochastic gradient-based approach, 2019. arXiv:1902.04562.

[5] Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In *Operations research & management science in the age of analytics*, pages 130–166. Informs, 2019.

Shape optimization and uncertainty

Shape optimization on Riemannian manifolds

Lidiya Pryymak[1], **Kathrin Welker**[2]

1. *Lidiya.Pryymak@math.tu-freiberg.de TU Bergakademie Freiberg, Germany*
2. *Kathrin.Welker@math.tu-freiberg.de TU Bergakademie Freiberg, Germany*

Abstract spaces, where each point describes some arbitrary shape, are called shape spaces. Often, these shapes are described by smooth 2D curves on S^1 and the corresponding shape space is denoted by $B_e(S^1, \mathbb{R}^2)$. Shape spaces are commonly used for applications in image processing, flag algebra and shape optimization, to only name a few. In this talk, we mainly focus on shape optimization, whose applications include the design of medical devices, such as cardiovascular stents and engineering applications, like the construction of aerodynamic wings. The structure of the underlying shape space has a great impact on the outcoming shape optimization algorithm. Often, the shape space is equipped with a Riemannian structure. In this talk, we consider various Riemannian metrics and look into the inherited geodesic equation. Hereby, we focus on the Steklov-Poincaré metric.

Important extensions of shape optimization methods on $B_e(S^1, \mathbb{R}^2)$ include the optimization of multiple shapes and the optimization of piecewise-smooth shapes. Therefore, we introduce a novel shape space which can be indentified with a Riemannian product manifold. Finally, we provide a numerical result regarding a fluid-mechanical problem constrained by the Navier-Stokes equations.

SHAPE OPTIMIZATION AND UNCERTAINTY

Second shape derivative for an interface identification problem constrained by nonlocal models

Volker Schulz[1], **Matthias Schuster**[2]
1. volker.schulz@uni-trier.de Universität Trier, Germany
2. schusterm@uni-trier.de Universität Trier, Germany

1. Introduction

Nonlocal operators are typically integral operators and therefore allow interactions between two distinct points in space. Moreover, the solution of a nonlocal problem in general needs to fulfill less regularity requirements compared to the solution of a system of partial differential equations. Thus, some physical phenomena like fracture propagation in continuum mechanics or anomalous diffusion effects can be modeled more accurately by nonlocal equations compared to their 'classic' description by partial differential equations(see, e.g., [1, 3]). Since shape optimization methods have been proven useful for identifying interfaces in models governed by partial differential equations, we show in this talk how shape optimization techniques can also be applied to an interface identification problem constrained by a nonlocal Dirichlet problem. Here, we focus on deriving the second shape derivative of the corresponding reduced functional and we further investigate a second order optimization algorithm.

2. Problem Description

We assume that the domain Ω is divided in two subdomains Ω_1 and Ω_2 with $\Omega = \Omega_1 \dot\cup \Gamma \dot\cup \Omega_2$, where $\Gamma := \partial\Omega_1 \cap \partial\Omega_2$ is the so-called interface. Furthermore, Γ is supposed to be an element of a suitable shape space. Then, given $f_\Gamma, \bar u \in H^2(\Omega)$ and appropriate boundary data g, we are considering the following tracking type problem:

$$\begin{aligned} \min_{u_\Gamma,\,\Gamma} \; & ||u_\Gamma - \bar u||_{L^2(\Omega)} + \alpha Reg(u_\Gamma, \Gamma) \\ s.t. \quad & -\mathcal{L}_\Gamma u_\Gamma = f_\Gamma \quad \text{on } \Omega, \\ & u_\Gamma = g \quad \text{on } \mathcal{I}, \end{aligned} \tag{2.1}$$

where $Reg(u_\Gamma, \Gamma)$ is a regularization term and the nonlocal convection-diffusion operator $-\mathcal{L}$ is defined as

$$-\mathcal{L}_\Gamma u_\Gamma(x) = \int_{\mathbb{R}^d} u_\Gamma(x)\gamma_\Gamma(x,y) - u_\Gamma(y)\gamma_\Gamma(y,x)\, dy.$$

As a consequence, the nonlocal boundary $\mathcal{I} := \{y \in \mathbb{R}^d \setminus \Omega : \int_\Omega \gamma_\Gamma(x,y) + \gamma_\Gamma(y,x)\, dx > 0\}$ is usually a domain with nonzero measure and is relying on the so-called kernel $\gamma_\Gamma : \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$. Further, we assume that the kernel γ_Γ and therefore the corresponding nonlocal operator $-\mathcal{L}_\Gamma$ as well as the forcing term f_Γ are dependent on Γ as follows

$$\gamma_\Gamma = \sum_{i,j=1,2} \gamma_{ij}\chi_{\Omega_i\times\Omega_j} + \sum_{i=1,2} \gamma_{i\mathcal{I}}\chi_{(\Omega_i\times\mathcal{I})\cup(\mathcal{I}\times\Omega_i)} \quad \text{and} \quad f_\Gamma = f_1\chi_{\Omega_1} + f_2\chi_{\Omega_2},$$

where $\gamma_{ij}, \gamma_{i\mathcal{I}} : \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ are again nonnegative functions and $f_1, f_2 \in H^2(\Omega)$. Thus, problem (2.1) tries to find a shape Γ such that the corresponding solution u_Γ of the nonlocal Dirichlet problem approximates the given data $\bar u$ as good as possible.
In this talk, we describe how the second shape derivative of the corresponding reduced functional can be computed by applying the averaged adjoint method [2]. As we will see, this approach in combination with the finite element method naturally yields an optimization algorithm, which will also be presented and investigated in numerical experiments.

Acknowledgements

This work has been supported by the German Research Foundation (DFG) within the Research Training Group 2126: 'Algorithmic Optimization'.

References

[1] Stewart A Silling. Reformulation of elasticity theory for discontinuities and long-range forces. *Journal of the Mechanics and Physics of Solids*, 48(1):175–209, 2000.

[2] Kevin Sturm. *On shape optimization with non-linear partial differential equations*. Technische Universitaet Berlin (Germany), 2015.

[3] Jorge L Suzuki, Mamikon Gulian, Mohsen Zayernouri, and Marta D'Elia. Fractional modeling in action: A survey of nonlocal models for subsurface transport, turbulent flows, and anomalous materials. *Journal of Peridynamics and Nonlocal modeling*, 5(3):392–459, 2023.

Shape optimization and uncertainty

Optimization of multiple shapes: from PDE constraints under uncertainty to variational inequality constraints

Caroline Geiersbach[1], **Volker Schulz**[2], **Tim Suchan**[3], **Kathrin Welker**[4]

1. caroline.geiersbach@wias-berlin.de *Weierstraß-Institut, Germany*
2. volker.schulz@uni-trier.de *Universität Trier, Germany*
3. suchan@hsu-hh.de *Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, Germany*
4. kathrin.welker@math.tu-freiberg.de *TU Bergakademie Freiberg, Germany*

Using concepts from differential geometry to formulate (multi-)shape optimization problems as optimization problems on Riemannian manifolds is an idea that has only come up in the current millenium. However, it enables the analysis of these types of problems also from a theoretical point of view and thus allows benefitting from algorithms with established convergence properties. Practically-motivated optimization problems frequently contain some physical behavior of a system, which can often be described by a partial differential equation (PDE) constraint on the optimization problem. If uncertainties have to be accounted for, then even more sophisticated algorithms have to be employed for the optimization of the non-deterministic optimization problem. A further challenge arises if the physical behavior, instead of a PDE constraint, has to be represented by a variational inequality (VI).

This presentation aims to present an algorithm for optimization on Riemannian shape manifolds while being able to incorporate inequality constraints enforced by an Augmented Lagrange approach, which we call the stochastic augmented Lagrangian method in shape spaces. This algorithm is then applied to the problem of minimizing viscous energy dissipation with the Navier-Stokes equations as the PDE constraint. These problems yield non-smooth shapes as an optimal solution, which motivates the usage of a Riemannian shape manifold that contains these types of shapes.

On the topic of VIs, the incorporation of VI constraints on the optimization problem is often addressed by a regularization of the resulting non-differentiability to again obtain a differentiable problem. We demonstrate a method that avoids regularization of the non-differentiability using a similar example as before, however with a VI constraint of the second kind instead of a PDE constraint.

Shape optimization and uncertainty

Quasi-Monte Carlo methods for PDEs on randomly moving domains

Ana Djurdjevac[1], Vesa Kaarnioja[2], Claudia Schillings[3], André-Alexander Zepernick[4]

1. adjurdjevac@zedat.fu-berlin.de, Freie Universität Berlin, Germany
2. vesa.kaarnioja@fu-berlin.de, Freie Universität Berlin, Germany
3. c.schillings@fu-berlin.de, Freie Universität Berlin, Germany
4. a.zepernick@fu-berlin.de, Freie Universität Berlin, Germany

The problem of modelling processes with partial differential equations posed on randomly moving domains arises in various applications like biology or engineering. We will consider the case when such a random domain is generated by the evolution of some initial domain driven by a random velocity field over a fixed time interval.

Since the domain is random it is not straightforward to define notions like the expectation of the PDE solution. A way to deal with this setting is to pull-back the considered equation to a fixed reference domain and to study the solution of the reformulated problem. In order to approximate the expectation of the solution we will use quasi-Monte Carlo methods. For that reason, we will present the needed regularity analysis on the weak formulation of the pull-back of the Poisson equation based on the assumption of a certain parametric regularity of the given velocity field. Our theoretical results will be illustrated by numerical examples, which will also be presented.

THEORETICAL APPROACHES TO MODERN MACHINE LEARNING METHODS

A meanfield control perspective on the training of neurODEs

Benoît Bonnet-Weill

benoit.bonnet-weill@centralesupelec.fr, Université Paris-Saclay, CentraleSupélec, CNRS Laboratoire des Signaux et Systèmes

In this talk, I will discuss some of the results of [4], obtained in collaboration with C. Cipriani, M. Fornasier, and H. Huang, in which we propose a reformulation of a certain class of deep learning problems as optimal control problems in the space of probability measures. The roots of our work trace back to the pioneering articles of Weinan E and co-authors [7, 8]. The incentive for developing such a point of view is twofold. Firstly, the latter builds on the so-called residual block regularisation of neural networks proposed in [10], which has since been known to improve the stability thereof as the number of layers increases. Secondly, embedding residual networks into continuous-time dynamical systems grants access to the broad literature of mathematical control theory, with the help of which one may hope to improve the overall explainability of learning algorithms (see e.g. the works [1, 9, 11, 13, 14] in this direction and the recent survey [12]).

After exposing the conceptual path leading to the reformulation of deep residual learning procedures as mean-field optimal optimal control problems, I shall present a general family of first-order optimality conditions that we derived for this class of problem, and show that the latter can be established by following either of two possible paths. On the one hand, one can derive such optimality extrinsically as a consequence of an abstract Lagrange multiplier rule in the Banach space of Radon measures following [2], in which the subspace of probability measures appears as a convex constraint set. On the other hand, one may also adopt the intrinsic viewpoint developed in [3,5,6], where a mean-field counterpart of the classical Pontryagin Maximum Principle – involving the existence of a state-costate pair solution of an Hamiltonian flow in Wasserstein spaces – is derived. We then show how these optimality conditions may be used to derive quantitative generalisation bounds for quadratically regularised learning problems in the convex regime.

References

[1] A. Agrachev and A. Sarychev. Control on the Manifolds of Mappings with a View to the Deep Learning. *Journal of Dynamical and Control Systems,* 28(4):989–1008, 2022.

[2] L. Ambrosio, M. Fornasier, M. Morandotti, and G. Savaré. Spatially Inhomogeneous Evolutionary Games. *Communication on Pure and Applied Mathematics,* 74(7):1353–1402, 2021.

[3] B. Bonnet. A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems. *ESAIM COCV*, 25(52), 2019.

[4] B. Bonnet, C. Cipriani, M. Fornasier, and H. Huang. A Measure Theoretical Approach to the Mean-field Maximum Principle for Training NeurODEs. *Nonlinear Analysis, Theory, Methods and Applications,* 227:113161, 2023.

[5] B. Bonnet and H. Frankowska. Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces. *Applied Mathematics and Optimization,* 84:1281–1330, 2021.

[6] B. Bonnet and F. Rossi. The Pontryagin Maximum Principle in the Wasserstein Space. *Calculus of Variations and Partial Differential Equations,* 58:11, 2019.

[7] W. E. A Proposal on Machine Learning via Dynamical Systems. *Communications in Mathematics and Statistics,* 1(5):1–11, 2017.

[8] W. E, J. Han, and Q. Li. A Mean-Field Optimal Control Formulation of Deep Learning. *Research in the Mathematical Sciences,* 6(10), 2019.

[9] B. Geshkovski and E. Zuazua. Turnpike in optimal control of PDEs, ResNets, and Beyond. *Acta Numerica,* 31:135–263, 2022.

[10] K. He, X. Zhang, S. Ren, and S. Jian. Deep Residual Learning for Image Recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,* pages 770–778, 2016.

[11] J.-F. Jabir, D. Šiška, and L. Szpruch. Mean-Field Neural ODEs via Relaxed Optimal Control. *arxiv preprint arXiv:1912.05475,* 2021.

[12] D. Ruiz-Balet and E. Zuazua. Neural ODE Control for Classification, Approximation, and Transport. *SIAM Review,* 65(3):735–773, 2023.

[13] A. Scagliotti. Deep Learning Approximation of Diffeomorphisms via Linear-Control Systems. *arxiv preprint arXiv:2110.12393,* 2022.

[14] P. Tabuada and B. Gharesifard. Universal Approximation Power of Deep Residual Neural Networks Through the Lens of Control. *IEEE Transactions on Automatic Control,* 2022.

THEORETICAL APPROACHES TO MODERN MACHINE LEARNING METHODS

On the emergence of clusters in Transformers

Borjan Geshkovski

borjan.geshkovski@inria.fr Inria & Laboratoire Jacques-Louis Lions, Sorbonne Université

1. Introduction

The introduction of Transformers in 2017 marked a milestone in the development of neural network architectures. Central to this is *self-attention*, a novel mechanism which distinguishes Transformers from traditional architectures, and which plays a substantial role in their superior practical performance. We present a mathematical framework for analyzing Transformers based on their interpretation as interacting particle systems, in which time plays the role of layers. Our analysis reveals that clusters emerge in long time, confirming previous empirical findings, and shedding light on the role of the attention mechanism.

2. Main result

As first done in [3], we define an idealized model of the Transformer architecture that consists in viewing the discrete layer indices as a continuous time variable, and which focuses exclusively on two key components of the Transformers architecture: *self-attention* and *layer-normalization*. This results in the dynamics

$$\dot{x}_i(t) = \mathbf{P}^{\perp}_{x_i(t)}\left(\frac{1}{Z_{\beta,i}(t)}\sum_{j=1}^{n} e^{\beta\langle x_i(t),x_j(t)\rangle}x_j(t)\right) \tag{SA}$$

for $i \in [n]$ and $t \geq 0$, where

$$Z_{\beta,i}(t) = \sum_{k=1}^{n} e^{\beta\langle x_i(t),x_k(t)\rangle} \tag{2.1}$$

and $\mathbf{P}^{\perp}_x = I_d - xx^{\top}$ is the orthogonal projector to $\mathsf{T}_x\mathbb{S}^{d-1}$. We prove the following result (see [1, 2]).

Theorem 2.1 *Let $d, n \geq 2$ and $\beta \geq 0$, and suppose that either $d \geq n$, or $\beta \gtrsim_d n^2$, or $\beta \lesssim n^{-1}$. Consider the unique solution $(x_i(\cdot))_{i\in[n]} \in C^0(\mathbb{R}_{\geq 0}; (\mathbb{S}^{d-1})^n)$ to the Cauchy problem for* (SA)*, corresponding to an initial sequence of points $(x_i(0))_{i\in[n]} \in (\mathbb{S}^{d-1})^n$ distributed uniformly at random. Then almost surely there exists $x^* \in \mathbb{S}^{d-1}$ such that*

$$\lim_{t\to+\infty} x_i(t) = x^*$$

for all $i \in [n]$.

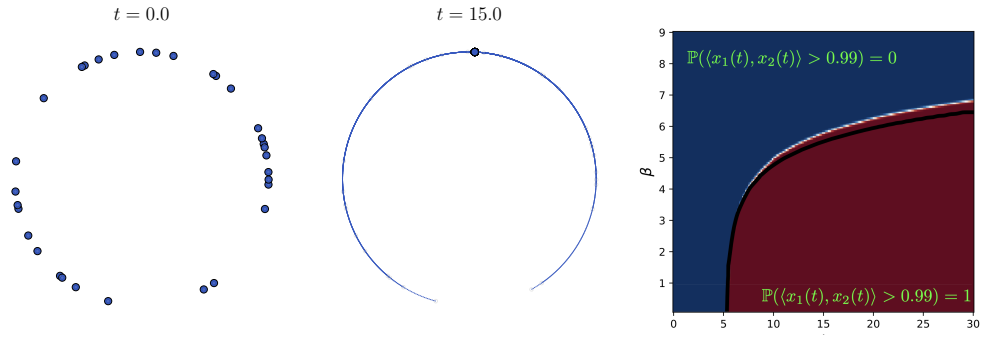


Fig. 1 The evolution of trajectories can be fully described in very large dimension ($d \gg$ poly n), beyond solely long time asymptotics with rates, as seen in the phase diagram above. See [1].

References

[1] Borjan Geshkovski, Cyril Letrouit, Yury Polyanskiy, and Philippe Rigollet. A mathematical perspective on transformers. *arXiv preprint arXiv:2312.10794*, 2023.

[2] Borjan Geshkovski, Cyril Letrouit, Yury Polyanskiy, and Philippe Rigollet. The emergence of clusters in self-attention dynamics. *Advances in Neural Information Processing Systems*, 36, 2024.

[3] Michael E Sander, Pierre Ablin, Mathieu Blondel, and Gabriel Peyré. Sinkformers: Transformers with doubly stochastic attention. In *International Conference on Artificial Intelligence and Statistics*, pages 3515–3530. PMLR, 2022.

THEORETICAL APPROACHES TO MODERN MACHINE LEARNING METHODS

Implicit diffusion: efficient optimization through stochastic sampling

Peter Bartlett[1], **Quentin Berthet**[1], **Mathieu Blondel**[1], **Valentin De Bortoli**[1], **Arnaud Doucet**[1], **Anna Korba**[2], **Felipe Llinares-López**[1], **Pierre Marion**[3], **Courtney Paquette**[1]

1. *Google DeepMind*
2. *ENSAE, CREST, Paris, France*
3. `pierre.marion@epfl.ch`, *EPFL, Lausanne, Switzerland*

Sampling from a target distribution is a ubiquitous task at the heart of various methods in machine learning, optimization, and statistics. Increasingly, sampling algorithms rely on iteratively applying large-scale parameterized functions (e.g. neural networks with trainable weights) to samples, such as in denoising diffusion models [2]. This iterative *sampling operation* implicitly maps a parameter $\theta \in \mathbb{R}^p$ to a distribution $\pi^\star(\theta)$. In this work, our focus is on optimization problems over these implicitly parameterized distributions. For a space of distributions $\mathcal{P}$ (e.g. over $\mathbb{R}^d$), and a function $\mathcal{F} : \mathcal{P} \to \mathbb{R}$, our main problem of interest is

$$\min_{\theta \in \mathbb{R}^p} \ell(\theta) := \min_{\theta \in \mathbb{R}^p} \mathcal{F}(\pi^\star(\theta))$$

This setting encompasses for instance learning parameterized Langevin diffusions, contrastive learning of energy-based models or finetuning denoising diffusion models, as illustrated by Figure **??**.

Applying first-order optimizers to this problem raises the challenge of computing gradients of functions of the target distribution with respect to the parameter: we have to *differentiate through a sampling operation*, where the link between θ and $\pi^\star(\theta)$ can be *implicit*. To this aim, we propose to exploit the perspective of *sampling as optimization* over the space of probability distributions $\mathcal{P}$ [3].

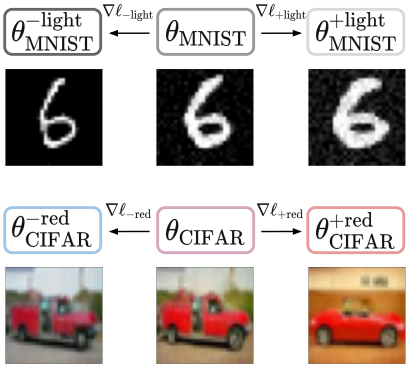


Fig. 1 Optimizing through sampling with **Implicit Diffusion** to finetune denoising diffusion models. The reward is the average brightness for MNIST and the red channel average for CIFAR-10.

This allows us to draw a link between optimization through stochastic sampling and *bilevel optimization* [1]. These motivating similarities, while useful, are not limiting, and we also consider settings where the sampling iterations are not readily interpretable as an optimization algorithm.

Main Contributions. In this work, we introduce the algorithm of **Implicit Diffusion**, an effective and principled technique for optimizing through a sampling operation. Our main contributions are the following:

- We present a general framework describing parameterized sampling algorithms, and introduce Implicit Diffusion optimization, a **single-loop** optimization algorithm to optimize through sampling.
- We provide **theoretical guarantees** under various conditions, in the continuous and discrete time settings.
- We showcase its performance in experimental settings.

References

[1] Mathieu Dagréou, Pierre Ablin, Samuel Vaiter, and Thomas Moreau. A framework for bilevel optimization that enables stochastic and global variance reduction algorithms. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 26698–26710. Curran Associates, Inc., 2022.

[2] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020.

[3] Anna Korba and Adil Salim. Sampling as first-order optimization over a space of probability measures, 2022. Tutorial at ICML 2022. Accessible at `https://akorba.github.io/resources/Baltimore_July2022_ICMLtutorial.pdf`, consulted on 01/30/2024.

THEORETICAL APPROACHES TO MODERN MACHINE LEARNING METHODS

Residual Networks: from infinitely deep neural ODEs to Transformers in actions

Michael E. Sander
michael.sander@ens.fr Ecole Normale Supérieure and CNRS, France

1. Introduction

From convolutional ResNets to Transformers, residual connections are ubiquitous in state-of-the-art deep learning models. The continuous depth analogues of residual networks, neural ODEs, have been widely adopted, but the connection between the discrete and continuous models still lacks a solid mathematical foundation. We will show that for a formal correspondence between residual networks and neural ODEs to hold, the residual functions must be smooth with depth, and we will present an implicit regularization result of deep residual networks towards neural ODEs. We will then present two applications of this analogy to the design and study of new architectures. First, we will introduce a drop-in replacement for any residual network that can be trained with the same accuracy, but with much less memory. Second, by viewing the attention mechanism as an interacting particle system, where the particles are the tokens, we will study the impact of attention map normalization on the Transformer model.

2. Residual Neural Networks and Neural ODEs

A ResNet [2] of depth N iterates, starting from $x_0 \in \mathbb{R}^d$

$$x_{n+1} = x_n + \frac{1}{N} f(x_n, \theta_n^N), \tag{2.1}$$

and outputs a deep representation $x_N \in \mathbb{R}^d$ where f is a neural network called residual function, which can typically be a convolutional neural network [2] or, more recently, a Transformer layer [7]. On the other hand, a Neural ODE [1] uses a neural network $\phi_\Theta(x, s)$, that takes time s into account, to parameterise a vector field in a differential equation, as follows,

$$\frac{dx}{ds} = \phi_\Theta(x(s), s) \quad \text{with} \quad x(0) = x_0. \tag{2.2}$$

It outputs a deep representation $x(1) \in \mathbb{R}^d$, the solution of Eq.(2.2). Neural ODEs enable learning without storing the trajectory $x(s)$, significantly reducing the memory needed for backpropagation, which is frequently the main bottleneck when training deep models.
From a theoretical perspective, we will show that for a formal correspondence between (2.1) and (2.2) to hold as $N \to +\infty$, the $f(x, \theta)$ must be smooth with depth [4]. Moreover, we will show that, under proper assumptions on f, such a smoothness property holds during training if it is imposed at initialization of gradient flow. This corresponds to an implicit regularization result of ResNets towards neural ODEs [3]. From a practical viewpoint, we will use the analogy between (2.1) and (2.2) to introduce a drop-in replacement for any ResNet that can be trained with the same accuracy, but with much less memory [5], based on momentum updates. Second, we will view the attention mechanism as an interacting particle system over tokens, in order to study the impact of attention map normalization on the Transformer model using partial differential equations [6].

Acknowledgements. Joint works with Pierre Ablin, Gérard Biau, Mathieu Blondel, Pierre Marion, Gabriel Peyré and YuHan Wu.

References

[1] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems,* 31, 2018.

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition,* pages 770–778, 2016.

[3] Pierre Marion, Yu-Han Wu, Michael E Sander, and Gérard Biau. Implicit regularization of deep residual networks towards neural odes. *arXiv preprint arXiv:2309.01213,* 2023.

[4] Michael Sander, Pierre Ablin, and Gabriel Peyré. Do residual neural networks discretize neural ordinary differential equations? *Advances in Neural Information Processing Systems,* 35:36520–36532, 2022.

[5] Michael E Sander, Pierre Ablin, Mathieu Blondel, and Gabriel Peyré. Momentum residual neural networks. *International Conference on Machine Learning,* pages 9276–9287. PMLR, 2021.

[6] Michael E Sander, Pierre Ablin, Mathieu Blondel, and Gabriel Peyré. Sinkformers: Transformers with doubly stochastic attention. *International Conference on Artificial Intelligence and Statistics,* pages 3515–3530. PMLR, 2022.

[7] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems,* 30, 2017.

UNCERTAIN OPTIMIZATION AND RELATED TOPICS

Application of random ordinary differential equation in smart thermostat thermal models for residential buildings

Jorge Arias[1], **Miguel Sama**[2]

1. jarias310@alumno.uned.es UNED, Valor, Desarrollo e Innovación, S.L., Gijón, Spain
2. msama@ind.uned.es Universidad Nacional de Educación a Distancia, UNED, Madrid, Spain

1. Description

This work aims to study the thermal behavior of residential buildings by using the data provided by smart thermostats and weather forecast data. For this, we consider an equivalent ODE circuit model depending on four parameters related to the heater power, the solar energy, heat capacity, and the thermal resistance of the building. We consider a random version of the model to overcome natural model uncertainty. More specifically we consider the following Ordinary Differential Equation with random data wich was introduced in [1]

$$\frac{dT}{dt}(t,\omega) + \frac{1}{R(\omega)C(\omega)}T(t,\omega) = \frac{1}{R(\omega)C(\omega)}T^e(t) + \frac{1}{C(\omega)}\left[P(\omega)dU(t) + A(\omega)I^c(t)\right] \text{ a.e } \omega \in \Omega$$

Here:

- $T(t,\omega)$: indoor temperature.
- $T^e(t)$: exterior temperature.
- $I^c(t)$: corrected solar radiation.
- $dU(t)$: heater usage time fraction.

While the random parameters to be identified have the following physical intepretation

- $R(\omega)$: thermal resistance.
- $C(\omega)$: heat capacity.
- $P(\omega)$: effective boiler power.
- $A(\omega)$ factor of solar radiation.

Based on the data available we show how to solve effectively this model.

Acknowledgements

This research is partially supported by Project PID2020-112491GB-I00/AEI/10.13039/501100011033 through the Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación, Spain.

References

[1] J. Arias, A.A. Khan, J. Rodriguez-Uría, and M. Sama (2022). Analysis of smart thermostat thermal models for residential building. Applied Mathematical Modelling, 110, 241-261.

Uncertain optimization and related topics

Optimal portfolios with asymmetric information and the meaning of noise for risk-averse traders

Carlos Escudero
cescudero@mat.uned.es Universidad Nacional de Educación a Distancia (UNED), Spain

We will study the problem of portfolio optimization in a financial market in which traders with asymmetric information are present. We will summarize some of our results on the mathematical formalization of such an economic problem [1–4]. Our main goal is to compare the utilization of different anticipating stochastic calculi in this context. Precisely, we will employ the Russo-Vallois forward and Skorokhod stochastic integrals to interpret the noise of the stochastic differential equations that model the asymmetrically informed investors. Theoretical analyses and illustrative numerical examples showcase the critical role that risk aversion has on the wealth dynamics. In particular, we will see that the results that follow in the case of risk neutrality become invalid if we assume a convex utility modeling risk aversion. This latter case produces other unexpected outcomes such as, for instance, the fact that less-informed-traders could surpass more-informed-traders in certain market conditions. Our findings signal the intricate interplay between the anticipating interpretation of noise and nonlinear utilities, which might produce counter-intuitive results from the financial viewpoint.

Acknowledgements

This work has received funding from the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 777822. It has been also partially supported by the Government of Spain (Ministerio de Ciencia e Innovación) and the European Union through Projects PID2021-125871NB-I00, CPP2021-008644/AEI/10.13039/501100011033/Unión Europea NextGenerationEU/PRTR, and TED2021-131844B-I00/AEI/10.13039/501100011033/Unión Europea NextGenerationEU/PRTR.

References

[1] J. C. Bastons, C. Escudero, (2018). A triple comparison between anticipating stochastic integrals in financial modeling. Communications on Stochastic Analysis, 12, 73-87.

[2] M. Elizalde, C. Escudero, (2022). Chances for the honest in honest versus insider trading. SIAM Journal on Financial Mathematics, 13, SC39-SC52.

[3] M. Elizalde, C. Escudero, T. Ichiba, (2022). Optimal investment with insider information using Skorokhod & Russo-Vallois integration. arXiv:2211.07471.

[4] C. Escudero, (2018). A simple comparison between Skorokhod & Russo-Vallois integration for insider trading. Stochastic Analysis and Applications, 36, 485-494.

Uncertain optimization and related topics

Uncertainty in the steelmaking process: impact and challenges

Alejandro Fernández[1], **Carlos A. Moreno-Camacho**[2]

1. alejandro.fernandez-alonso@arcelormittal.com ArcelorMittal Global R&D, Spain
2. carlos-alberto.moreno@arcelormittal.com ArcelorMittal Global R&D, Spain

In response to the global efforts towards cleaner practices, ArcelorMittal, a leading steel manufacturer operating in over 60 countries, is adapting its production strategies. One pivotal move in this transition is the adoption of Electric Arc Furnace (EAF) technology, which emphasizes using recycled scrap metal instead of the more polluting Blast Oxygen Furnace route. This shift is primarily driven by the global decarbonization strategy, which is promoting an increased demand for scrap metal across the world. This trend is reshaping market dynamics and significantly impacting production costs.

Unlike the traditional method of using hot metal, incorporating scrap metal introduces complexity to the production process due to variations in chemical composition, including residual elements like copper. Managing these variations carefully is crucial to maintain the quality standards of the liquid steel produced in each heat.

Decision support models in steelmaking commonly overlook variability in compound content, relying solely on nominal values. However, neglecting this variability can lead to two main issues. Firstly, inaccurate estimates of compound content may result in unmet production requirements. Secondly, when utilizing scrap, overly conservative solutions may prioritize costly high-quality input materials, leading to excessive costs. In this context, stochastic optimization techniques appear as a valuable tool to address with the mentioned situation. To address the blending problem, stochastic optimization techniques such as Robust Optimization, Probabilistic Constraints, and Stochastic Programming can be valuable tools, each with its own pros and cons worth exploring.

Incorporating uncertainty into decision-making can provide valuable insights aligned with risk profiles, thereby minimizing production costs and supporting the scrap purchasing process. The introduction of probabilistic scrap characterization represents a significant advancement in managing compositional uncertainty within the steelmaking process.

Uncertain optimization and related topics

Nondominated solutions of uncertain optimization problems

César Gutiérrez
cesargv@uva.es Universidad de Valladolid, Spain

This talk concerns an uncertain unconstrained optimization problem with countably many scenarios, which is dealt with the robust approach. Specifically, a new class of solutions is introduced that generalizes the so-called pointwise efficient solutions (see [2–4] and the references therein). In addition, necessary and sufficient optimality conditions are derived in convex problems. These characterizations extend and clarify some recent results in [1].

Acknowledgements

This research was partially supported by the Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación, Spain, through project PID2020-112491GB-I00 / AEI / 10.13039/501100011033.

References

[1] Alexander Engau, Devon Sigler. Pareto solutions in multicriteria optimization under uncertainty. *European J. Oper. Res.*, 281(2):357–368, 2020.

[2] Elisabeth Köbis. On robust optimization: relations between scalar robust optimization and unconstrained multicriteria optimization. *J. Optim. Theory Appl.*, 167:969–984, 2015.

[3] Fouad Ben Abdelaziz, Pascal Lang, Raymond Nadeau. Pointwise efficiency in multiobjective stochastic linear programming. *J. Opl Res. Soc.*, 45(11):1324–1334, 1994.

[4] Jonas Ide, Anita Schöbel. Robustness for uncertain multi-objective optimization: a survey and analysis of different concepts. *OR Spectrum*, 38:235–271, 2016.

UNCERTAIN OPTIMIZATION AND RELATED TOPICS

Necessary and sufficient conditions for differentiability of interval-valued functions: applications

Tiago M. Da Costa[1], **Beatriz Hernández-Jiménez**[2], **Rafaela Osuna-Gómez**[3], **Antonio Pascual-Acosta**[4]

1. grafunjo@yahoo.es *Universidad de Tarapacá, Chile*
2. mbherjim@upo.es *Universidad Pablo de Olavide, Sevilla, Spain*
3. rafaela@us.es *Universidad de Sevilla, Sevilla, Spain*
4. antoniop@us.es *Universidad de Sevilla, Sevilla, Spain*

After a review on differentiability notions for interval-valued functions you can find in the literature, we present necessary and sufficient conditions for generalized Hukuhara differentiability of interval-valued functions and counterexamples of some equivalences previously presented in the literature, for which important results are based on.

Differential Calculus is a branch of Mathematics that allows us to solve problems where the change of variables can be modeled in a numerical continuum to determine, from it, the variation of these elements in specific moment or interval. The Optimization Theory is a basic part of Applied Mathematics, and the development of differential calculus has enabled powerful mathematical tools for this area.

The Differential Calculus has provided essential mathematical tools to areas as physics, biology, engineering, economics, among others. In particular, since Fermat and Lagrange's work, Differential Calculus has played a leading role in the Optimization Theory. In order to optimize a differentiable function or to solve an optimization problem with constraints, derivative is crucial in both situations and numerical algorithms for computing approximately optimal solutions because the main iterative optimization methods are based on the evaluation of hessian matrices or gradients.

Under the hypothesis that observations and estimates in the real world are incomplete to accurately represent the actual data, the Interval Analysis was introduced by Moore with the aim of managing the imprecision or lack of accurate information that appears on many mathematical models or computational of some real-world deterministic phenomena.

Moreover, interval differentiability and its application in fuzzy environment is an active research area as you can see in literature. But it has not been developed without problems, in order to define correctly the operations between intervals and to establish the appropriate differentiability concept due to the no linearity of the space of intervals. Therefore, it is of interest to establish the definitions and equivalences correctly, such that they allow a successful development of the theory and applications based on them.

So, in this talk we present necessary and sufficient conditions for generalized Hukuhara differentiability of interval-valued functions and counterexamples of some equivalences previously presented in the literature, for which important results are based on.

Acknowledgements

The research has been supported by MICINN through grant PID2021-123051NB-100 and by project FIUS20/0117.

References

[1] Moore, R.E.: Interval Analysis. Englewood Cliffs NJ. Prince-Hall. (1966).

[2] Ghosh, D.: Newton method to obtain efficient solutions of the optimization problems with interval-valued objective functions. J Appl Math Comput. **53**,709–731 (2017).

[3] Ghosh, D., Singh, A., Shukla, K.K,, Manchanda, K.: Extended Karush-Kuhn-Tucker condition for constrained interval optimization problems and its application in support vector machines. Inf Sci. **504**: 276–292 (2019).

[4] Osuna-Gómez, R., Da Costa, T.M, Chalco-Cano, Y., Hernández-Jiménez, B.: Quasilinear approximation for interval-valued functions via generalized Hukuhara differentiability. Comp Appl Math. **41**, 149 (2022).

[5] Stefanini, L.: A generalization of Hukuhara difference for interval and fuzzy arithmetic. Soft Methods for Handling Variability and Imprecision, Series on Advances in Soft Computing: Springer (2008).

[6] Stefanini, L., Bede, B.: Generalized Hukuhara differentiability of interval-valued functions and interval differential equations. Nonlinear Anal. **71**(34),1311–1328 (2009).

Uncertain optimization and related topics

On stability in set-valued dynamics

Elvira Hernández
ehernandez@ind.uned.es UNED, Spain

In this talk we study several notions of stability of set-valued discrete-time dynamical systems. More exactly, we develop a general framework in order to unify several results existing in the literature and we generalize dynamical results of the unidimensional case. We start by reviewing the preliminary results. Then, we formulate appropriate notions of set dynamical systems as attractor, stability and invariant sets. To obtain stability results we prove some properties of the Lyapunov mappings.

Acknowledgements

The author is supported by the Ministerio de Ciencia e Innovación of Spain under the project PID2021-122442NB-I00.

Uncertain optimization and related topics

Stability of interval optimization problems

Rubén López

rlopezm@academicos.uta.cl Universidad de Tarapacá, Chile

We study the stability of interval optimization problems; i.e., optimization problems with interval-valued objective functions. We focus on set-type solutions that are defined by means of the Kulisch-Miranker order between intervals. To this end, we use a suitable notion of variational convergence for vector functions. We provide geometric and metric characterizations of this convergence notion. We describe the behavior of level, colevel and solution sets under perturbations of the data of the problem. We show that the coercivity properties and coercive existence conditions for these problems are preserved locally within certain classes of functions. We compare the variational convergence with other convergence notions from the literature. Finally, we study the behavior of operations with interval functions under perturbations.

Acknowledgements

The author was partially supported by project Fondecyt 1220687 (ANID, Gobierno de Chile).

References

[1] Iván Aguirre-Cipe, Rubén López, Exequiel Mallea-Zepeda, Lautaro Vásquez. A study of interval optimization problems. *Optim. Lett.*, 15:859–877, 2021.

[2] Daishi Kuroiwa. Existence theorems of set optimization with set-valued maps. *J. Inform. Optim. Sci.*, 24:73-8, 2003.

[3] Rubén López. Global stability of interval optimization problems. *Optimization*, 69:2431-2451, 2020.

[4] Rafaela Osuna-Gómez, Yurilev Chalco-Cano, Beatríz Hernández-Jiménez, Gabriel Ruiz-Garzón. Optimality conditions for generalized differentiable interval-valued functions. *Inform. Sci.*, 321:136-146, 2015.

Uncertain optimization and related topics

Scalarization and robustification in vector games: a non componentwise approach

Elisa Caprari[1], **Lorenzo Cerboni Baiardi**[2], **Elena Molho**[3]

1. elisa.caprari@unipv.it Università di Pavia, Italy
2. lorenzo.cerboni@unibo.it Università di Bologna, Italy
3. elena.molho@unipv.it Università di Pavia, Italy

The notion of robustness has been extended to uncertain multicriteria games in [4] where, at each strategy profile, an uncertain parametric vector payoff is replaced by a vector collecting the componentwise worst cases. By the componentwise approach to robustness, possible interdependences among vector components are neglected. As a consequence, the componentwise worst case may be unachievable as uncertain parameters vary and componentwise robustness may induce an unduly pessimistic behavior of decision makers. A further extension of the notion of robustness to uncertain vector games is presented in [3]. The authors obtain a deterministic set-valued robust counterpart of an uncertain vector game by replacing uncertain parametric payoffs by set-valued payoffs where all the parameters' occurrences are considered. At a given strategy profile, the set-valued payoff of player i is the set that includes all possible realizations of i's payoff as uncertain parameters vary in their domains. As a consequence, the drawbacks of the componentwise approach are addressed, even in the special case of multicriteria games. We extend to noncooperative game theory the abstract scalarization scheme presented in [2] for the special case of vector optimization problems and we provide necessary and sufficient robust equilibrium conditions through scalarization. Interestingly, we highlight that any scalarization scheme that can be applied to the set-valued robust counterpart of a vector game providing necessary and sufficient robust equilibrium conditions is encompassed within the axiomatic approach introduced in [2]. In order to avoid the use of set-valued analysis, an alternative approach to robustify a vector game can be introduced. One can scalarize the original parametric vector-valued payoff functions of the original game and subsequently consider the associated parametric scalar game where the uncertainty on parameters can be treated by considering a robust counterpart built according to the classical approach introduced in [1]. This approach, more suitable for applications, where scalarizations are often implicitely considered in the model, does not consider the coherence with the robustness notion defined on the original uncertain vector game. We investigate the commutativity of scalarization and robustification of uncertain vector games obtaining equivalent robust equilibrium conditions that rely on the application of the standard notion of robust equilibrium on the parametric scalar game obtained by appropriately scalarizing the original vector parametric game.

References

[1] Aghassi, Michele and Bertsimas, Dimitris. Robust game theory. *Mathematical Programming*,107:231-273 2006.

[2] Caprari, Elisa and Cerboni Baiardi, Lorenzo and Molho, Elena. Scalarization and robustness in uncertain vector optimization problems: a non componentwise approach. *Journal of Global Optimization*, 84:295-320 2022.

[3] Crespi, Giovanni P and Kuroiwa, Daishi and Rocca, Matteo. Robust Nash equilibria in vector-valued games with uncertainty. *Annals of Operations Research*, 289:185-193 2020.

[4] Yu, H and Liu, HM. Robust multiple objective game theory. *Journal of Optimization Theory and Applications*, 159:272-280 2013.

UNCERTAIN OPTIMIZATION AND RELATED TOPICS

Uncertainty quantification in partial differential equations with random data

Miguel Sama

msama@ind.uned.es Universidad Nacional de Educación a Distancia, UNED, Spain

1. Description

Partial differential equations (PDEs) involving random parameters are useful models in science since they allow to model natural uncertainties. We will discuss the nonlinear inverse problem of estimating stochastic parameters in elliptic partial differential equations with random data. More specifically, in this talk, we focus on the following specific model studied in [1].

Let $(\Omega, \mathbb{F}, \mathbb{P})$ be a complete probability space, where Ω is a nonempty set of elementary events, $\mathbb{F}$ is a σ-algebra of subsets of Ω, and $\mathbb{P} : \mathbb{F} \to [0,1]$ is a probability measure. Let $D \subset \mathbb{R}^2$ be a bounded domain and let ∂D be its sufficiently smooth boundary $\partial D = \Gamma_1 \cup \Gamma_2$ with $\Gamma_1 \cap \Gamma_2 = \emptyset$. The following stochastic linear elasticity system models the displacement in the elastic body induced by the stochastic load f and the stochastic boundary traction h:

$$-\nabla \cdot [2\mu(\omega,x)\epsilon_u + \lambda(\omega,x)\mathrm{tr}\,(\epsilon_u)\, I] = f(\omega,x), \text{ in } D, \tag{1.1a}$$

$$u(\omega,x) = 0, \text{ on } \Gamma_1, \tag{1.1b}$$

$$[2\mu(\omega,x)\epsilon_u + \lambda(\omega,x)\mathrm{tr}\,(\epsilon_u)\, I]\, n = h(\omega,x), \text{ on } \Gamma_2. \tag{1.1c}$$

Here the random field $u(\omega,x)$ is the displacement vector, I is the identity map, n is the outward-pointing unit normal to ∂D, and the random fields $\mu : \Omega \times D \to \mathbb{R}$ and $\lambda : \Omega \times D \to \mathbb{R}$ are the Lamé parameters. Denoting by ∇u, the gradient of the vector-valued random field u, the linearized strain tensor

$$\epsilon_u := \frac{1}{2}\left(\nabla u(\omega,x) + \nabla u(\omega,x)^\top\right),$$

is the local deformation of the elastic body, whereas $\mathrm{tr}(\epsilon_u)$ is the trace of ϵ_u. We note that the derivatives in the elasticity system are with respect to x. The stress tensor σ measures the response of the elastic object described by the strain.

In the context of (1.1), our focus will be on the stochastic inverse problem of estimating the stochastic Lamé parameters $\mu(\omega,x)$ and $\lambda(\omega,x)$ from some statistical information concerning the corresponding solution u.

Acknowledgements

This research is partially supported by Project PID2020-112491GB-I00/AEI/10.13039/501100011033 through the Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación, Spain.

References

[1] M. Dambrine, A.A. Khan, M. Sama, H.J. Starkloff, (2023). Stochastic Elliptic Inverse Problems. Solvability, Convergence Rates, Discretization, and Applications. Journal of Convex Analysis, 30(3), 851-885.

Uncertain optimization and related topics

On some stochastic aspects of stochastic elliptic inverse problems

Hans-Jörg Starkloff

hjstark@math.tu-freiberg.de Technische Universität Bergakademie Freiberg, Germany

Stochastic elliptic problems arise mainly by substituting deterministic parameters in elliptic problems, as for example coefficients, forcing terms or boundary conditions, by certain random parameters. Then one issue in the consideration of random equations is the measurability of desired solutions. Based on the fact that there exist different measurability concepts it is important to use the appropriate measurability concept for each problem. Hereby the Borel, weak and strong measurability concepts are of main interest.

In the talk these measurability concepts are presented and some of the relations between them are discussed. This is important, because in elliptic problems also non-separable Banach spaces play a certain role and in these spaces the measurability concepts mentioned above do not necessarily coincide. Based on these findings measurability properties of solutions of elliptic problems are investigated.

Furthermore it will be shown exemplarily, which stochastic elliptic inverse problems can be treated as abstract elliptic inverse problems and which such stochastic inverse problems require a specific stochastic investigation.

Uncertain optimization and related topics

A review on the radius of robust feasibility of uncertain mathematical programs

M.A. Goberna[1], **V. Jeyakumar**[2], **G. Li**[2], **José Vicente-Pérez**[3]

1. *Universidad de Alicante, Spain*
2. *University of New South Wales, Australia*
3. *jose.vicente@ua.es Universidad de Alicante, Spain*

The radius of robust feasibility provides a numerical value for the largest possible uncertainty set that guarantees feasibility of a robust counterpart of a mathematical program with uncertain constraints. The objective of this review of the state-of-the-art in this field is to present this useful tool of robust optimization to its potential users and to avoid undesirable overlapping of research works on the topic as those we have recently detected. In this talk we overview the existing literature on the radius of robust feasibility in continuous and mixed-integer linearly constrained programs, linearly constrained semi-infinite programs, convexly constrained programs, and conic linearly constrained programs. We also analyze the connection between the radius of robust feasibility and the distance to ill-posedness for different types of uncertain mathematical programs.

References

[1] M.A. Goberna, V. Jeyakumar, G. Li, J. Vicente-Pérez. The radius of robust feasibility of uncertain mathematical programs: A survey and recent developments. *European J. Oper. Res.* 296:749–763, 2022.

Index

Author

Author Index

D

E

F

G

H

I

J

K

L

M

N

O

P

R

S

T

U

V

W

Y

Z