Book of abstracts

PGMO DAYS 2023

November 28th and 29th, 2023

EDF Labs Paris-Saclay

Preface

This volume contains the extended abstracts of the talks presented at the conference PGMODAYS 2023 held on November 28th – 29th, 2023 at EDF Labs Paris-Saclay.

We especially acknowledge the support of EDF and FMJH. We thank CNRS, Institut polytechnique de Paris, Université Paris-Saclay and Inria. We also thank ROADEF, SMAI-MODE, SMF, GdR MOA, GdR RO, and GdR Jeux. We are grateful to the plenary speakers and to the organizers of the invited sessions.

November 15, 2023
Palaiseau

The organizers of the PGMO days

Organization Committee

Scientific Program Committee

Wim van Ackooij	EDF
Gregoire Allaire	Ecole Polytechnique
Claudia D'Ambrosio	CNRS LIX, Ecole Polytechnique
Christian Artigues	LAAS, CNRS
Walid Ben Ameur	Télécom SudParis
Frederic Bonnans	INRIA
Claire Boyer	Sorbonne Université
Luce Brotcorne	INRIA
Jean-Baptiste Caillau	Univ. Côte d'Azur and CNRS/Inria
Antonin Chambolle	CEREMADE, CNRS and Université Paris-Dauphine PSL, Paris
Johanne Cohen	LRI, CNRS and Université Paris-Sud
Michel De Lara	CERMICS, École des Ponts ParisTech
Sourour Elloumi	UMA, Ensta Paris and CEDRIC, Cnam
Oliver Fercoq	Telecom Paris
Stéphane Gaubert	INRIA and CMAP, Ecole Polytechnique
Jerome Malick	CNRS, Lab. J. Kuntzmann
Vianney Perchet	ENSAE Paris *(chair)*
Patrice Perny	Sorbonne Université
Alain Quilliot	UCA, CNRS
Frédéric Roupin	Université Paris 13
Joseph Salmon	Université de Montpellier

Organizers of Invited Sessions

Émilie Chouzenoux and Antonin Chambolle: *First Order Optimization: Acceleration and Convergence*
Julien Grand-Clément: *Robust Optimization and Applications*
Andrea Simonetto: *Quantum computing and combinatorial optimisation*
Mathias Staudigl: *Hierarchical Optimization in Energy Systems*
Luca Nenna: *Optimal Transport and Applications*
Sorin-Mihai Grad: *New roads to optimality*
Matías R. Bender and Elias Tsigaridas: *Nonlinear Algebra and its Application*
Michel De Lara: *Algorithms in Generalized Convexity. Applications to Machine Learning and Sparse Optimization*
Hanene Mohamed: *Optimization in car-sharing systems : Probabilistic Models and Analysis*
Guilherme Mazanti and Laurent Pfeiffer: *Mean field games and control*
Eric Gourdin, David Coudert, Sébastien Martin and Dritan Nace: *Optimization in Telecom Networks.*

Table of Contents

Integration of climate variability and climate change in renewable energy planning

Philippe Drobinski

CNRS, École Polytechnique, Institut Pierre Simon Laplace, France

The trajectory outlined in the Paris Agreement to keep global warming below 2° dictates not only the timing but also the speed at which the transformation of our energy system must take place to decarbonize energy production. Complying with the Paris Agreement requires reducing the carbon content of energy by about 75% and therefore making a rapid transition from fossil production to production based on low-carbon technologies. Among these technologies are those based on renewable energies. The variability of the climate itself induces a fluctuating or even an intermittent production of variable renewable energy (solar, wind, marine), challenging the balance of the electricity grid. In this context, to speak of energy transition is to face the problem of increasing the pene- tration of low-carbon energy production while limiting the variability so as to ensure the socio-technical feasibility and economic viability. The problem is not simple, and the delicate balance between urgency (drastically reducing emissions) and utopia (choosing a strategy for low carbon energies and analyzing opportunities and obstacles) needs to be clearly defined.

Submodular Functions on Modular Lattices

Michel Goemans

Massachusetts Institute of Technology, USA

Submodular set functions are central to discrete optimization and are ubiquitous in many areas, including machine learning. To some extent, submodular functions can be viewed as the analog of convex functions in the continuous setting. In this talk, I will first introduce and present basic properties of submodular functions and algorithms for associated optimization problems. I will then discuss extensions to the much less studied setting of submodular functions on lattices, especially modular lattices. This perspective will allow us to derive a surprising (and easy) result, connecting two problems whose solutions are seemingly unrelated.

Randomization techniques for solving large scale linear algebra problems

Laura Grigori

École Polytechnique Fédérale de Lausanne and Paul Scherrer Institut, Switzerland

In this talk we discuss recent progress in using randomization for solving large scale linear problems. We present a randomized version of the Gram-Schmidt process for orthogonalizing a set of vectors and its usage in the Arnoldi iteration. This leads to introducing new Krylov subspace methods for solving large scale linear systems of equations and eigenvalue problems. The new methods retain the numerical stability of classic Krylov methods while reducing communication and being more efficient on modern massively parallel computers.

Ease-controlled Random Reshuffling Gradient Algorithms for nonconvex finite sum optimization

Laura Palagi

Department of Computer, Control and Management Engineering
Università di Roma La Sapienza, Italy, laura.palagi@uniroma1.it

We consider minimizing the sum of a large number of smooth and non-convex functions, which is a typical problem encountered when training deep neural networks on huge datasets. We propose ease-controlled modifications of the traditional online gradient schemes, either incremental (IG) or random reshuffling (RR) gradient methods, which converges to stationary points under weak and basic assumptions. Indeed, besides the compactness of level sets, we require the lonely assumption of Lipschitz continuity of the gradients of the component functions. The algorithmic schemes control the IG/RR iteration by using a watchdog rule and a derivative-free line search that activates only sporadically to guarantee convergence. The schemes also allow controlling the updating of the learning rate used in the main IG/RR iteration, avoiding the use of preset rules, thus overcoming another tricky aspect in implementing online methods. We also propose a variant to further reduce the need for the objective function computation. We performed an extensive computational test using different Deep Networks Architectures and a benchmark of large datasets of varying sizes and we compare performance to state-of-the-art optimization methods for ML. The tests show that the schemes are efficient, in particular when dealing with ultra-deep architectures.

Joint paper with Corrado Coppola, corrado.coppola@uniroma1.it Giampaolo Liuzzi, giampaolo.liuzzi@uniroma1.it Ruggiero Seccia, ruggiero.seccia @uniroma1.it

PGMO 2023 PhD Prizes

The Gaspard Monge Program for Optimization, Operations Research, and their Interactions with Data Sciences, awards every year, under the scientific patronage of ROADEF and SMAI-MODE, two PhD prizes (exaequo). All the fields in Optimization, and Operations Research, including their Interfaces, are eligible. The applicants must have been defended their PhD in France, during the previous civil year.

The two 2023 PGMO PhD prizes were awarded to

- *Ngoc Hoang Anh Mai*, for his PhD at LAAS CNRS, under the supervision of Jean-Bernard Lasserre and Victor Magron, on "Practical polynomial optimization through positivity certificates with and without denominators"".
- *Rémi Robin*, for his PhD at Sorbonne Université, under the supervision of Ugo Boscain and Mario Sigalotti, at Laboratoire Jacques-Louis Lions, in the CAGE Inria team, on "Contrôle et optimisation de systèmes physiques : applications à la mécanique quantique et au confinement magnétique dans les stellarators".

The 2023 PhD prize committee was chaired by Pierre Fouilhoux (LIPN). It was composed of the following researchers:

- Jean-Michel Coron (LJLL)
- Gersende Fort (IMT)
- Anna Korba (CREST, ENSAE)
- Olga Battaia (Kedge Business School)
- Adam Ouorou (Orange)
- Mathurin Massias (INRIA)
- Francisco Silva (Université de Limoges)
- Bruno Zillioto (CNRS)

Maximum of Reachable Values of Discrete-Time Systems

Assalé Adjé[1]

[1]Université de Perpignan Via Domitia, France
assale.adje@univ-perp.fr

Keywords: Discrete-time system, Sequences suprema, Lyapunov functions, Semidefinite programming

In this presentation, we are interested in finding a reachable value of a discrete-time dynamical system that maximizes a real valued function. This type of maximization problems arises in the analyze of the robustness or the performance of the system; for inverse problems or even in formal verification of systems. The optimal value in this context can be viewed as the value which penalizes the most the system with respect to some criteria.

More precisely, the problem is formulated from an objective real-valued continuous function and a discrete-time dynamical system. This latter is characterized by, the dynamics which is a continuous self map on a finite dimensional vector space and initial conditions set which is compact. The constraints set of the problem is the whole set of reachable values which is a set of terms of recurrence sequences starting from an element inside the initial conditions set. In this context, the optimal solution is composed of an initial condition and an integer which represents the rank for which the term is maximal. The compactness of the initial conditions set and the continuity assumptions constitute a comfortable framework for the optimal initial condition. The difficulty of the problem lies on the characterization of the optimal integer solution.

Then, in this presentation, we propose characterizations of the optimal integer solution based on homeomorphisms of geometric sequences. When the system is Lyapunov stable and a Lyapunov function can be computed such a homeomorphism can be computed. This also leads to a computable formula of an upper bound of this optimal integer solution. We illustrate our approach on discrete-time stable affine dynamical systems for which semidefinite programming can be employed to compute such a bound [1]. We also aboard the problem of the computation of the minimal upper-bound and its formulation as a nonlinear semidefinite program.

References

[1] A. Adjé, Quadratic Maximization of Reachable Values of Stable Discrete-Time Affine Systems. *preprint*, https://doi.org/10.48550/arXiv.1810.09778, 2023.

First order inertial optimization algorithms with threshold effects associated with dry friction

Samir Adly[1] Hedy Attouch[2] Manh Hung Le[1]

[1]XLIM, Université de Limoges, France, samir.adly@unilim.fr, manh-hung.le@unilim.fr
[2]IMAG, Université Montpellier, France, hedy.attouch@umontpellier.fr

Keywords: proximal-gradient algorithms; inertial methods; dry friction; Hessian-driven damping; soft thresholding

In a Hilbert space setting, we consider new first order optimization algorithms which are obtained by temporal discretization of a damped inertial autonomous dynamic involving dry friction. The function f to be minimized is assumed to be differentiable (not necessarily convex). The dry friction potential function φ, which has a sharp minimum at the origin, enters the algorithm via its proximal mapping, which acts as a soft thresholding operator on the sum of the velocity and the gradient terms. After a finite number of steps, the structure of the algorithm changes, losing its inertial character to become the steepest descent method. The geometric damping driven by the Hessian of f makes it possible to control and attenuate the oscillations. The algorithm generates convergent sequences when f is convex, and in the nonconvex case when f satisfies the Kurdyka-Lojasiewicz property. The convergence results are robust with respect to numerical errors, and perturbations. The study is then extended to the case of a nonsmooth convex function f, in which case the algorithm involves the proximal operators of f and φ separately. Applications are given to the Lasso problem and nonsmooth d.c. programming. This talk is based on the paper [1]

References

[1] S. Adly, H. Attouch, M.H. Le, *First order inertial optimization algorithms with threshold effects associated with dry friction*, Comput. Optim. Appl. (2023).

Accessibility Constraints in Shape Optimization

Grégoire Allaire[1] Martin Bihr[2] Beniamin Bogosel[1] Matias Godoy[1]

[1]CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France
[2]Safran Tech, Rue des jeunes Bois, 78117 Chateaufort, France.
gregoire.allaire@polytechnique.fr, martin.bihr@safrangroup.com,
beniamin.bogosel@polytechnique.edu, mgodoycampbell@gmail.com

Keywords: accesibility, distance functions, additive manufacturing, shape optimization

In this work, described in detail in the article [1], we are interested in defining, quantifying and optimizing the accessibility of a shape with respect to a given source region. This question is motivated by practical aspects related to support structures in additive manufacturing. Support structures are used to guarantee the correct fabrication of a given part and these structures are removed at the end of the manufacturing process. To remove the supports, the contact region between the part and supports needs to be accessed with particular tools.

Two notions of accessibility in a straight line from a given source region are proposed, which from a mathematical point of view are expressed using distance functions, solutions of eikonal equations. The advantage of these criteria used for measuring the accessibility is the possibility to compute their shape derivatives. This allows us to implement numerical algorithms which optimize the accessibility with respect to variable targets and obstacles. Notably, it is possible to optimize simultaneously the mechanical performances of a structure and the accessibility of its supports.

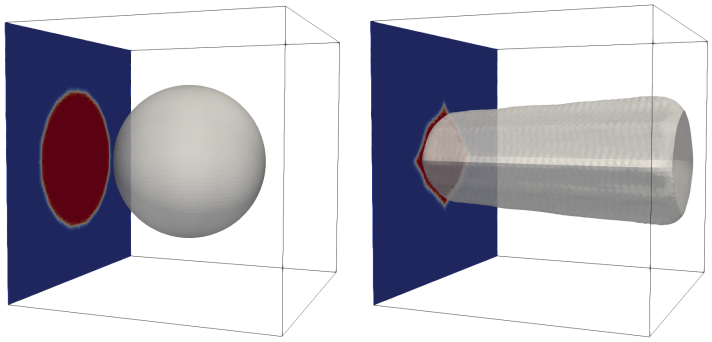


Figure 1: Numerical example where we minimize the area of the region on the target surface (marked in red) which is inaccessible in a straight normal line from the source surface.

References

[1] G. Allaire, M. Bihr, B. Bogosel, M. Godoy, Accessibility constraints in structural optimization via distance functions, *Journal of Computational Physics*, 484, 2023.

Classic-inspired quantum generation scheme for the parallel sequencing problem

Thibaud Ardant[1] Pascale Bendotti[1] Rodolphe Griset[1] Joseph Mikael[1]

[1]EDF, France, thibaud.ardant@polytechnique.edu, pascale.bendotti@edf.fr, rodolphe.griset@edf.fr, joseph.mikael@edf.fr

Keywords: Quantum computation, Grover Search, Scheduling Problem, Generation Schema

The Resource-Constrained Project Scheduling Problem (RCPSP) is to find a minimum makespan schedule for the jobs of a project under precedence constraints with limited renewable resource availability [1]. The RCPSP is an NP-hard problem known to be computationally challenging to solve. In the present work, we explore quantum computing approaches for the parallel sequencing variant defined as a special case of the RCPSP, where all jobs have a unit processing time and a unique predecessor. This variant is a generalization of Hu's polynomial case [2] to non unit renewable resources.

The key principle is to derive a feasible solution, namely a schedule, from a job permutation, i.e., an order of the jobs respecting the precedence constraints. Our approach relies on the quantum amplitude amplification algorithm [3] which requires to mark elements by a function called oracle. Taking benefit from the quantum superposition, this algorithm achieves a quadratic speedup on a large family of search problems for which good classical heuristics exist. The oracle is inspired from a classical parallel Schedule Generation Scheme (SGS) [4]. Parallel SGS is a method iterating over time periods to build a schedule based on priority rules. The oracle simulates the generation scheme on a predefined time horizon T, subsequently marking the permutations for which all jobs can be completed within the given horizon.

The classic-inspired quantum generation scheme is as follows. A preprocessing step computes an upper bound $\bar{T}$ on the optimal makespan using a classical heuristic. This upper bound is used as an input for the oracle which marks improving solutions. The quantum algorithm starts by generating a superposition of all states corresponding to valid permutations w.r.t the precedence constraints. Then the amplitude amplification algorithm amplifies the density of marked solutions. Finally a measurement retrieves an improving solution if exists. The oracle and the generation of the superposition restricted to valid permutations have been implemented using Qiskit library. Preliminary experiments performed on small instances exhibit the capability of the oracle to mark the improving valid solutions, thus tending to provide a proof of concept.

A perspective would be to set upper bound $\bar{T}$ iteratively to a value corresponding only to the few best solutions by performing a quantum counting [3], thus estimating the number of marked solutions. Such a scheme would solve the problem to optimality while taking advantage of quadratic speed-up obtained from both amplitude amplification and estimation compared to their deterministic counterpart.

References

[1] P. Brucker, A. Drexl, R. Mohring, K. Neumann, and E. Pesch. Resource-constrained project scheduling: Notation, classification, models, and methods. European journal of operational research, 112(1):3−41, 1999.

[2] T. C. Hu. Parallel sequencing and assembly line problems. Operations research, 9(6):841−848, 1961

[3] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. Contemporary Mathematics, 305:53−74, 2002

[4] R. Kolisch and S. Hartmann. Heuristic algorithms for the resource-constrained project schedul-ing problem: Classification and computational analysis. Springer, 1999

Characterization of the solutions of single-leader-multi-followers games with common constraints

Didier Aussel Gemayqzel Bouza Allende Sebastien Lepaul

University of Perpignan, France aussel@univ-perp.fr
Universidad de La Habana, Cuba gema@matcom.uh.cu
EdF R& D, Paris Saclay, France sebastien.lepaul@edf.fr

Nowadays non-cooperative games have many applications. A particular case is the following electricity contract problem. Consider there are certain packages the aggregator wants to offer to his/her clients. Each package has a limited amount of Mw and its price is decided by the aggregator. The clients choose how much to buy from each package for minimizing their total expenses. So, the solutions are Generalized Nash equilibrium points that depend on the prices of the packages decided by the aggregator, who wants to maximizes his/her profits. This is an example of a single-leader-multi-followers games with common constraints. In this paper we discuss on possible formulations of this kind of models. We obtain the properties that fulfill the solutions for a large class of problems, i.e. those that are satisfied for almost all perturbation of the objective functions and the constraints. Possible numerical solution approaches are considered.

The controller placement problem

Erison Ballasheni[2] Dritan Nace[1] Alban Zyle[1]

[1]Universite de Technologie de Compiègne, France, dritan.nace@hds.utc.fr, alban.zyle@etu.utc.fr
[2]Polytechnic University of Tirana, Albania, eballasheni@fti.edu.al,

Keywords: Controllers placement, Software-Defined Networking (SDN), network resilience, maximal coverage, network attacks

This study addresses the critical issue of controller placement in networks to fortify against cyber-attacks [1]. Node-to-node connections, essential conduits for inter-node communication, form the backbone of network functionality. Controllers, strategically positioned at select nodes play a pivotal role in facilitating the establishment of these connections. Node attacks can take various forms, one of the most frequent being the Denial-of-Service (DoS) Attacks The research centers on the deployment of controllers to safeguard the entire network. The objective is to minimize communication loss in the event of an attack. This problem aligns with the Maximal Coverage Location Problem (MCLP) [2] category, acknowledged for its NP-hard nature. The study extends this framework to the Equitable Controller Location Problem [3],[4], aiming to ensure equitable protection for all nodes under limited controller availability. A novel mathematical model is introduced for controller placement, leveraging lexicographic optimization [5] principles tailored to this context. The study introduces and compares both lexicographic and proportional fairness approaches for controller deployment.

References

[1] D. Santos, A. de Sousa, C. Mas-Machuca, and J. Rak, "Assessment of connectivity-based resilience to attacks against multiple nodes in sdns," *IEEE Access*, vol. 9, pp. 58 266–58 286, 2021.

[2] R. Church and C. ReVelle, "The maximal covering location problem," *Papers of the Regional Science Association*, vol. 32, no. 1, pp. 101–118, 1974. [Online]. Available: http://dx.doi.org/10.1007/BF01942293

[3] H. D. Luss, *Equitable Resource Allocation: Models, Algorithms and Applications*. Hoboken, NJ, USA: John Wiley & Sons, 2012.

[4] H. Luss, D. Nace, M. Poss, and M. C. Santos, "Equitable sensor location problems," in *ROADEF 2016*, Compiegne, France, 2016.

[5] D. Nace and M. Pioro, "Max-min fairness and its applications to routing and load-balancing in communication networks: a tutorial," *IEEE Commun. Surv. Tutor.*, vol. 10, no. 4, pp. 5–17, 2008.

QPLayer: efficient differentiation of convex quadratic optimization

A. Bambade[1] F. Schramm[2] A. Taylor[2] J. Carpentier[2]

[1]EDF R&D, France, antoine.bambade@edf.fr, [2]Inria & ENS Paris, France, fabian.schramm@inria.fr adrien.taylor@inria.fr, justin.carpentier@inria.fr

Keywords: Differentiable Optimization, Machine Learning, Quadratic Programming

Optimization layers within neural network architectures have become increasingly popular for their ability to solve a wide range of machine learning tasks and to model domain-specific knowledge. However, designing optimization layers requires careful consideration as the underlying optimization problems might be infeasible during training. Motivated by applications in learning, and control, this work focuses on convex quadratic programming (QP) layers. The specific structure of this type of optimization layer can be efficiently exploited for faster computations while still allowing rich modeling capabilities. We leverage primal-dual augmented Lagrangian techniques for computing derivatives of both feasible and infeasible QP solutions. More precisely, we propose a unified approach which tackles the differentiability of the *closest* feasible QP solutions in a classical ℓ_2 sense. The obtained Jacobian covers for feasible QPs the traditional implicit differentiation [Dontchev and Rockafellar(2009)] when it is valid and a weaker notion (i.e., conservative Jacobian by [Bolte et al.(2021)Bolte, Le, Pauwels, and Silveti-Falls]) when it is infeasible. We then harness this approach to enrich the expressive capabilities of existing QP layers. More precisely, we show how differentiating through infeasible QPs during training enables to drive towards feasibility at test time a new range of QP layers. These layers notably demonstrate superior predictive performance in some conventional learning tasks. Additionally, we present alternative formulations that enhance numerical robustness, speed, and accuracy for training such layers. The effectiveness of our approach is demonstrated in a few standard learning experiments, obtaining three to ten times faster computations than alternative state-of-the-art methods while being more accurate and numerically robust. Along with these contributions, we provide an open-source C++ software package called QPLayer for efficiently differentiating convex QPs and which can be interfaced with modern learning frameworks.

References

[Bolte et al.(2021)Bolte, Le, Pauwels, and Silveti-Falls] J. Bolte, T. Le, E. Pauwels, and A. Silveti-Falls. Nonsmooth implicit differentiation for machine-learning and optimization. In *NeurIPS*, pages 13537–13549, 2021. URL https://dblp.org/rec/conf/nips/BolteLPS21.

[Dontchev and Rockafellar(2009)] A. L. Dontchev and R. T. Rockafellar. Implicit functions and solution mappings. 2009.

Decision-Aware Learning with REINFORCE enables new Architectures and Applications

Léo Baty[1] Louis Bouvier[2] Mathieu Blondel[2] Youssouf Emine[3] Guillaume Dalle[4] Thibaut Vidal[3] Axel Parmentier[1]

[1]CERMICS, France, leo.baty@enpc.fr, louis.bouvier@enpc.fr, axel.parmentier@enpc.fr
[2]Google Brain Paris, France, mblondel@google.com
[3]Polytechnique Montréal, Canada, youssouf.emine@polymtl.ca, thibaut.vidal@cirrelt.ca
[4]EPFL Lausanne, Switzerland, guillaume.dalle.pro@gmail.com

Keywords: combinatorial optimization, machine learning, REINFORCE gradient

Using Combinatorial Optimization (CO) layers in Machine Learning (ML) pipelines is a recent area of interest (Kotary et al. 2021). These new hybrid methods provide fast differentiable heuristics for difficult CO problems. For instance, it can be used to tackle hard stochastic transportation problems, such as the stochastic vehicle scheduling (Dalle et al. 2022) or the dynamic vehicle routing (Baty et al. 2023) problems. Given an input instance x, a statistical model φ_w outputs a vector $\theta = \varphi_w(x)$. A CO oracle $f\colon \theta \mapsto \arg\max_{y \in \mathcal{Y}(x,\theta)} g(y, \theta)$ then outputs a feasible solution y.

To learn these pipelines, the goal is to find parameters w such that $f(\varphi_w(x))$ gives a good solution y of a hard problem $\min_y c(y, x)$ for any given instance x. However, current state-of-the-art methods face several limitations: most of them only support linear CO layers of the form $\arg\max_{y \in \mathcal{Y}(x)} \theta^\top y$, or need (near) optimal solutions to imitate in a supervised learning way. This limits a lot the possible applications and pipeline architectures for these methods.

We introduce what we call *perturbed oracles*, which can transform any CO oracle f into a distribution on the feasible solution space. This is done by perturbing the input θ of f using a distribution $p(\eta|\theta)$. It then turns the learning problem into a moment problem $\min_w \mathbb{E}_{\eta \sim p(\eta|\theta)}[c(f(\eta))]$, with non-zero gradients. Meaningful gradients can be computed by using the log-derivative/ REINFORCE trick, a well-known technique in ML and Reinforcement Learning.

This flexible approach unlocks new applications and architectures, such as learning constraints for the two-stage knapsack, and learning to sparsify graphs to scale the stochastic vehicle scheduling problem on large instances. Our numerical experiments show the efficiency of the approach.

References

Baty, L., K. Jungel, P. S. Klein, A. Parmentier, and M. Schiffer (2023). "Combinatorial optimization enriched machine learning to solve the dynamic vehicle routing problem with time windows". In: *arXiv preprint arXiv:2304.00789*.

Dalle, G., L. Baty, L. Bouvier, and A. Parmentier (July 27, 2022). "Learning with Combinatorial Optimization Layers: A Probabilistic Approach". In: arXiv:2207.13513. DOI: 10.48550/arXiv.2207.13513.

Kotary, J., F. Fioretto, P. Van Hentenryck, and B. Wilder (2021). "End-to-end constrained optimization learning: A survey". In: *arXiv preprint arXiv:2103.16378*.

An Extension of the Bertsimas & Sim Result for Discrete, Linear, and Γ-Robust Min-Max Problems

Yasmine Beck[1] Ivana Ljubić[2] Martin Schmidt[1]

[1]University of Trier, Germany,
yasmine.beck@uni-trier.de,martin.schmidt@uni-trier.de
[2]ESSEC Business School, France,
ljubic@essec.edu

Keywords: Bilevel Optimization, Robust Optimization, Mixed-Integer Programming, Branch-and-Cut, Knapsack Interdiction

Due to their nested structure, bilevel problems are intrinsically hard to solve—even if all variables are continuous and all parameters of the problem are exactly known. In this paper, we study mixed-integer linear bilevel problems with lower-level objective uncertainty, which we address using the notion of Γ-robustness. We provide an extension of the famous result by Bertsimas and Sim for Γ-robust single-level problems stating that the Γ-robust counterpart of a min-max bilevel problem can be solved by solving a polynomial number of min-max problems of the nominal type. Moreover, we discuss the situation in which the problems of the nominal type are not solved exactly but in which either an α-approximation algorithm or a primal heuristic together with a lower bounding scheme is used. For general Γ-robust bilevel problems, however, we illustrate that the previous ideas cannot be carried over completely. Nevertheless, we present a primal heuristic that is based on the solution of a polynomial number of problems of the nominal type. To assess the performance of the presented methods, we perform a computational study on 560 instances for both settings. For the primal heuristic, we observe that the optimality gap is closed for a significant portion of the considered instances. If our result for the min-max setting is exploited, we report speed-up factors exceeding 400 and 32 in the mean and the median, respectively, when compared to recently published problem-tailored solution approaches.

Estimation-of-Distribution Algorithms for Multi-Valued Decision Variables

Firas Ben Jedidia[1] Benjamin Doerr[2] Martin S. Krejca[2]

[1]École Polytechnique, IP Paris, France,
firas.ben-jedidia@polytechnique.edu
[2]LIX, CNRS, École Polytechnique, IP Paris, France,
benjamin.doerr@polytechnique.edu, martin.krejca@polytechnique.edu

Keywords: estimation-of-distribution algorithms, univariate marginal distribution algorithm, evolutionary algorithms, genetic drift, LeadingOnes benchmark

With apparently all research on estimation-of-distribution algorithms (EDAs) concentrated on pseudo-Boolean optimization and permutation problems, we undertake the first steps towards using EDAs for problems in which the decision variables can take more than two values, but which are not permutation problems. To this aim, we propose a natural way to extend the known univariate EDAs to such variables. Different from a naïve reduction to the binary case, it avoids additional constraints.

Since understanding genetic drift is crucial for an optimal parameter choice, we extend the known quantitative analysis of genetic drift to EDAs for multi-valued variables. Roughly speaking, when the variables take r different values, the time for genetic drift to become critical is r times shorter than in the binary case. Consequently, the update strength of the probabilistic model has to be chosen r times lower now.

To investigate how desired model updates take place in this framework, we undertake a mathematical runtime analysis on the r-valued LeadingOnes problem. We prove that with the right parameters, the multi-valued UMDA solves this problem efficiently in $O(r \log(r)^2 n^2 \log(n))$ function evaluations.

Overall, our work shows that EDAs can be adjusted to multi-valued problems and gives advice on how to set their parameters.

This work is part of the PGMO-funded project *Mathematical Analysis of Complex Randomized Search Heuristics* (PI: Benjamin Doerr).

References

[1] F. Ben Jedidia, B. Doerr, M. S. Krejca, Estimation-of-Distribution Algorithms for Multi-Valued Decision Variables, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, 230–238, 2023.

The No-Meet Matroid

Walid Ben-Ameur[1] Natalia Kushik[1] Alessandro Maddaloni[1]
José Neto[1] Dimitri Watel [2]

[1]Samovar, Telecom SudParis, Institut Polytechnique de Paris, Palaiseau, France
walid.benameur@telecom-sudparis.eu, natalia.kushik@telecom-sudparis.eu,
alessandro.maddaloni@telecom-sudparis.eu, jose.neto@telecom-sudparis.eu
[2]Samovar, ENSIIE, Evry, France, dimitri.watel@ensiie.fr

Keywords: Matroids, Graphs, Cycles, Orientation, Complexity

Consider a subset of vertices S of a directed network and assume $|S|$ vehicles move on the network: they start from S (one vehicle at each vertex of S) and at every step they move simultaneously to a neighbour vertex: we study if they can do this infinitely many times without ever meeting at any vertex (no-meet condition).
Meeting problems on graphs have gained their interest previously. Various results have been obtained in the area of merging or meeting of two or more tokens (or vehicles / robots in the literature). In particular, one can refer to a (deterministic) rendezvous problem where two given robots should meet at some point, given a set of instructions (see for example, [1]). The meeting time for random walks and an arbitrary number of tokens (for connected undirected graphs) has been also evaluated [2]. We study the contrary problem, *no-meet*, which turns out to be connected to matroid theory [3]:
We show that the subsets that meet the no-meet condition constitute the independent sets of a matroid. We prove that the independence oracle is polynomial-time and provide a more efficient algorithm to compute the size of a basis, namely polynomial algorithms to decide whether a subset has the no-meet condition and to calculate a largest subset that has the no-meet condition. We also report several connections with other matroid classes. For example, we show that the class of no-meet matroids strictly contains transversal matroids while it is strictly contained inside gammoids.
Then we focus on the orientation problem of the undirected edges of a mixed graph to either maximize or minimize the rank of a subset. Some NP-hardness results and inapproximability results are proved in the general case. Polynomial-time algorithms are described for subsets of size 1.

References

[1] A. Dessmark, P. Fraigniaud, D.R. Kowalski, Deterministic Rendezvous in Graphs *Algorithmica*, 46(1):69–96, 2006.

[2] N.H. Bshouty, L. Higham, J. Warpechowska-Gruca, Meeting Times of Random Walks on Graphs *Inf. Process. Lett.*, 69(5):259–265, 1999.

[3] J.G. Oxley, Matroid Theory (Oxford Graduate Texts in Mathematics) *Oxford University Press, Inc.*, 2006

Topology of spatiotemporal trajectories

Katherine Benjamin[1]

[1]University of Oxford, United Kingdom, katherine.benjamin@maths.ox.ac.uk

Keywords: Topological data analysis, Multiparameter persistent homology, Gröbner bases

In recent years, multiparameter persistent homology (MPH) has been developed to address the limitations of traditional single-parameter persistence. While much progress has been made on the theoretical aspects of MPH, practical applications have mostly been limited to the special case of two parameters. Building on recent developments in the computation of MPH in more than two parameters, we propose and implement a new Gröbner basis algorithm for the computation of multiparameter persistence landscapes in this setting. We apply this work to compute landscapes of three-parameter persistence modules built on spatiotemporal data arising in biology.

Integer linear formulations for the maximum flow blocker problem

Isma Bentoumi[1] Fabio Furini[2] A. Ridha Mahjoub[1] Sébastien Martin[3]

[1]LAMSADE, UMR CNRS 7243, University Paris-Dauphine-PSL, France
ridha.mahjoub@lamsade.dauphine.fr, isma.bentoumi@dauphine.eu
[2] Department of Computer, Control and Management Engineering, Sapienza University of Rome.
fabio.furini@uniroma1.it
[3]Huawei Technologies France, France
sebastien.martin@huawei.com

Keywords: Combinatorial optimization, Bi-level optimization, Blocker, Interdiction, Maximum flow

Given a directed graph with capacities and blocker costs associated with its arcs, we study the maximum flow blocker problem (MFBP) which asks to find a minimum-cost subset of arcs to be removed from the graph in such a way that the remaining maximum-flow value does not exceed a given threshold. The MFBP has applications in telecommunication networks and in the monitoring of civil infrastructures, among others.

To the best of our knowledge, the MFBP has not been addressed in the literature but a closely-related problem, called maximum flow interdiction problem (MFIP), has been largely studied. This problem consists in finding a subset of arcs to be removed from the graph such that the remaining maximum flow is minimized. The set of interdicted arcs is limited by a budget. We refer the interested reader to [1] and [2] for a complete overview of the MFIP.

We propose two integer linear programming (ILP) formulations for the MFBP which are used to derive the first exact algorithms to optimally solve the problem. The first one relies on the bilevel structure of the MFBP. It has an exponential number of constraints and accordingly it is solved via a Branch-and-Cut algorithm. The second one has a polynomial number of variables and constraints and it is solved via an ILP solver. This formulation is obtained thanks to a property that we have demonstrated, which establishes a relationship between the MFBP and the MFIP. The two formulations are compared through an extensive computational study. Our tests aim at evaluating the performance of the exact algorithms and at determining the features of the instances which can be solved to proven optimality.

References

[1] R. K. Wood. Deterministic network interdiction. *Mathematical and Computer Modeling*,**17**(2):1-18, 1993.

[2] J. Royset and K. R. Wood. Solving the bi-objective maximum-flow network-interdiction problem. *INFORMS Journal on Computing*,**19**(2):175-184, 2007.

The Virtual Network Function Reconfiguration Problem

Hanane Biallach[1,2] Mustapha Bouhtou[2] Kristina Kumbria[1] Dritan Nace[1] Artur Tomaszewski[3]

[1]Université de Technologie de Compiègne, France, kristina.kumbria@hds.utc.fr, dritan.nace@hds.utc.fr
[2]Orange Labs Châtillon, France, hanane.biallach@gmail.com, mustapha.bouhtou@orange.com
[3]Warsaw University of Technology, Poland, artur.tomaszewski@pw.edu.pl

Keywords: 5G network slicing, VNFs reconfiguration, Knapsack Problem, Column Generation

One of the challenges in the administration of 5G networks revolves around the need to reconfigure network slices [1, 2]. This entails the process of reconfiguring and relocating Virtual Network Functions (VNFs) to align with the specific demands of the network slices and the resource capacities within data centers. We consider a problem of allocating Virtual Network Functions (VNFs) in 5G networks, treating it as a multiple multidimensional knapsacks reconfiguration problem where items represent VNFs and knapsacks represent servers on which VNFs are able to run. When provided with both the starting and ending allocations of items to knapsacks, the challenge involves defining a sequence of steps for transferring items between source and destination knapsacks. The ultimate goal is to reduce the overall count of steps during which items reside outside the knapsacks, thereby minimizing the total downtime for Virtual Network Functions (VNFs). We define two exact, compact Integer Linear Programming (ILP) formulations of the VNFs reconfiguration problem, and derive one approximate solution algorithm based on the Column Generation (CG) method [3]. Numerical tests and comparisons illustrate the performance of the algorithms and the quality of the provided solutions [4, 5, 6].

References

[1] 5g; procedures for the 5g system (3gpp ts 23.502 version 15.2.0 release 15), v15.2.0

[2] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, Network slicing in 5g: Survey and challenges, *IEEE Communications Magazine*, vol. 55, no. 5, pp. 94–100, 2017.

[3] Guy Desaulniers, Jacques Desrosiers and Marius M. Solomon, Column generation, *GERAD 25th Anniversary Series*, 2005.

[4] K. Kumbria, D. Demko, D. Nace, A. Tomaszewski, M. Bouhtou, and H. Biallach, VNFs reconfiguration in 5g networks, *12th International Workshop on Resilient Networks Design and Modeling (RNDM)*, 2022.

[5] H. Biallach, M. Bouhtou, and D. Nace, An efficient heuristic for the virtual network function reconfiguration problem, *12th International Workshop on Resilient Networks Design and Modeling (RNDM)*, 2022.

[6] H. Biallach, Optimization of VNF reconfiguration problem for 5G network slicing

Optimal Trees and Multivariate Functional Data

Rafael Blanquero[1] Emilio Carrizosa[1] Cristina Molero-Río[2] Dolores Romero Morales[3]

[1]Universidad de Sevilla, Spain, rblanquero@us.es, ecarrizosa@us.es
[2]École Polytechnique, France, molero@lix.polytechnique.fr
[3]Copenhagen Business School, Denmark, drm.eco@cbs.dk

Keywords: optimal randomized regression trees, multivariate functional data, critical intervals detection, nonlinear programming

In this talk, we tailor optimal regression trees to deal with multivariate functional data. A compromise between prediction accuracy and sparsity, as a proxy for interpretability, is sought. In particular, whilst fitting the tree model, the detection of a reduced number of predictor variables and the proportion of the domain used by the model is performed. This is achieved through the inclusion of LASSO-type regularization terms. The resulting optimization problem can be formulated as a nonlinear continuous optimization model with linear constraints. We illustrate the performance of our approach on real-world and synthetic datasets. For more details, see [1].

References

[1] R. Blanquero, E. Carrizosa, C. Molero-Río, D. Romero Morales, On optimal regression trees to detect critical intervals for multivariate functional data, *Computers & Operations Research*, 152:106152, 2023.

Learning Energy Networks with Generalized Fenchel-Young Losses

Mathieu Blondel, Felipe Llinares-López,
Robert Dadashi, Léonard Hussenot, Matthieu Geist
[1]Google

Keywords: energy networks, generalized conjugates, loss functions

Energy-based models, a.k.a. energy networks, perform inference by optimizing an energy function, typically parametrized by a neural network. This allows one to capture potentially complex relationships between inputs and outputs. To learn the parameters of the energy function, the solution to that optimization problem is typically fed into a loss function. The key challenge for training energy networks lies in computing loss gradients, as this typically requires argmin/argmax differentiation. In this paper, building upon a generalized notion of conjugate function, which replaces the usual bilinear pairing with a general energy function, we propose generalized Fenchel-Young losses, a natural loss construction for learning energy networks. Our losses enjoy many desirable properties and their gradients can be computed efficiently without argmin/argmax differentiation. We also prove the calibration of their excess risk in the case of linear-concave energies. We demonstrate our losses on multilabel classification and imitation learning tasks.

Neural networks for deterministic two-player games

Olivier Bokanowski[1] Xavier Warin [2]

[1] Université Paris Cité, LJLL, France boka@math.univ-paris-diderot.fr
[2] EDF R&D & FiME, France xavier.warin@edf.fr

Keywords: two-player games, neural networks, deterministic optimal control, dynamic programming principle, state constraints, front propagation

Abstract

We consider the value of some deterministic optimal control problem for two-player games. We propose time-marching schemes and deep neural network approximations, using a Lagrangian approach of [1] which was developed in the context of one-player optimal control problems. Our analysis aims to give error estimates in an average norm for the considered value. Some academic numerical examples related to reachability will be considered, where level set functions are used in order to represent the reachable sets and SGDA type algorithms are utilized for min/max optimization (see e.g. [2]).

References

[1] O. Bokanowska, A. Prost, X. Warin. *Neural networks for first order HJB equations and application to front propagation with obstacle terms,* ArXiv:2210.04300, 2022 (to appear in PDEA)

[2] M. Razaviyayn, T. Huang, S. Lu, M. Nouiehed, M. Sanjabi and M. Hong, *Nonconvex Min-Max Optimization: Applications, Challenges, and Recent Theoretical Advances,* in IEEE Signal Processing Magazine, vol. 37, no. 5, pp. 55-66, 2020.

An Entropy Penalized Approach for Stochastic Optimal Control Problems

Thibaut Bourdais Nadia Oudjane Francesco Russo

ENSTA Paris, France, t.bourdais34@gmail.com
EDF R& D and Fime, France, nadia.oudjane@edf.fr
ENSTA Paris, France, francesco.russo@ensta-paris.fr

We propose an alternative technique to dynamic programming for solving stochastic control problems. We consider a weak formulation that is written as an optimization (minimization) problem on the space of probabilities. We then propose a regularized version of this problem obtained by splitting the minimization variables and penalizing the entropy between the two probabilities to be optimized. We show that the regularized problem provides a good approximation of the original problem when the weight of the entropy regularization term is large enough. Moreover, the regularized problem has the advantage of giving rise to optimization problems that are easy to solve in each of the two optimization variables when the other is fixed. We take advantage of this property to propose an alternating optimization algorithm whose convergence to the infimum of the regularized problem is shown. The relevance of this approach is illustrated by solving a high-dimensional stochastic control problem aimed at controlling consumption in electrical systems.

Renault's Forward Logistic Problem

Mathis Brichet[1] Axel Parmentier[1] Maximilian Schiffer[2]

[1]CERMICS, Ecole des Ponts ParisTech, France,
mathis.brichet@eleves.enpc.fr, axel.parmentier@enpc.fr
[2]BAIS, Technical University of Munich, Germany, schiffer@tum.de

Keywords: Network design, multicommodity flow, large scale combinatorial optimization

The inbound supply chain of a carmaker is designed to transport tens of thousands of types of car parts from thousands of suppliers to dozens of plants. This logistics operation represents tens of millions of euros and tens of thousands of tons of CO2 emissions per year. The arrival of a new vehicle in the catalog of a carmaker requires a redesign of the network. The resulting problem is a rich version of a multicommodity network design problem [1]. Its difficulty stems from the combination of a large network and a huge number of commodities to be routed, nearly a million. The considered problem is therefore orders of magnitude bigger than what classical resolution methods from the litterature can solve [2, 3, 4]. We model this network design problem and propose a heuristic algorithm to solve it in reasonable time. We provide a data analysis of their instance and a lower bound to analyze algorithm performance. Numerical experiments show significant improvement over Renault's current solution but also point to a large margin of possible improvement. We finally conclude on the future research perspectives arising from the insights collected in this work.

References

[1] T.G. Crainic, M. Gendreau, B. Gendron, Network Design with Applications to Transportation and Logistics, *Springer International Publishing*, 2021.

[2] T.G. Crainic, B. Gendron, Exact Methods for Fixed-Charge Network Design, *Network Design with Applications to Transportation and Logistics*, 29–89, 2020.

[3] B. Gendron, M. Larose, Branch-and-price-and-cut for large-scale multicommodity capacitated fixed-charge network design, *EURO Journal on Computational Optimization*, 2014.

[4] B. Gendron, S. Hanafi, R. Todosijevi, Matheuristics based on iterative linear programming and slope scaling for multicommodity capacitated fixed charge network design, *European Journal of Operational Research*, 2018.

Displacement smoothness of entropic optimal transport and applications

Guillaume Carlier[1] Lénaïc Chizat[2] Maxime Laborde[2]

[1] Ceremade, Université Paris Dauphine, PSL France, carlier@ceremade.dauphine.fr
[2] EPFL, Lausanne, Switzerland, [3] Université Paris Cité

Keywords: entropic optimal transport, displacement smoothness, gradient flows, semi geostrophic equations

Entropic optimal transport (EOT) has received a lot of attention in recent years due to its numerical tractability (Sinkhorn algorithm) as well as its connections to the Schrödinger bridge problem. In this talk, I will first give some Lipschitz dependence results for the potentials of EOT (in C^k norm) with respect to the marginals (for the Wasserstein metric). As a first application, well posedness for Wasserstein gradient flows (with possibly several species) of energies involving EOT terms will follow. Finally, I will describe another application which concerns the so-called semi geostrophic equation, which has an Hamiltonian structure, well-posedness and convergence of its entropic approximation will be explained (work in progress with Jean-David Benamou and Hugo Malamut).

References

[1] G. Carlier, L. Chizat, M. Laborde, *Displacement Smoothness of Entropic Optimal Transport*, arxiv.

Latency-Sensitive Service Chaining with Security Isolation Constraints

Yannick Carlinet[1] Nancy Perrot[1] Laurent Valeyre[1] Jean-Philippe Wary[1]

[1]Orange Innovation, France, firstname.lastname@orange.com

Keywords: Service Chain, Edge Cloud, Edge Computing, Latency-sensitive Service, Security, Security Isolation

Multi-access edge computing (MEC) is a distributed computing architecture that brings computing and storage resources closer to the network edge. It enables applications and services to be hosted and processed at the network edge, rather than in centralized data centers or cloud platforms. By bringing computing closer to the end-users, MEC can reduce network latency, improve application performance, and support emerging use cases that require low latency and high bandwidth, such as augmented reality, virtual reality, or autonomous vehicles. MEC can also enable new business models and revenue streams for network operators and service providers, by providing value-added services, such as content caching, video optimization, or network slicing.

However, there are also some challenges that need to be addressed, including:

- Resource Constraints: Edge nodes typically have limited computing and storage resources compared to centralized data centers, which can limit the types and scale of applications that can be deployed on edge nodes.
- Security: MEC introduces new security challenges, such as securing distributed computing resources and protecting sensitive data at the edge of the network. More specifically, an exploit can enable a rogue (or compromised) application to have access to non-authorized parts of RAM, thus allowing to steal or poison the data of the other applications running on the same physical host (sidechannel attack or compromised host). In a cloud environment, applications with greater security requirements would run on dedicated hardware. However, this is no longer possible in MEC because of the scarcity of resources.
- Orchestration: MEC requires the management and orchestration of resources across multiple edge nodes and clusters, while meeting the requirements of service chains, including security and end-to-end latency constraints.

We focus on the orchestration challenges in MEC. In particular, this work brings the following contributions:

- Description and modelling of an optimization problem, called SLSCP (Security and Latency-aware Service Chaining Problem) ;
- Algorithms that provide solutions for SLSCP ;
- Set-up of a platform that emulates a realistic Edge Computing environment and integration of the designed algorithms

Stochastic Primal Dual Hybrid Gradient Algorithm with Adaptive Step-Sizes

Antonin Chambolle[1] Claire Delplancke[2] Matthias J. Ehrhardt[3]
Carola-Bibiane Schönlieb[4] Junqi Tang[5]

[1]CEREMADE, Université Paris-Dauphine, France, chambolle@ceremade.dauphine.fr
[2]EDF Lab Paris-Saclay, France, claire.delplancke@edf.fr,
[3]Department of Mathematical Sciences, University of Bath, UK m.ehrhardt@bath.ac.uk,
[4]DAMTP, University of Cambridge, UK, cbs31@cam.ac.uk,
[5]School of Mathematics, University of Birmingham, UK j.tang.2@bham.ac.uk,

Keywords: primal-dual algorithm, stochastic algorithm, adaptive step-sizes, computed tomography.

In this work we propose a new primal-dual algorithm with adaptive step-sizes. The stochastic primal-dual hybrid gradient (SPDHG) algorithm with constant step-sizes [1, 2] has become widely applied in large-scale convex optimization across many scientific fields due to its scalability. While the product of the primal and dual step-sizes is subject to an upper-bound in order to ensure convergence, the selection of the ratio of the step-sizes is critical in applications. Up-to-now there is no systematic and successful way of selecting the primal and dual step-sizes for SPDHG. In this work, we propose a general class of adaptive SPDHG (A-SPDHG) algorithms, and prove their convergence under weak assumptions. We also propose concrete parameters-updating strategies which satisfy the assumptions of our theory and thereby lead to convergent algorithms. Numerical examples on computed tomography demonstrate the effectiveness of the proposed schemes.

References

[1] Antonin Chambolle, Matthias J. Ehrhardt, Peter Richtárik, and Carola-Bibiane Schönlieb. Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications. *SIAM Journal on Optimization*, 28(4):2783-2808, 2018.

[2] Matthias J. Ehrhardt, Pawel Markiewicz, and Carola-Bibiane Schönlieb. Faster PET reconstruction with non-smooth priors by randomization and preconditioning. *Physics in Medicine & Biology*, 64(22):225019, 2019.

Abstract Cutting Plane Method applied to Sparse Optimization

Jean-Philippe Chancelier [1] Michel De Lara[1]
Seta Rakotomandimby[1]
[1]École des Ponts ParisTech

Keywords: abstract convexity, abstract cutting plane method, generalized convexity, sparse optimization

In usual convexity, a function is closed convex if and only if it is the supremum of its affine minorants. In abstract convexity, affine functions are replaced by other elementary functions belonging to some set H. By definition, a function is abstract H-convex if it is the supremum of its H-minorants [3, 5, 4]. Few abstract H-convex minimization algorithms have been studied [3, Chapter 9], [4, Chapter 9], such as the abstract cutting plane method and the abstract branch-and-bound method. Recent work [1] on the ℓ_0 pseudonorm has highlighted elementary functions that makes the ℓ_0 pseudonorm abstract H-convex — the so-called E-CAPRA affine functions. As explicit formulas for the E-CAPRA subdifferential of the ℓ_0 pseudonorm have been calculated [2], it has been made possible to implement abstract H-convex minimization algorithms to the special case of the ℓ_0 pseudonorm.

We present our numerical tests on the E-CAPRA cutting plane method applied to the following sparse optimization problems: the minimization of the ℓ_0 pseudonorm in a blunt closed cone, and the computation of the spark of a matrix (known to be a NP-hard problem [6]).

References

[1] J.-P. Chancelier and M. De Lara. Capra-convexity, convex factorization and variational formulations for the l_0 pseudonorm. *Set-Valued and Variational Analysis*, 30:597–619, 2022.

[2] A. L. Franc, J.-P. Chancelier, and M. De Lara. The capra-subdifferential of the l0 pseudonorm. *Optimization*, pages 1–23, 2022. accepted for publication.

[3] D. Pallaschke and S. Rolewicz. *Foundations of mathematical optimization*, volume 388 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1997.

[4] A. Rubinov. *Abstract convexity and global optimization*, volume 44 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.

[5] I. Singer. *Abstract Convex Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons, Inc., New York, 1997.

[6] A. M. Tillmann and M. E. Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, 2014.

Expected Runtime of the Compact Genetic Algorithm on the LeadingOnes Benchmark

Marcel Chwiałkowski[1] Benjamin Doerr[2] Martin Krejca[2]

[1]École Polytechnique, IP Paris, France, marcel.chwialkowski@polytechnique.edu
[2]LIX, École Polytechnique, IP Paris, Palaiseau, France, firstname.lastname@polytechnique.edu

Keywords: Estimation-of-distribution algorithms, runtime analysis.

Estimation-of-distribution algorithms (EDAs) are iterative methods that heuristically solve optimization problems by maintaining an explicit probabilistic model of the search space. More precisely, in each iteration an EDA randomly samples some potential solutions according to its probabilistic model, evaluates their objective-function value (referred to as *fitness*), and updates the model so that it has higher probability of sampling a solution with better fitness.

Over the past years the performance of EDAs has been studied extensively via theoretical means. The results suggest that EDAs can have an advantage over other randomized search heuristics in particular when coping with local optima or with noise [1, 2].

The compact genetic algorithm (cGA) [3] is one of the simpler EDAs. In each iteration it samples only two solutions. Despite its simplicity, there is still no rigorous result on its performance on the classic LeadingOnes benchmark, one of the most-studied benchmarks in the theory of randomized search heuristics.

In this work, we study the performance of the cGA on LeadingOnes. The effectiveness of the cGA highly depends on its sole parameter μ, the so-called *hypothetical population size*. We prove that when $\mu = \omega(n \log^2 n)$, then with high probability the cGA samples an optimal solution in $\mathcal{O}(\mu n \log n)$ iterations.

This work is part of the PGMO-funded project *Mathematical Analysis of Complex Randomized Search Heuristics* (PI: Benjamin Doerr).

References

[1] B. Doerr. The runtime of the compact genetic algorithm on Jump functions. *Algorithmica*, 83:3059–3107, 2021.

[2] T. Friedrich, T. Kötzing, M. S. Krejca, and A. M. Sutton. The compact genetic algorithm is efficient under extreme Gaussian noise. *IEEE Transactions on Evolutionary Computation*, 21:477–490, 2017.

[3] F. G. L. Georges R. Harink and D. E. Goldberg. The compact genetic algorithm. *IEEE Transactions on Evolutionary Computation*, 1963.

Optimal Sets for the L_∞ Star Discrepancy

François Clément[1] Carola Doerr[1] Kathrin Klamroth[2] Luís Paquete[3]

[1] Sorbonne Université, CNRS, LIP6, Paris France, francois.clement@lip6.fr, carola.doerr@lip6.fr
[2] University of Wuppertal, Wuppertal, Germany, klamroth@uni-wuppertal.de
[3] University of Coimbra, CISUC, Department of Informatics Engineering, Coimbra, Portugal, paquete@dei.uc.pt

Keywords: Discrepancy, Computational geometry, Non-Linear Programming

Discrepancy measures are metrics designed to quantify how well-spread a point set is in a given space. Among the different measures, the L_∞ star discrepancy is the most important, measuring the largest deviation between the volume of any box $[0, q(, q \leq 1$, and the proportion of points that are inside this box. Point sets of low discrepancy are used in many contexts, including Quasi-Monte Carlo integration, financial mathematics and one-shot optimization. Despite their wide-spread use, few people have directly tried to find optimal sets for this measure. In dimension 1, the optimal sets are known for all n [1], but already in dimension 2 results are limited to $n \leq 6$ by White [2]. More recent works have focused either on very few points in higher dimensions ([3],[4]) or other measures [5].

In this talk, we introduce two non-linear programming formulations to find optimal L_∞ star discrepancy sets. With these, we are able to compute optimal sets for $n \leq 20$ in dimension 2 and $n \leq 8$ in dimension 3. These sets present a very different structure to that of known low-discrepancy sets and sequences such as Sobol' or Kronecker , possibly suggesting a new construction method. Furthermore, our formulations can be easily adapted to other related problems such as adding one or many points to an existing set, finding optimal lattices or optimal sets for the 4-corner star discrepancy which considers all boxes anchored in one of the corners of $[0, 1]^d$.

References

[1] H. Niederreiter, Discrepancy and convex programming, *Ann. Mat. Pura Appl.*, 93:89-97, 1972.

[2] B.E. White, On optimal extreme-discrepancy point sets in the square, *Numer. Math.*, 27:157-164, 1976/1977.

[3] Tim Pillards, Bart Vandewoestyne, and Ronald Cools, Minimizing the L_2 and L_∞ star discrepancies of a single point in the unit hypercube, *Journal of Computational and Applied Mathematics*, 197:282-285, 12, 2006.

[4] Gerhard Larcher and Friedrich Pillichshammer, A note on optimal point distributions in $[0, 1)^s$, *Journal of Computational and Applied Mathematics*, 206(2):977-985, 2007

[5] Aicke Hinrichs and Jens Oettershagen, Optimal point sets for quasi-monte carlo integration of bivariate periodic functions with bounded mixed derivatives,*Monte Carlo and Quasi-Monte Carlo Methods*, 2014.

A Planning Game for Global Decarbonization Strategies

A. Couthures[1] B. Mroué[1] S. Lasaulce[1,2] I. C. Morărescu[1]

[1]Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France
[2]Khalifa University, Abu Dhabi, UAE
anthony.couthures@univ-lorraine.fr, bouchra.mroue@univ-lorraine.fr,
samson.lasaulce@univ-lorraine.fr, constantin.morarescu@univ-lorraine.fr

Keywords: Game Theory, Decarbonization, Climate Change, Modeling, Potential game

The ongoing concerns surrounding climate change and its irreversible impacts necessitate a deeper understanding of the commitment levels and behavioral nuances of governments toward decarbonization. In our previous work, we established the dynamics of government players' actions regarding carbon emission levels, emphasizing the interplay between GDP benefits and the repercussions of climate change.

In this year's presentation, we further enrich our study by introducing two decision-making methodologies: a planning method and a receding horizon method. This allows us to delve deeper into the strategic decisions made by governments and ascertain the critical trade-off between global temperature rise and GDP. Such a rigorous examination provides valuable insights into certain behavioral tendencies that appear invariant to the choice of the damage function.

Furthermore, we augment our study with the introduction of new damage functions, offering a broader and more comprehensive perspective on the potential consequences of emissions. By examining the utilities, we aim to understand better the optimal agreeing time for both the planning and receding horizon methods for in favor of government and climate.

In the culmination of our analysis, we establish conditions on the benefits and damage functions that guarantee the uniqueness of a pure Nash equilibrium. This is instrumental in not only characterizing the strategic interactions among governments but also in identifying scenarios where cooperative outcomes are more likely.

References

[1] P. Chander, Game Theory and Climate Change, *Columbia University Press*, 2018.

[2] Bouchra Mroué, Anthony Couthures, Samson Lasaulce, Irinel-Constantin Morărescu, Une approche par la théorie des jeux du problème du choix des niveaux nationaux d'émissions de CO2, *GRETSI 2023*, 2023.

[3] M. Finus, Game Theory and International Environmental Cooperation: Any Practical Application?, *Edward Elgar Publishing*, 2002.

Minimizing a Separable Sum Coupled by a Difference of Functions and Linear Constraints

Minh N. Dao[1]

[1]RMIT Univeristy, Australia, minh.dao@rmit.edu.au

Keywords: Alternating direction method of multipliers, Bregman distance, composite optimization problem, difference of functions, linear constraints.

In this work, we develop a splitting algorithm for solving a broad class of linearly constrained composite optimization problems, whose objective function is the separable sum of possibly nonconvex nonsmooth functions and a smooth function, coupled by a difference of functions. This structure encapsulates numerous significant nonconvex and nonsmooth optimization problems in the current literature including the linearly constrained difference-of-convex problems. Relying on the successive linearization and alternating direction method of multipliers, the proposed algorithm exhibits the global subsequential convergence to a stationary point of the underlying problem. We also establish the convergence of the full sequence generated by our algorithm under the Kurdyka–Łojasiewicz property and some mild assumptions. The efficiency of the proposed algorithm is tested on a robust principal component analysis problem and a nonconvex optimal power flow problem.

Fisher Market Model based Resource Allocation for 5G Network Slicing

Mandar Datar[1] Naresh Modina[2] Rachid El Azouzi[3] Eitan Altman[4]

[1] Orange Innovation,Châtillon, France mandar.datar@orange.com
[2] (CEDRIC) CNAM - Paris, France naresh.modina@cnam.fr,
[3] CERI/LIA, University of Avignon, France Rachid.Elazouzi@univ-avignon.fr
[4] INRIA Sophia Antipolis-Mediterranean, France eitan.altman@inria.fr

Keywords: Network slicing, Resource allocation, Fisher Market, Market Equilibrium, Trading post mechanism, Decentralized learning.

Network slicing (NS) is a potential technology to support a higher degree of heterogeneity and flexibility required by next-generation services in 5G networks. One of the essential elements in network slicing is resource allocation, which needs to ensure slice tenants the protection of their service level agreements (SLAs) while optimizing network efficiency. We propose a resource-sharing scheme based on the Fisher market model and the Trading post mechanism that can achieve efficient resource utilization while catering to multi-level fairness, dynamic load conditions and SLA protection. In the proposed scheme, each service provider (SP) is given a budget representing its infrastructure share or purchasing power in the market. The SPs acquire different resources by spending their budgets to offer the service to different classes of users, which are classified according to their service needs and priorities. We assume that SPs employ the α fairness criteria while delivering the service to their subscribers. The proposed allocation scheme aims to find a market equilibrium(ME) that provides allocation and resource pricing whereby each SP's needs are met, and resources are fully utilized. We show that the ME solution problem can be formulated as a convex optimization problem whose primal and dual optimal solution provides equilibrium allocation and pricing. We build a privacy-preserving learning algorithm enabling SPs to reach ME in a decentralized fashion. We theoretically evaluate the proposed allocation scheme's performance by comparing it with the Social Optimal and Static Proportional sharing schemes. Finally, we run extensive numerical simulations to assess the fairness and efficiency properties of the proposed scheme.

References

[1] D. T. Nguyen and L. Le and V. Bhargava, A Market-Based Framework for Multi-Resource Allocation in Fog Computing, *IEEE/ACM Transactions on Networking*, 27:1151-1164, 2019.

[2] Y. K. Cheung and R. Cole and Yixin Tao, Dynamics of Distributed Updating in Fisher Markets, *Proceedings of the 2018 ACM Conference on Economics and Computation*, 2018.

On the Optimal Rate for the Convergence Problem in Mean-Field Control

Samuel Daudin[1] François Delarue[2] Joe Jackson[2]

[1]Université Côte d'Azur, France, sdaudin@unice.fr
[2]Université Côte d'Azur, France, francois.delarue@unice.fr,
[3] University of Chicago, USA,jjackso1@utexas.edu

Keywords: Mean-Field Control, Convergence Problem, Rate of Convergence, PDEs in the Space of Probability Measures

The goal of this work is to obtain optimal rates for the convergence problem in mean field control. Our analysis covers cases where the solutions to the limiting problem may not be unique nor stable. Equivalently the value function of the limiting problem might not be differentiable on the entire space. Our main result is then to derive sharp rates of convergence in two distinct regimes. When the data is sufficiently regular, we obtain rates proportional to $N^{-1/2}$, with N being the number of particles. When the data is merely Lipschitz and semi-concave with respect to the first Wasserstein distance, we obtain rates proportional to $N^{-2/(3d+6)}$. Noticeably, the exponent $2/(3d+6)$ is close to $1/d$, which is the optimal rate of convergence for uncontrolled particle systems driven by data with a similar regularity. The key argument in our approach consists in mollifying the value function of the limiting problem in order to produce functions that are almost classical sub-solutions to the limiting Hamilton-Jacobi equation (which is a PDE set on the space of probability measures). These sub-solutions can be projected onto finite dimensional spaces and then compared with the value functions associated with the particle systems. In the end, this comparison is used to prove the most demanding bound in the estimates. The key challenge therein is thus to exhibit an appropriate form of mollification. We do so by employing sup-convolution within a convenient functional Hilbert space. To make the whole easier, we limit ourselves to the periodic setting. We also provide some examples to show that our results are sharp up to some extent.

This is based on the preprint [1]

References

[1] Samuel Daudin, François Delarue, and Joe Jackson. On the Optimal Rate for the Convergence Problem in Mean Field Control. *arXiv:2305.08423*, 2023.

Algorithms in Generalized Convexity

Jean-Philippe Chancelier [1] Michel De Lara[1]
Seta Rakotomandimby[1]
[1]École des Ponts ParisTech

Keywords: generalized convexity, algorithm, sparse optimization

In generalized convexity, the duality product is replaced by a coupling c, the Fenchel conjugate is replaced by the conjugation associated to the coupling, and closed convex functions are replaced by c-convex functions (defined as functions that are equal to their biconjugates) [2]. This field has been mostly the subject of theoretical works [3, 4, 5] and, more rarely, of algorithmic ones [3, Chapter 9], [4, Chapter 9]. Recent work [1] on the ℓ_0 pseudonorm has highlighted a so-called E-Capra coupling and conjugation which could lead to the application of generalized convex algorithms to sparse optimization.

In this context, we intend to develop new algorithms in generalized convexity from both theoretical and numerical standpoints. Here, we focus on the theoretical perspective and we present: how to define, in generalized convexity, Bregman distance and proximal operator; how to construct and analyze dual problems; how to design sparsity-inducing regularizers.

References

[1] J.-P. Chancelier and M. De Lara. Capra-convexity, convex factorization and variational formulations for the l_0 pseudonorm. *Set-Valued and Variational Analysis*, 30:597–619, 2022.

[2] J. J. Moreau. Fonctionnelles convexes. *Séminaire Jean Leray*, 2:1–108, 1966-1967.

[3] D. Pallaschke and S. Rolewicz. *Foundations of mathematical optimization*, volume 388 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1997.

[4] A. Rubinov. *Abstract convexity and global optimization*, volume 44 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.

[5] I. Singer. *Duality for Nonconvex Approximation and Optimization*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2006.

Optimal location of charging station: a Single-Leader-Follower approach with cardinality constraints

Aussel Didier Daniel Lasluisa David Salas

University of Perpignan, France
Universidad de Chile, dlasluisa@dim.uchile.cl,david.salas@uoh.cl

When deciding about the location of charging stations for electrical vehicles, different conflicting aims can enter into the scope: minimal cost for investors, minimal inconvenience for EV owners, limited number of stations... In order to take into account all these features in an optimal way, we analyzed different Single-Leader-Follower models in which cardinality constraints are intrinsically integrated.

Decarbonizing OCP

Dimitris Bertsimas[1] Ryan Cory-Wright[2] Vassilis Digalakis Jr[3]

[1]Massachusetts Institute of Technology, USA, dbertsim@mit.edu
[2]Imperial College Business School, UK, r.cory-wright@imperial.ac.uk
[2]HEC Paris, France, digalakis@hec.fr

Keywords: Decarbonization, robust optimization, renewable energy integration, energy analytics

I will present our collaboration with OCP, one of the world's largest fertilizer producers, to decarbonize its electricity supply through the installation of solar panels and batteries. We develop an optimization framework that minimizes OCP's time-discounted investment cost plus the cost of satisfying its remaining demand via the national grid. We immunize our model against deviations between forecast and realized solar generation output via a combination of robust and distributionally robust optimization, to account, respectively, for variability in daily solar generation and in seasonal weather patterns induced by climate change. OCP is currently designing its renewable investment strategy, using insights gleaned from our optimization model, and has pledged to invest over 12.3 billion USD in its green initiative by 2027, a subset of which involves decarbonization. Under a 2 billion USD investment, the proposed methodology reduces the carbon emissions which arise from OCP's electricity needs by 95% while generating a net present value of 200 million USD over 20 years. The methodology could be applied to decarbonize other industrial consumers in a financially sustainable fashion, which is imperative in the developing world.

References

[1] D. Bertsimas, R. Cory-Wright, V. Digalakis Jr, Decarbonizing OCP, *arXiv preprint arXiv:2209.06341*, 2022.

Runtime Analyses of Multi-Objective Evolutionary Algorithms in the Presence of Noise

Matthieu Dinot[1] Benjamin Doerr[2] Ulysse Hennebelle[1]
Sebastian Will[2]

[1]École Polytechnique, IP Paris, Palaiseau, France, firstname.lastname@polytechnique.edu
[2]LIX, École Polytechnique, IP Paris, Palaiseau, France, firstname.lastname@polytechnique.edu

Keywords: Multi-objective optimization, evolutionary algorithms, runtime analysis, noisy optimization.

In single-objective optimization, it is well known that evolutionary algorithms also without further adjustments can stand a certain amount of noise in the evaluation of the objective function. In contrast, this question is not at all understood for multi-objective optimization.

In this work [1], we conduct the first mathematical runtime analysis of a simple multi-objective evolutionary algorithm (MOEA) on a classic benchmark in the presence of noise in the objective function. We prove that when bit-wise prior noise with rate $p \leq \alpha/n$, α a suitable constant, is present, the *simple evolutionary multi-objective optimizer* (SEMO) without any adjustments to cope with noise finds the Pareto front of the OneMinMax benchmark in time $O(n^2 \log n)$, just as in the case without noise. Our proofs suggest that the strong robustness of the MOEA stems from its implicit diversity mechanism designed to enable it to compute a population covering the whole Pareto front.

Interestingly this result only holds when the objective value of a solution is determined only once and the algorithm from that point on works with this, possibly noisy, objective value. We prove that when all solutions are reevaluated in each iteration, then any noise rate $p = \omega(\log(n)/n^2)$ leads to a super-polynomial runtime. This is very different from single-objective optimization, where it is generally preferred to reevaluate solutions whenever their fitness is important and where examples are known such that not reevaluating solutions can lead to catastrophic performance losses.

This work is part of the PGMO-funded project *Mathematical Analysis of Complex Randomized Search Heuristics* (PI: Benjamin Doerr).

References

[1] M. Dinot, B. Doerr, U. Hennebelle, and S. Will. Runtime analyses of multi-objective evolutionary algorithms in the presence of noise. In *International Joint Conference on Artificial Intelligence, IJCAI 2023*, pages 5549–5557. ijcai.org, 2023.

An Improved Tie-Breaking Rule for the Non-Dominated Sorting Genetic Algorithm II

Benjamin Doerr[1] Tudor Ivan[2] Martin S. Krejca[1]

[1]LIX, École Polytechnique, IP Paris, France, firstname.lastname@polytechnique.edu
[2]École Polytechnique, IP Paris, France, tudor.ivan@polytechnique.edu

Keywords: Multi-objective optimization, evolutionary algorithms, runtime analysis.

The *non-dominated sorting genetic algorithm II* (NSGA-II) [1] is the most intensively used multi-objective evolutionary algorithm (MOEA) in real-world applications. While theoretical research on the NSGA-II has only started very recently [3], this first work was quickly followed-up by around a dozen research papers. The lower bounds proven in [2] show that the expected run time of the NSGA-II often scales linearly with the population size N of the algorithm. Consequently, different from many other evolutionary algorithms, the NSGA-II does not profit from larger population sizes (as long as the population size is large enough for the NSGA-II to be effective). This shows that the choice of the right parameter value is particularly critical for the NSGA-II.

Analyzing the lower bound proofs, we find that this undesirable scaling is caused by the way the NSGA-II breaks ties between solutions of the same quality (in the sense of non-dominated sorting and crowding distance). We propose a simple modification of the tie-breaking rule of the NSGA-II that takes into account the number of individuals with same objective value. We investigate the performance of the NSGA-II with this *balanced* tie breaking on the two benchmark functions OneMinMax and OneJumpZeroJump. We prove upper bounds that remove the dependence of the expected run time on the population size N for decent ranges of this parameter. Hence in these regimes, the choice of the population size is uncritical.

This work is part of the PGMO-funded project *Mathematical Analysis of Complex Randomized Search Heuristics* (PI: Benjamin Doerr).

References

[1] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6:182–197, 2002.

[2] B. Doerr and Z. Qu. From understanding the population dynamics of the NSGA-II to the first proven lower bounds. In *Conference on Artificial Intelligence, AAAI 2023*, pages 12408–12416. AAAI Press, 2023.

[3] W. Zheng, Y. Liu, and B. Doerr. A first mathematical runtime analysis of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II). In *Conference on Artificial Intelligence, AAAI 2022*, pages 10408–10416. AAAI Press, 2022.

Some statistical insights into Physics-Informed Neural Networks

Nathan Doumèche[1,2] Claire Boyer[2] Gérard Biau[2]

[1]EDF R&D, France, nathan.doumeche@edf.fr

[2]Sorbonne University, France, nathan.doumeche@sorbonne-universite.fr, claire.boyer@sorbonne-universite.fr, gerard.biau@sorbonne-universite.fr

Keywords: Physics-informed neural networks, PDE solving, hybrid modeling.

Physics-informed machine learning consists in training neural networks by using an empirical risk which is penalized by a system of differential equations. These models combine the expressiveness of neural networks with the interpretability provided by physical modeling. Their good practical performance has been illustrated both in the context of partial differential equations solving, and in the context of hybrid modeling, which consists in combining an imperfect physical model with noisy observations. However, their theoretical properties have yet to be established. In this paper, we show that the classical training of these networks suffers from systematic overfitting. We then show that adding a ridge regularization to their empirical risk ensures the consistency of the resulting estimator. However, although the convergence of the theoretical risk of the estimator is guaranteed, its convergence to a solution verifying the regularized physical constraint requires a more detailed analysis. We show how adding a Sobolev penalty to the risk makes it possible and comment this result with numerical illustrations.

References

[1] N. Doumèche, C. Boyer, and G. Biau, Convergence and error analysis of PINNs, *arXiv:2305.01240*, 2023.

[2] G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-informed machine learning, *Nature Reviews Physics*, 422-440(3), 2021.

[3] A. Arzani, J.-X. Wang, and R.M. D'Souza, Uncovering near-wall blood flow from sparse data with physics-informed neural networks, *Physics of Fluids*, 071905(33), 2021.

Existence and Lipschitz regularity of the trajectories minimizing the total variation in a congested setting

Annette Dumas[2] Filippo Santambrogio[2]

[1]University of Lyon 1 (ICJ), France, dumas@math.univ-lyon1.fr
[2]University of Lyon 1 (ICJ), France, santambrogio@math.univ-lyon1.fr

Keywords: Mean Field Games, piecewise constant, optimal trajectories, regularity

The problem I will present is motivated by the study of a Mean Field Game model whose theory was simultaneously introduced by Lasry and Lions ([1, 2]) and Caines, Huang and Malhamé ([3]) in 2006. The model consists in studying a population in a city where each agent jumps to move from one place to another. Each inhabitant minimizes a cost composed of the number of jumps and an increasing function of the density of the population. The solution to this problem is a probability measure on the trajectories which is a Nash equilibrium.

The probability measure Q on the trajectories can be seen as a trajectory of the density of the population $\rho_t = e_t \sharp Q$ which leads us to the following variational problem:

$$\min_{\substack{\rho\in E,\rho\geq 0\\ \forall t\in[0,T],\ \int_\Omega \rho(t,x)dx=1}} \int_0^T \int_\Omega \left(|\partial_t \rho(t,x)| + V(t,x)\rho(t,x) + f(\rho(t,x))\right) dxdt + \psi_0(\rho(0)) + \psi_T(\rho(T)),$$

where $E = \mathrm{BV}([0,T], L^1(\Omega)) \cap L^2([0,T]\times\Omega)$. The number of jumps is associated with the total variation of the density and the additional cost is associated with the increasing function of the density. Density constraints are also added to the problem. We will see that the solution exists, is unique and is Lipschitz in time, despite the discontinuous trajectories taken by each agent. With additional hypothesis on the data, boundedness or continuity in space can be obtained with Dirichlet conditions in time.

The aspect of the solutions are given by the Euler-Lagrange equations which show that in space, either the solution is constant, or it follows the critical points of the cost. Numerical simulations are carried out on a simple example by using the Fast Dual Proximal Gradient method which validates the theoretical framework.

References

[1] J.-M. Lasry, P.-L. Lions, *Jeux à champ moyen. I - Le cas stationnaire*, C. R. Acad. Sci. Paris, Ser. I 343 (2006)

[2] J.-M. Lasry, P.-L. Lions, *Jeux à champ moyen. II - Horizon fini et contrôle optimal,* C. R. Acad. Sci. Paris, Ser. I 343 (2006)

[3] M. Huang, R. P. Malhamé and P. E. Caines, *Large Population Stochastic Dynamic Games: Closed-Loop MckeanVlasov Systems and the Nash Certainty Equivalence Principle*, Communications in Information and Systems, Vol. 6, No. 3, 2006, pp. 221-252.

Optimal Price Menu Design of Electric Vehicle Charging Stations

Alix Dupont[1,2] Panagiotis Andrianesis[3] Yezekael Hayel[1]
Tania Jiménez[1]
Jean-Baptiste Breal[2] Raphael Payen[2]

[1] Avignon University, Computer Science Lab, France
[2] EDF, R&D Lab, Palaiseau, France
[3] Technical University of Denmark, Department of Wind and Energy Systems, Lyngby, Denmark

Keywords: Electric vehicles, charging station, price design, optimization

The growing adoption of Electric Vehicles (EVs) presents many challenges and opportunities for the Charging Station Operators (CSOs), who are called to determine their pricing strategies considering the EV user charging behavior. Related literature, such as [1, 2], considers a CSO who sets prices for different delays in the time required to charge. However, EV users are supposed to give personal information such as their departure time, which also constrains them to leave at the announced departure time (with potential overstay penalties).

In this work, we consider that the EV users arriving at a charging station view charging more as an opportunity rather than as an indispensable need; hence, the parking duration depends only on their on-site activities. Arguably, this behavior is likely to be more common in the future (in the presence of high EV penetration and high charging point availability). In such settings, an appropriate CSO "price menu" proposal would consist of different prices for different levels of charging rates (unlike cases where prices are associated with different delay completion times).

We formulate the proposed price menu design problem as a bi-level optimization problem. At the lower level, the EV users decide to choose one of the options in the menu or they may choose not to charge. Their choice depends on their utility and the price per charging rate offered in the menu. At the upper level, the CSO designs the prices per charging rates. We provide analytical and numerical results on the optimal price menu design, considering in the upper level (i) the maximization of the CSO's profit, and (ii) the maximization of the *social welfare*. We investigate the properties of the optimal price menu under both *complete information* of the EV users utility, and *incomplete information*, where the CSO has only knowledge on the distribution of the utility across the EV users population. In the latter, the bi-level optimization problem yields to a Mixed Integer Linear Programming (MILP) in both the CSO's profit and *social welfare* maximization cases.

References

[1] C. Lu, J. Wu, J. Cui, Y. Xu, C. Wu and M. C. Gonzalez, *Deadline Differentiated Dynamic EV Charging Price Menu Design*, IEEE Transactions on Smart Grid, 14(1):502-516, 2023

[2] O. Wu, S. Yucel , Y. Zhou, *Smart Charging of Electric Vehicles: An Innovative Business Model for Utility Firms*, Manufacturing & Service Operations Management, 24(5):2481-2499, 2021.

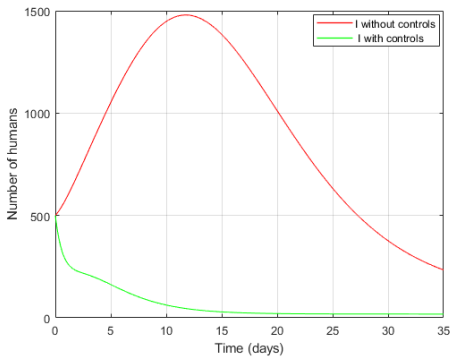
Optimal Control for General Impulsive VS-EIAR Epidemics Models with Application to COVID-19

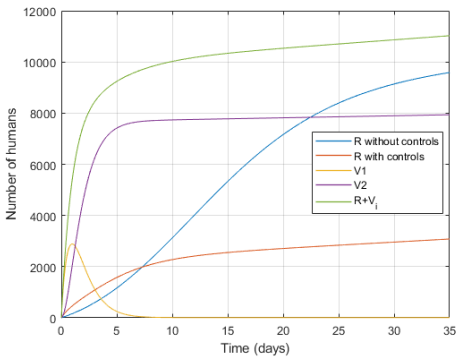
Mohammed Elghandouri[1,2] Khalil Ezzinbi[1,2]

[2]Cadi Ayyad University, FSSM, Departement of Mathematics, B.P: 2390, Marrakesh, Morocco.
[1]Sorbonne University, EDITE (ED130), IRD, UMMISCO, F-93143, Paris, France.
medelghandouri@gmail.com, ezzinbi@uca.ac.ma

Keywords: Optimal Control, epidemic Model, Impulsive Epidemic Model, Mathematical Modeling, Ordinary Differential Equations, COVID-19.

We interested in a VS-EIAR epidemiological model considering vaccinated individuals $\{V_i : i = 1, \cdots, n\}$. The dynamic of the VS-EIAR model involves $n+6$ ordinary differential equations that describe the change of vaccinated, susceptible, infected, exposed, asymptomatic and deaths people groups. Our aim is to reduce the number of suscebtible, exposed, infected, and asymptomatic people by using vaccination doses for suscepbtible individuals and treatment for infected individuals. To do so, we use optimal control theory to control the dynamic of our considered epidemic model in a terminal optimal time τ^*. Pontryagins maximum principle (PMP) will be used to establish the existence of an optimal control time $(v^*(t), u^*(t))$. An impulsive VS-EIAR epidemic model will be taken into account with a especial attention since the immigration or the travel of some people to make our study more applicable. A numerical simulation is given to show how the theoretical study can be applied.





References

[1] Z. Abbasi, I. Zamani, A. H. Amiri Mehra, Mohsen Shafieirad, Asier Ibeas, Optimal Control Design of Impulsive SQEIAR Epidemic Models with Application to COVID-19, Chaos, Solitons & Fractals, Volume 139, 2020, 110054.

A Mean Field Game Model for Renewable Investment under Long-Term Uncertainty and Risk Aversion

Celia Escribe[1,2] Josselin Garnier[2] Emmanuel Gobet[2]

[1]CIRED-CNRS, France, celia.escribe@polytechnique.edu
[2]CMAP, CNRS, Ecole Polytechnique, France

Keywords: Stochastic control, Mean field games, Nash equilibrium, Renewable energy, Electricity markets

We consider a stylized model for investment into renewable power plants under long-term uncertainty. We model risk-averse agents facing heterogeneous weather conditions and a common noise including uncertainty on demand trends, future fuel prices and the average national weather conditions. The objective of each agent is to maximize multistage profit by controlling investment in discrete time steps. We analyze this model in a noncooperative game setting with N players, where the interaction among agents occurs through the spot price mechanism. Our model extends to a mean field game with common noise when the number of agents is infinite. We prove that the N-player game admits a Nash equilibrium. Moreover, we prove that under proper assumptions, any sequence of Nash equilibria to the N-player game converges to the unique solution of the MFG game. Finally, our numerical experiments highlight the impact of the risk aversion parameter and emphasize the difference between our model which captures heterogeneity and representative agent models.

References

[1] Rene Aid, Roxana Dumitrescu, and Peter Tankov. The entry and exit game in the electricity markets: A mean-field game approach. arXiv:2004.14057 [math], April 2020.

[2] Frederic J. Bonnans, Pierre Lavigne, and Laurent Pfeiffer. Discrete-time mean field games with risk-averse agents. *ESAIM: Control, Optimisation and Calculus of Variations*, 27: 44, 2021a. ISSN 1292-8119, 1262-3377.

The ℓ_0 Bregman Relaxation

M'hamed Essafri[1] Luca Calatroni[2] Emmanuel Soubies[1]

[1]Université de Toulouse, IRIT, CNRS, France. (mhamed.essafri@irit.fr, emmanuel.soubies@cnrs.fr)
[2]CNRS, Université Côte d'Azur, Laboratoire i3S, France. (calatroni@i3s.unice.fr),

Keywords: ℓ_0 regularization, non-quadratic data terms, linear inverse problems, non-convex optimization, exact continuous relaxations.

Sparse optimization approaches have attracted extensive attention due to their practical applications in various fields such as statistics, computer vision, signal/image processing, and machine learning. A typical formulation is given by

$$\hat{\mathbf{x}} \in \left\{ \arg \min_{\mathbf{x}\in\mathbb{R}^N} G_0(\mathbf{x}) := F_{\mathbf{y}}(\mathbf{A}\mathbf{x}) + \lambda \|\mathbf{x}\|_0 \right\}, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{M\times N}$ (with $M \ll N$) is a linear operator, $\mathbf{y} \in \mathbb{R}^M$ denotes the observed data, and the term $\|\cdot\|_0$ is the ℓ_0 pseudo-norm on $\mathbb{R}^N$ that counts the number of non-zero components of its arguments. The function $F_{\mathbf{y}} : \mathbb{R}^M \to \mathbb{R}_{\geq 0}$ is a measure of fit between $\mathbf{A}\mathbf{x}$ and $\mathbf{y}$, and $\lambda > 0$ controls the trade-off between data fidelity and sparsity. Due to the nature of the ℓ_0 function, problem (1) is non-continuous, non-convex, and belongs to the class of NP-hard problems [1].

In this talk, I will introduce the ℓ_0 Bregman Relaxation for (1), which is defined as a non-convex but continuous function providing a good approximation to the ℓ_0 term. First, I will present results on the existence of global minimizers. Then, I will show that replacing the ℓ_0 term in (1) with the ℓ_0 Bregman Relaxation leads to an exact relaxation, denoted $\tilde{G}$, of G_0. This generalizes existing work [2, 3] dedicated to the least squares case. The relaxation $\tilde{G}$ is said to be exact as it "reduces" the non-convexity of (1) (e.g., less local minimizers, wider basins of attraction) while preserving its global solution(s). Finally, a tailored Bregman proximal gradient algorithm is used for the numerical solution of the relaxed problem. Several numerical results for problems involving widely-used non-quadratic data terms (Kullback-Leibler divergence, logistic regression) will be shown.

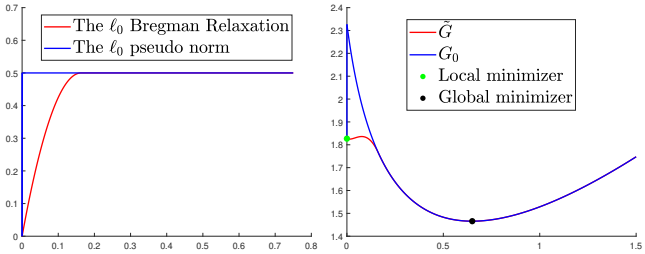


Figure 1: Illustration of the ℓ_0 Bregman Relaxation for 1D Kullback-Leibler sparse regression on $\mathbb{R}_{\geq 0}$.

References

[1] T. T. Nguyen, C. Soussen, J. Idier, and E-H. Djermoune. NP-hardness of ℓ_0 minimization problems: revision and extension to the non-negative setting. Proceedings of SAMPTA, Bordeaux, 2019.
[2] Marcus Carlsson. On Convex Envelopes and Regularization of Non-convex Functionals Without Moving Global Minima. *Journal of Optimization Theory and Applications*, 183(1), 66–84, 2019.
[3] E. Soubies, L. Blanc-Féraud, and G. Aubert. A continuous exact ℓ_0 penalty (CEL0) for least squares regularized problem. *SIAM Journal on Imaging Sciences*, 8(3), 2015.

Monitoring the Convergence Speed of PDHG to Find Better Primal and Dual Step Sizes

Olivier Fercoq[1]

[1]LTCI, Télécom Paris, Institut Polytechnique de Paris, France, olivier.fercoq@telecom-paris.fr

Keywords: Primal-Dual algorithm; adaptive step sizes; power iterations

Primal-dual algorithms for the resolution of convex-concave saddle point problems usually come with one or several step size parameters. Within the range where convergence is guaranteed, choosing well the step size can make the difference between a slow or a fast algorithm. A usual way to set step sizes is to ensure that there is a fair balance between primal and dual variable's amount of change. This can be done adaptively [GLY15] and gives a much better performance than fixed step sizes.

In this work, we show that on quadratic problems, it is possible to find even better step sizes. We base our method on a spectral radius estimation procedure and try to minimize this spectral radius, which is directly related to the rate of convergence. Building on power iterations [Hag21], we could produce spectral radius estimates that are always smaller than 1 and work also in the case of conjugate principal eigenvalues.

For strongly convex quadratics, we show that our step size rules yields an algorithm as fast as inertial gradient descent. Moreover, since our spectral radius estimates only rely on residual norms, our method can be readily adapted to more general convex-concave saddle point problems.

References

[GLY15] Tom Goldstein, Min Li, and Xiaoming Yuan. Adaptive primal-dual splitting methods for statistical learning and image processing. *Advances in neural information processing systems*, 28, 2015.

[Hag21] William W Hager. *Applied numerical linear algebra*. SIAM, 2021.

A new version of the Kurdyka-Łojasiewicz property to better study the convergence of stochastic optimization algorithms in a non-convex framework.

Jean-Baptiste Fest[1] Emilie Chouzenoux[1] Audrey Repetti [2]

[1] CVN, CentraleSupélec, Inria, Université Paris-Saclay, France,

[2] School of Engineering & Physical Sciences, School of Mathematical & Computer Sciences, UK.

Corresponding author: jean-baptiste.fest@inria.fr

Keywords: Stochastic algorithms, Non-convexity, Kurdyka-Łojasiewicz inequality.

Typical well-built first order stochastic schemes satisfy desirable properties, including conditional descent inequalities and almost sure (a.s.) convergence guarantees on the objective function F, as well as on the involved gradient [4, 3]. In the convex framework, the geometry of F often allows technical shortcuts, leading to the a.s. convergence of the algorithm iterates to a global minimiser. By contrast, for non-convex objective functions, such type of convergence cannot be easily guaranteed. We introduced in [2] a new avenue of work to bridge the gap between descent-type inequalities and a.s. convergence of the associated stochastic process despite the lack of convexity of F. To do so, we build upon the strategy of [1] which initially consists in using the Kurdyka-Lojasiewicz (KL) property to prove the finite length of a discrete deterministic process (without any convexity assumptions). However, the stochastic setting does not allow the use the KL theory as such. As a response to this challenge, we propose an alternative version of the generic KL inequality, so as (i) to be consistent with the tools involved in usual proofs of the stochastic approximation theory and (ii) to promote the existence of a.s. convergence results for the iterates.

References

[1] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, 146(1-2):459–494, 2014.

[2] E. Chouzenoux, J.-B. Fest, and A. Repetti. A Kurdyka-Lojasiewicz property for stochastic optimization algorithms in a non-convex setting. Technical report, 2023. https://arxiv.org/abs/2302.06447.

[3] M. Duflo. *Random iterative models*, volume 34. Springer Science & Business Media, 2013.

[4] H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications. In *Optimizing Methods in Statistics*, pages 233–257. Elsevier, 1971.

Hybrid games in route planning for sailing vessels and their mean field limit

Simone Cacace[1] Roberto Ferretti[2] Adriano Festa[3]

[1]Dipartimento di Matematica Guido Castelnuovo, Sapienza UniversitÃ di Roma,
L.go S. Leonardo Murialdo, 1, 00146 Roma, Italy. simone.cacace@uniroma1.it
[2]Dipartimento di Matematica e Fisica, Università Roma Tre,
L.go S. Leonardo Murialdo, 1, 00146 Roma, Italy. roberto.ferretti@uniroma3.it
[3]Dipartimento di Scienze Matematiche Giuseppe Luigi Lagrange, Politecnico di Torino,
Corso Duca degli Abruzzi, 24, 10129 Torino TO adriano.festa@polito.it

Keywords: Hybrid systems, Route planning, Non-zero sum games, Mean Field Games.

In the last decades, the sport of sailing has experienced an increasing impact of new technologies, and notably of scientific computing. Among all computational problems relevant for sailing, we focus here on *route planning* and *race strategy*, i.e., the optimization of the yacht route.

In the most typical and basic form, the route planning problem requires reaching a windward mark in minimum time within a variable wind field. According to the previous literature and experimental evidence, this problem is approached from a stochastic viewpoint, in which the wind field is modeled as having both a deterministic and a stochastic component, since even when we have a reliable forecast of its evolution, the wind field is typically affected by random fluctuations that represent a crucial part of the problem.

We discuss the extension of the classic dynamic programming approach to hybrid control systems, where a continuous controller can switch between a finite collection of states, paying a cost of switching. This approach requires the resolution of a system of quasi-variational Hamilton-Jacobi inequalities that we propose to approximate via a semi-Lagrangian scheme obtained by the direct discretization of the dynamical programming principle. We discuss the application of such a framework to model a sailing boat navigation problem for the optimization of the strategic choices on a race course [3].

The model can be extended to the non-zero sum games, naturally including the presence of 3 or more players [1, 2]. For the number of players going to infinity, it is possible to introduce a new model as its mean-field limit. This is the case of very crowded races, where the interaction between many players would be impossible to model correctly with other techniques.

References

[1] A. Bressan and F.S. Priuli, Infinite horizon noncooperative differential games, *J. Differ. Equations*, **227** (2006), 230–257

[2] S. Cacace, R. Ferretti and A. Festa, Hybrid differential games and their application to a match race problem, *Appl. Math. Comp.*, **372** (2020), 124966

[3] R. Ferretti and A. Festa, Optimal route planning for sailing boats: a hybrid formulation, *J. of Optim. Theory and Applications*, **181** (2019), 1015–1032

Accommodating fairness in a shared-energy allocation problem with uncertainties

Zoé Fornier[1] Vincent Leclère[2] Pierre Pinson[3]

[1]METRON, France, zoe.fornier@metronlab.com
[2]Ecole Nationale des Ponts et Chaussées, France, vincent.leclere@enpc.fr
[2]Imperial College of London, UK, p.pinson@imperial.ac.uk

Keywords: Energy Aggregator, Fairness, Uncertainties, Stochastic Optimization

With the recent emergence of prosumers (who are both energy producers and consumers), some companies propose to aggregate prosumers in the energy market. This is because prosumers are generally too small to directly enter the electricity market themselves. To implement those aggregators, there is a practical need to study how those who compose the portfolio, based on their assets and production/consumption profiles, are going to be treated fairly. We introduce and discuss ways to accommodate fairness by design in the modeling of the problem. Following this paradigm, the allocation decided by the aggregator has to be fair at the cost of an objective deterioration. This problem translates to a shared-resource fair allocation problem and can be modeled as a mixed-integer linear program in the deterministic case. First, we revisit conventional approaches to accommodate fairness including, e.g., the minimax approach, favoring the least well-off agent and the proportional approach, derived from Nash's bargaining problem. We also examine measuring the proportional savings of agents instead of their costs. There, we minimize the agent with least-proportional savings, or we add equal proportional savings constraints. Those different approaches are compared based on application case studies of limited size, yet respecting realistic features of assets to be aggregated in electricity markets. Importantly, we eventually take into consideration the uncertainties in the energy market: in prices or energy loads. Accommodating fairness in stochastic resource allocation problems has been less studied in the literature. As a basis, we extend the previous approaches and related investigations to a two-stage stochastic program, by rethinking the previous fairness constraints or objectives in an uncertain case. Using the same case studies, we finally discuss the specifics of the stochastic setup compared to the deterministic one.

Large Scale Analysis of a Stochastic Model for Free-Floating Car-Sharing

Christine Fricker[1] Hanene Mohamed[2] Alessia Rigonat[1]

[1]DI ENS, INRIA Paris, France, christine.fricker@inria.fr, alessia.rigonat@inria.fr
[2]Université Paris Nanterre, France, hanene.mohamed@parisnanterre.fr

Keywords: Mean-field, Multi-Scale Markov Processes, Car-Sharing , Fleet Optimization Problem

Bike-sharing and car-sharing systems are an ecological alternative for urban transportation. In particular in car-sharing systems, free-floating is becoming increasingly popular, but the imbalance problem shows the need for stochastic modelling and analysis. We propose a new stochastic model suited for free-floating. Indeed, it is commonly admitted that a free-floating system can be modelled as a station-based system, by dividing the service area into small geographical zones. However, in a station-based system, the stations have a fixed capacity. In contrast, the parking spaces available in free-floating car-sharing systems are those left by other users of the public space. The dynamics of car-sharing systems interact with those of private cars enlightening two different time scales. Assuming that private cars are more numerous, we have a fast process for the number of private cars in each zone and a slow process for the number of shared cars. A stochastic averaging principle governs the free-floating cars behavior and must be managed in the mean-field analysis. A coupling argument under an additional technical assumption is used. A phase transition with explicit threshold is exhibited between a non-saturated regime where free-floating cars can always be parked and a saturated one where free-floating cars do not find an available parking space with positive probability. This probability is completely determined by the more numerous private cars which implies that the operator cannot act on the proportion of saturated areas, those with no parking spaces. This allows to solve the dimensioning problem. The solution is completely different in this new model, compared to the station-based case: the more free-floating cars there are in the system, the more satisfied the users are. See [1] for details.

References

[1] C. Fricker, H. Mohamed, A. Rigonat, M. Trépanier. A new stochastic model for carsharing suited to free-floating. 2023. *⟨hal-03938964⟩*

A Model for Dimensioning Ressources in a Car-Sharing System with Booking in Advance and Cancellation

Christine Fricker[1] Florian Verdier[2]

[1]DI-ENS, Inria de Paris, France, christine.fricker.fr

Keywords: reservation in advance, reneging, impatience, fluid limit, infinite-server queue, fleet size optimization

In station-based car-sharing systems where users pick up a car and had to return it in the same station, the parking place is reserved. As to the car, the user can reserve it in advance for a time slot, with the option of cancelling. The issue is to evaluate the number of cars needed to meet the demand in the event of cancellation.

As a first approximation, when the demand is high, we assume that a user will pick up a car wherever there is one available for the time slot. We therefore consider a stochastic model in which C cars (servers) are available to users (customers) arriving at rate λ. When the user arrives, if a car is available for the entire time slot in the future, a reservation is made for that slot. If not, the request is rejected. Before picking up the car, the reservation can be cancelled by releasing the time slot. The user can also return the car before the end of the reservation period. This is equivalent to abandoning the service. It is reasonable to assume that interarrival times, pre-booking times, booking slots, reneging times and abandonment times are respectively i.i.d with a general distribution, and an exponential distribution for inter-arrival times.

The number of reservations is difficult to study when C is finite. If C is infinite, this process, generalizing the number of customers in a $M/G/\infty$ queue can be expressed as the number of points of a Poisson point process in some set of dimension 5. We obtain its fluid limit and fluctuations when the arrival rate becomes large. The fluid limit is stricly increasing converging to an explicit C_{max} when t is large giving the number of This result generalizes a result for $M/G/\infty$ queue by Fricker and Jaibi[]. The proof is based on the theory of Poisson point processes, in particular martingales associated to them, and convergence theorems. If C is finite, simulations show that there is a phase transition between two regimes, a saturated one if $C_{max} = C$ and a non-saturated one otherwise. The long time value of the fluid limit is C_{max} in the non-saturated regime, while in the saturated regime, the fluid limit does not reach C. This long time value is not known and the object of future work. This work was carried out during Florian Verdier's internship at INRIA and Polytechnique Montreal. Paper in preparation.

References

[1] C. Fricker and M. R. Jaïbi, On the fluid limit of the $M/G/\infty$ queue, *Queueing Systems*, 56, 255–265, 2007.

Improving Sparse Penalties with Nonconvexity and Coefficient Clustering

Anne Gagneux[1,2] Mathurin Massias[1] Emmanuel Soubies[2]

[1]Ens Lyon, Inria, LIP, Lyon, France. anne.gagneux@ens-lyon.fr mathurin.massias@inria.fr
[2]CNRS, IRIT, Univ Toulouse, Toulouse, France , emmanuel.soubies@irit.fr

Keywords: sparsity, nonconvexity, clustering, regularization

In this talk, we address the following question: how can we enforce parsimony and structure in a regularized regression model ? Specifically, we investigate strategies that promote not only sparsity but also clustering of correlated features. In fact, we aim to recover these groups *without prior knowledge about the clusters*. For such problems, the Sorted ℓ_1 norm (SLOPE [1], OWL [2]) is already widely used. Yet, SLOPE relies on the ℓ_1 penalty. It therefore shares a major drawback with the Lasso: it is biased and tends to shrink non-null coefficients.

In the same way nonconvex penalties have been introduced as unbiased ℓ_1 alternatives, we deal here with *sorted nonconvex penalties as unbiased SLOPE's counterparts*. Feng and Zhang [3] have already proposed an approach to address a family of sorted nonconvex penalties: they use a "majorize-minimization" method that generates a sequence of easier-to-solve convex surrogate problems.

Instead, we show that for a parameter η small enough in $\text{prox}_{\eta\Psi}$, we can exactly compute -with linear complexity- the prox of these penalties and apply proximal gradient algorithms to solve the penalized regression. We introduce a new writing for the prox of sorted penalties which only involves the scalar prox of the penalty (for which we usually have an explicit expression) and averaged values of parts of the input vector and the regularization sequence.

We experimentally show that using sorted nonconvex penalties do improve both unbiasedness of the solution and automatic features grouping, which validates our initial hypothesis. We also compare our method with the "majorize-minimization" approach and demonstrate a significant gain in both minimization and computation time.

References

[1] Bogdan, Malgorzata, et al, SLOPE-adaptive variable selection via convex optimization, *The annals of applied statistics* 9.3 (2015): 1103.

[2] Zeng, Xiangrong, and Mario AT Figueiredo, The Ordered Weighted ℓ_1 Norm: Atomic Formulation, Projections, and Algorithms, arXiv preprint arXiv:1409.4271 (2014).

[3] Feng, Long, and Cun-Hui Zhang, Sorted concave penalized regression, *The Annals of Statistics* 47.6 (2019).

An Empirical Case of Gaussian Processes Learning in High Dimension: the Likelihood versus Leave-One-Out Rivalry

David Gaudrie[1] Rodolphe Le Riche[2] Tanguy Appriou[1,2]

[1]Stellantis, Carrières-sous-Poissy, France, david.gaudrie@stellantis.com, tanguy.appriou@stellantis.com
[2]CNRS LIMOS, leriche@emse.fr

Keywords: Gaussian processes, surrogate modeling, curse of dimensionality, maximum likelihood

Gaussian Processes (GPs) are semi-parametric models commonly employed in various real-world applications such as statistical modeling and sensitivity analysis. They play an instrumental role in Bayesian optimization since they determine the sequence of iterates where the objective function is evaluated. GPs are particularly useful in the context of small data when there are of the order of a hundred to a thousand points to learn. However, GPs suffer particularly from the curse of dimensionality [2]: at a fixed number of data points, their predictive capability may decrease dramatically in high dimension ($d \gtrsim 40$) [3].

In this talk, we investigate such a phenomenon in details. We illustrate this loss of performance with increasing dimension on simple functions and analyze its underlying symptoms, in particular a tendency to become constant away from the data points. We show that the fundamental problem is one of learning and not one of representation capacity: maximum likelihood, the dominant loss function for such models, can miss regions of optimality of the GP hyperparameters. Failure of maximum likelihood is related to statistical model inadequacy [1]: a model with constant trend is sensitive to dimensionality when fitting quadratic functions while it much better handles dimension growth for linear functions or Gaussian trajectories generated with the right covariance. Our experiments also show that the leave-one-out loss function is less prone to the curse of dimensionality even for inadequate statistical models.

References

[1] F., Bachoc, Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification, *Computational Statistics & Data Analysis*, 66:55–69, 2013.

[2] K. Beyer, J., Goldstein, R., Ramakrishnan, U., Shaft, When Is "Nearest Neighbor" Meaningful?, *Lecture Notes in Computer Science*, vol 1540, https://doi.org/10.1007/3-540-49257-7_15, 1999.

[3] D. Ginsbourger, D., Dupuy, A., Badea, L., Carraro and O., Roustant, A note on the choice and the estimation of kriging models for the analysis of deterministic computer experiments, *Appl. Stochastic Models Bus. Ind.*, 25:115–131, 2009.

A Control Problem with Random State Constraints in Probabilistic and Almost-sure Form

Caroline Geiersbach[1] René Henrion[2]

[1]Weierstrass Institute Berlin, Germany, geiersbach@wias-berlin.de
[2]Weierstrass Institute Berlin, Germany, henrion@wias-berlin.de

Keywords: risk-averse PDE constrained optimization, probabilistic state constraints, almost-sure constraints

We consider a simple PDE-constrained optimization problem under uncertainty governed by a Poisson's equation with Dirichlet boundary conditions. The right-hand side is defined by a control plus a random source term. The resulting random states are supposed to stay within certain limits (uniformly over the domain) with given high probability. We provide an exact formula for the Clarke subdifferential of the associated probability function assigning to each control the probability of satisfying the state constraints. This formula, in particular its reduction in the differentiable case, can be used both for algorithmic purposes and for the derivation of necessary and sufficient optimality conditions. The latter require some additional results on convexity. Numerical results in 1 and 2 D are illustrated. Special emphasis is laid on the consideration of almost-sure constraints (the extreme situation for probability equal to one).

References

[1] C. Geiersbach and R. Henrion, Optimality conditions in control problems with random state constraints in probabilistic or almost-sure form, *arXiv:2306.03965*

Neural Benders Decomposition for Mixed Integer Programming

Shahin Gelareh[1] Rahimeh N Monemi[2]

[1]Universite d'Artois, shahin.gelareh@univ-artois.fr,
[2]Sharkey Predictim Globe, France contact@predictim-globe.com

Combinatorial Optimization: Mixed Integer Programming, Decomposition

In this study, we propose an imitation learning framework designed to enhance the Benders decomposition method. Our primary focus is addressing degeneracy in subproblems with multiple dual optima, among which Magnanti-Wong technique identifies the non-dominant solution. We develop two policies. In the first policy, we replicate the Magnanti-Wong method and learn from each iteration. In the second policy, our objective is to determine a trajectory that expedites the attainment of the final subproblem dual solution. We train and assess these two policies through extensive computational experiments on two real cases, confirming that the presence of such learned policies significantly enhances the efficiency of the decomposition process.

References

[1] Rahmaniani, R., Crainic, T.G., Gendreau, M., Rei, W., 2017. The benders decomposition algorithm: A literature review. European Journal of Operational Research 259, 801–817.

[2] Sherali, H., Fraticelli, B., 2002. A modification of Benders' decomposition algorithm for discrete subproblems: An approach for stochastic programs with integer recourse. Journal of Global Optimization 22, 319–342.

[3] Benders, J., 1962a. Partitioning procedures for solving mixed-variables programming problems. Computational Management Science 2(1), 3–19.

Approximation of Deterministic Mean Field Games

Justina Gianatti[1] Francisco J. Silva[2]

[1]Universidad Nacional de Rosario, Argentina, gianatti@cifasis-conicet.gov.ar
[2] Université de Limoges, France, francisco.silva@unilim.fr

Keywords: Deterministic mean field games, control-affine dynamics, Lagrangian equilibrium, approximation of equilibria, convergence results.

In this talk we study the numerical approximation of deterministic Mean Field Games where the dynamics of a typical agent is non-linear with respect to the state variable and affine with respect to the control variable. Particular instances of the problem considered here are mean field games where the typical player controls its acceleration. Our main result is the convergence of solutions of this approximation towards MFG equilibria. If time permits, we will also discuss the case of deterministic mean field games with state constraints.

Outlier Detection in Regression: Conic Quadratic Formulations

Andrés Gómez[1] José Neto[2]

[1]Daniel J. Epstein Department of Industrial and Systems Engineering
University of Southern California, Los Angeles, CA 90089, USA
gomezand@usc.edu
[2]SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris
91120 Palaiseau, France
jose.neto@telecom-sudparis.eu

Keywords: least trimmed squares, least quantile regression, robust linear regression, outlier detection

In many applications, when building regression models, it is important to account for the presence of outliers, i.e., corrupted input data points resulting, e.g., from recording errors. Problems consisting in simultaneously performing linear regression and discarding outliers may be formulated as mixed-integer optimization problems of the following form which involves cubic terms, each given by the product of a binary variable and a quadratic term of the continuous variables:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{z}} \ & \sum_{i=1}^{m} (y_i - \boldsymbol{a_i}^\top \boldsymbol{x})^2 (1 - z_i) \\ \text{s.t. } & (\boldsymbol{x}, \boldsymbol{z}) \in F \subseteq \mathbb{R}^n \times \{0,1\}^m, \end{aligned} \tag{1}$$

where $(\boldsymbol{a_i}, y_i) \in \mathbb{R}^{n+1}$ for all $i \in \{1, \ldots, m\}$ are given data and F is the feasible region. In particular, the formulation (1) includes as a special case the *least trimmed squares* problem which is NP-hard [1, 2], and also hard to approximate [3].

Existing approaches in the literature typically rely on the linearization of the cubic terms using big-M constraints. However the latter lead to weak relaxations and poor performance in practice. In this work, we derive stronger second-order conic relaxations not involving big-M constraints. Our computational experiments indicate that the proposed formulations are several orders of magnitude faster than existing big-M formulations in the literature for this problem.

References

[1] T. Bernholt, Robust estimators are hard to compute. Technical Report, No. 2005,52. University of Dortmund, 2005.

[2] T. Bernholt, Computing the Least Median of Squares Estimator in Time $O(n^d)$. Computational Science and Its Applications – ICCSA 2005. Springer Berlin Heidelberg. 697–706, 2005.

[3] D.M. Mount, N.S. Netanyahu, C.D. Piatko, R. Silverman and A.Y. Wu, On the Least Trimmed Squares Estimator. Algorithmica 69:148-183, 2014.

Provable non-accelerations of the heavy-ball method

Baptiste Goujaud Adrien Taylor Aymeric Dieuleveut

Ecole Polytechnique, Institut Polytechnique de PariS, France, baptiste.goujaud@polytechnique.edu
Inria, France, adrien.taylor@inria.fr
Ecole Polytechnique, Institut Polytechnique de Paris, France, aymeric.dieuleveut@gmail.com

In this work, we show that the heavy-ball (HB) method provably does not reach an accelerated convergence rate on smooth strongly convex problems. More specifically, we show that for any condition number and any choice of algorithmic parameters, either the worst-case convergence rate of HB on the class of L-smooth and μ-strongly convex quadratic functions is not accelerated (that is, slower than $1 - O(k)$), or there exists an L-smooth μ-strongly convex function and an initialization such that the method does not converge. To the best of our knowledge, this result closes a simple yet open question on one of the most used and iconic first-order optimization technique. Our approach builds on finding functions for which HB fails to converge and instead cycles over finitely many iterates. We analytically describe all parametrizations of HB that exhibit this cycling behavior on a particular cycle shape, whose choice is supported by a systematic and constructive approach to the study of cycling behaviors of first-order methods. We show the robustness of our results to perturbations of the cycle, and extend them to class of functions that also satisfy higher-order regularity conditions. Submitted

Beyond the Fermat Optimality Rules

Malek Abbasi[1] Sorin-Mihai Grad[2] Michel Théra[3]

[1]Department of Mathematics, University of Isfahan, Isfahan, Iran, malekabbasi72@gmail.com
[2]Unité de Mathématiques Appliquées, ENSTA Paris, 91120 Palaiseau, France, sorin-mihai.grad@ensta-paris.fr
[3]Mathematics and Computer Science Department, University of Limoges, michel.thera@unilim.fr

Keywords: Fermat optimality conditions, Moreau-Rockafellar subdifferential, Sup-subdifferential, Outer/upper/inner-semicontinuity, Kuratowski-continuity, Hamel basis

We propose a general framework for analyzing the behavior at its extrema of an extended real-valued function that can lack convexity or differentiability, and for which the classical Fermat rules of optimality do not apply. To this end, we employ the notions of *sup-subdifferential*, recently introduced by two of the authors together with A. Kruger in [1], and *partial sup-subdifferentials*. The sup-subdifferential is a *nonempty* enlargement of the Moreau-Rockafellar subdifferential, satisfying most of its fundamental properties and enjoying certain calculus rules. The partial sup-subdifferentials are obtained by breaking down the sup-subdifferential into *one-dimensional components* through the elements of a Hamel basis and play the same role as the partial derivatives in the Fermat optimality rules.

References

[1] M. Abbasi, A.K. Kruger, M. Théra, Enlargements of the Moreau-Rockafellar subdifferential, *Set-Valued Var Anal* 29(3):701–719, 2021

Using regression splines to approximate mixed integer nonlinear optimization models

Claudia D'Ambrosio[1] Vanesa Guerrero[2] Renan Spencer Trindade[1]

[1]LIX CNRS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, France,
dambrosio@lix.polytechnique.fr, rst@lix.polytechnique.fr
[2]Department of Statistics, Universidad Carlos III de Madrid, Getafe, Spain
vanesa.guerrero@uc3m.es

Keywords: Mixed Integer Nonlinear Programming, Splines, Surrogate models

Complex phenomena can be accurately described by means of data-driven mathematical models. However, being able to integrate these models within a mathematical optimization framework can be, in general, very challenging. In fact, many of these data-driven models are 'black-box', in the sense that they do not have an explicit mathematical formula which describes them. In other cases, even if an explicit expression exists, including it into a mathematical optimization model may make solving the problem computationally intractable. We propose to use a special kind of surrogate models, regression splines, to deal with functions of this kind which appear in Mixed Integer Nonlinear Programming (MINLP) problems. The choice of spline functions is not arbitrary. On one hand, they offer a good compromise between accuracy and complexity. On the other hand, their functional form allows us to exploit separability and approximate general non-convex MINLPs by a more tractable subclass of problems.

Variational inference with λ-exponential families

Thomas Guilmeau[1] Emilie Chouzenoux[1] Víctor Elvira[2]

[1]Université Paris-Saclay, CentraleSupélec, Inria, CVN, Gif-sur-Yvette, France,
thomas.guilmeau@inria.fr, emilie.chouzenoux@inria.fr
[2]School of Mathematics, University of Edinburgh, United Kingdom,
victor.elvira@ed.ac.uk

Keywords: Variational inference, λ-exponential families, generalized subdifferential

Variational inference methods aim at approximating difficult probability distributions by minimizing a statistical divergence over a family of parametric distributions. Exponential families, which include Gaussian distributions, are often chosen as the approximating family in such problems. This is because when the statistical divergence is the Kullback-Leibler divergence, it is possible to describe the optimal solutions using convex analysis techniques. However, it is often necessary to go beyond exponential families. For instance, heavy-tailed distributions cannot be represented as an exponential family.

Recently, λ-exponential families have been proposed to extend exponential families [1]. They include in particular Student and Cauchy distributions. The authors also introduced the theory of λ-duality to study λ-exponential families. This theory is a generalization of the convex duality theory using a suitable non-linear coupling instead of the standard scalar product. This is related to recent works in optimization [2]. Still, not much is known about these families, and there is a lack of results indicating how to use them to solve statistical tasks.

We study variational inference problems over λ-exponential families. To this end, we propose a theoretical framework based on λ-duality and extend the approach of [1]. This enables us to recover known variational inference results for standard exponential families and generalize them. In particular, we calculate generalized subgradients with our non-linear couplings for some functions of interest. This allows us to give novel sufficient optimality conditions for variational inference problems on λ-exponential families. We also propose a new proximal-like algorithm to address these problems. We illustrate numerically the usefulness of these results on problems involving heavy-tailed distributions.

References

[1] T.-K. L. Wong and J. Zhang, Tsallis and Rényi Deformations Linked via a New λ-Duality, *IEEE Transactions on Information Theory*, 68(8):5353-5373, 2022.

[2] A. Le Franc and J.-P. Chancelier and M. De Lara, The Capra-subdifferential of the ℓ_0 pseudonorm, *Optimization*, 1-23, 2022.

Screen & Relax for Sparse Support Identification

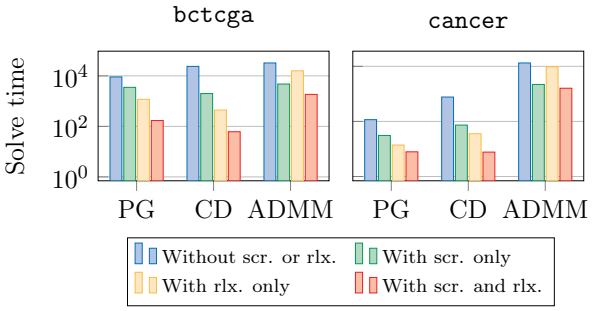
Théo Guyard[1,2] Cédric Herzet[1] Clément Elvira[3] Ayşe-Nur Arslan[4]

[1]Inria, Centre de l'Université de Rennes, France
[2]Insa Rennes, IRMAR UMR CNRS 6625, France
[3]CentraleSupélec Rennes, IETR UMR CNRS 6164, France
[4]Inria, Centre de l'Université de Bordeaux, France

Keywords: Sparse problems, screening tests, relaxing tests.

Sparse representations arise in many application fields such as machine learning and signal processing. They usually result from the resolution of some convex optimization problem of the form $\mathbf{x}^\star \in \arg\min_{\mathbf{x}} \; f(\mathbf{x}) + \lambda g(\mathbf{x})$, where $f(\cdot)$ is a loss function, $g(\cdot)$ is a sparsity-inducing regularization [1] and $\lambda > 0$ is a tradeoff parameter. The latter problem can be addressed with first-order optimization methods such as Proximal Gradient (PG), Coordinate Descent (CD) or Alternating Direction Method of Multipliers (ADMM). A popular strategy to exploit the problem structure is to enhance these algorithms with *screening tests* to detect the position of *zeros* in $\mathbf{x}^\star$. This allows shrinking the problem dimensionality and saving computational resources during the resolution [2]. We propose a complementary strategy, dubbed *relaxing*, to detect the position of *non-zeros* in $\mathbf{x}^\star$. We then leverage this knowledge to perform second-order iterations on the latter subset of entries, thus expecting improvements in the convergence rate.

The adjacent figure shows the improvements allowed by screening and relaxing on two large-scale datasets, namely `bctgca` and `cancer`, drawn from the `HDRM` database. We set $f(\cdot)$ as a Least-squares and Logistic loss, respectively, and $g(\cdot)$ as an Elastic-net regularization. For each dataset, we solve the corresponding problem for 20 logarithmically spaced values of λ to span different working regimes and we report the overall solution time. We observe an acceleration factor up to three orders of magnitude when implementing both screening and relaxing tests compared to a vanilla implementation of PG, CD and ADMM, which assesses the interest for the presented "Screen & Relax" strategy.



References

[1] Shen, L. et al. (2019). Structured sparsity promoting functions. *Journal of Optimization Theory and Applications*, 183, 386-421.

[2] Xiang, Z. J. et al. (2016). Screening tests for lasso problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(5), 1008-1027.

Safe Peeling for L0-Penalized Problems

Théo Guyard[1,2] Gilles Monnoyer[3] Cédric Herzet[1] Clément Elvira[4] Ayşe-Nur Arslan[5]

[1]Inria, Centre de l'Université de Rennes, France
[2]Insa Rennes, IRMAR UMR CNRS 6625, France
[3]UCLouvain, ELEN ICTEAM, Belgique
[4]CentraleSupélec Rennes, IETR UMR CNRS 6164, France
[5]Inria, Centre de l'Université de Bordeaux, France

Keywords: Sparse problems, ℓ_0-norm, mixed-integer programming, bound tightening.

In many application fields such as machine learning, signal processing or inverse problems, it is of paramount interest to solve optimization problems of the form

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + \lambda\|\mathbf{x}\|_0 \tag{1}$$

where $f(\cdot)$ is some loss function, $\|\cdot\|_0$ is the so-called ℓ_0-norm that promotes sparsity in the optimizers by counting the number of non-zeros in its argument and $\lambda > 0$ is an hyperparameter [1]. This problem can be solved effectively via Branch-and-Bound (BnB) algorithms by adding a bound constraint $\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$ to reformulate the ℓ_0-norm using a binary variable through Big-M relations [2]. In practice, the choice of the values in the bounds $\mathbf{l}$ and $\mathbf{u}$ must respect two conflicting imperatives. On the one hand, the additional constraint must not modify the solutions of the initial problem. Hence, the interval $[\mathbf{l}, \mathbf{u}] \subseteq \mathbb{R}^n$ should be chosen wide enough to contain all the optimizers of (1). On the other hand, the quality of the relaxations constructed across the BnB tree degrades with the spread of $[\mathbf{l}, \mathbf{u}]$. Hence, setting a too-wide interval is likely to increase the number of nodes explored during the BnB algorithm and penalize the solution time.

We propose a new strategy, dubbed *peeling*, to address this issue. By leveraging a duality property linking consecutive nodes, we locally tighten the interval $[\mathbf{l}, \mathbf{u}]$ which results in a strengthening of the relaxations constructed. This procedure is carried out in a *safe* manner, *i.e.*, we guarantee the validity of the BnB procedure and the optimality of the solutions obtained. Peeling can be seen as a generalization of the *screening* methodology recently introduced for ℓ_0-penalized problems [3]. Through different numerical experiments, we show that peeling allows significant gains in the number of nodes explored and in the solution time.

References

[1] Bertsimas, D. et al. (2016). Best subset selection via a modern optimization lens. *The Annals of Statistics*, 44(2), 813-852.

[2] Ben Mhenni, R. et al. (2022). Global optimization for sparse solution of least squares problems. *Optimization Methods and Software*, 37(5), 1740-1769.

[3] Atamturk, A. et al. (2020). Safe screening rules for L0-regression from perspective relaxations. *International Conference on Machine Learning*, PMLR, 421-430.

Control Strategies for Transport Networks under Demand Uncertainty with extensions to Bilevel Optimization

Simone Göttlich[1] Patrick Mehlitz[2] Thomas Schillinger[3]

[1]University of Mannheim, Germany, goettlich@uni-mannheim.de
[2]Brandenburg University of Technology Cottbus–Senftenberg, Germany, mehlitz@b-tu.de
[3]University of Mannheim, Germany, schillinger@uni-mannheim.de

Keywords: Optimal control, stochastic processes, transport equations

We consider transport networks with uncertain demands at the consumer nodes. Network dynamics are given by hyperbolic partial differential equations with linear flux functions and (non-)linear damping terms. At the network junctions suitable coupling conditions are discussed. Demands can be incorporated deterministically or stochastically. One way to model demand uncertainty is to use solutions to suitable stochastic differential equations, as for example Jacobi or Ornstein-Uhlenbeck processes. For the demand satisfaction, we solve a constrained optimal control problem. Controls in terms of network inputs are then calculated explicitly for different assumptions.

Extending the model, these explicit inflow controls can be used to set up a bilevel optimization framework. The inner problem consists of solving the aforementioned control problem, whereas the outer problem is concerned with reconstructing desired demand levels from past data. Existence of controls in Lebesgue spaces is ensured when interpreting the hyperbolic conservation law in the weak sense.

Some numerical experiments are presented to visualize various features of the model including potential updating strategies in the control problem and strategies of how to observe the network in order to obtain good reconstructions of the desired demand in the bilevel framework.

References

[1] S. Göttlich and T. Schillinger, Control strategies for transport networks under demand uncertainty, *Advances in Computational Mathematics*, 48:1-33, 2022.

[2] S. Göttlich, P. Mehlitz, T. Schillinger, Inverse demand tracking in transportation networks, *arXiv preprint*, 2023.

A new interior point solver for HiGHS

Julian Hall[1] Filippo Zanetti[2]

[1]University of Edinburgh, United Kingdom, jajhall@ed.ac.uk
[2]University of Edinburgh, United Kingdom, F.Zanetti@sms.ed.ac.uk

HiGHS is open-source software for large scale linear optimization, and has established itself as offering the best benchmark performance due to its innovative interior point (IPM) solver for LP, and powerful MIP solver. In particular, the IPM solver has been hailed as a game-changer for the field of open-source energy system modelling. However, for some instances, the IPM solver is still prohibitively expensive. This talk will discuss the development of a new IPM solver for HiGHS.

Keywords: Open-source software, Linear optimization, Interior point method, Sparse numerical linear algebra

Exact Algorithm for Decoupled Vehicle-to-Grid Problem

Jonathan Hornewall Nabil Absi Cecile Rottner Raphael Payen

IP Paris, France, jonhor@kth.se
École des Mines Saint Étienne, France, absi@emse.fr
EDF, France, cecile.rottner@edf.fr,raphael.payen@edf.fr

We study a Vehicle-to-Grid (V2G) smart charging optimization problem. The problem examines the perspective of an aggregator wishing to maximize profits by using the battery of a single electric vehicles to trade energy and ancillary services on a day-ahead market. The problem can be obtained as a relaxation of a more complex and operationally realistic multiple vehicle problem.

Our work focused on the development of an original and exact algorithm for optimally solving the considred problem in polynomial time. We present both mathematically rigorous proofs of optimality, and an in-depth time complexity analysis. The solution approach leverages a variant of forward dynamic programming.

The work was performed as part of an M2 internship in collaboration between EDF and École des Mines Saint-Étienne, and was funded by PGMO. It is part of the PGMO Vehicle-to-Grid project.

Nonlinear Sharp Minimum and the Stability of a Local Minimum on Metric Spaces

Van Ngai Huynh Michel Thera

University of QuyNhon, Viet Nam, ngaivn@yahoo.com
XLIM - Université de Limoges, France, michel.thera@unilim.fr

In this talk we put the emphasis on nonlinear (weak) sharp minimizers with respect to a gauge function. This concept plays an important role in both theoretical and numerical aspects of optimization. In the last part of the contribution we study the stability (in some appropriate sense) of local/global minimizers of an objective function f perturbed to $f + g$ by a function g belonging to a suitable class of Lipschitz functions defined on metric spaces.

Semidefinite Games

Constantin Ickstadt[1] Thorsten Theobald[1] Elias Tsigaridas[2]

[1] Goethe-Universität, Frankfurt, Germany, ickstadt@math.uni-frankfurt.de, theobald@math.uni-frankfurt.de
[2]Sorbonne Université, Paris University, CNRS, and Inria Paris. IMJ-PRG, France, elias.tsigaridas@inria.fr

Keywords: Semidefinite game, Bimatrix game, Semidefinite programming, Dantzig game, Number of Nash equilibria

We introduce and study the class of semidefinite games, which generalizes bimatrix games and finite N-person games, by replacing the simplex of the mixed strategies for each player by a slice of the positive semidefinite cone in the space of real symmetric matrices.

For semidefinite two-player zero-sum games, we show that the optimal strategies can be computed by semidefinite programming. Furthermore, we show that two-player semidefinite zero-sum games are almost equivalent to semidefinite programming, generalizing Dantzig's [1] result on the almost equivalence of bimatrix games and linear programming.

For general two-player semidefinite games, we prove a spectrahedral characterization of the Nash equilibria. Moreover, we give constructions of semidefinite games with many Nash equilibria. In particular, we give a construction of semidefinite games whose number of connected components of Nash equilibria exceeds the long standing best known construction for many Nash equilibria in bimatrix games, which was presented by von Stengel [2] in 1999.

References

[1] G. B. Dantzig. A proof of the equivalence of the programming problem and the game problem. In T. Koopmans, editor, *Activity Analysis of Production and Allocation*, volume 13 of *Cowles Commission Monograph*. John Wiley & Sons, 1951.

[2] B. von Stengel. New maximal numbers of equilibria in bimatrix games. *Disc. Comp. Geom.*, 21:557-568, 1999.

Optimal Allocation for a Fair Distribution of Green Energy in Microgrids

Natalia Jorquera-Bravo [1,2] Sourour Elloumi[1,2] Safia Kedad-Sidhoum[2] Agnès Plateau[2]

[1]Unité de Mathématiques Appliquées, ENSTA Paris, Institut Polytechnique de Paris, France,
natalia.jorquera@ensta-paris.fr, sourour.elloumi@ensta-paris.fr
[2]CEDRIC, Conservatoire National des Arts et Métiers, France,
safia.kedad_sidhoum@cnam.fr, agnes.plateau_alfandari@cnam.fr

Keywords: Energy Management, Fair Allocation, Production Planning

An energy collective is a community of energy system actors who decide to operate in a collaborative manner. In this way they allow to facilitate the production and distribution of renewable energy. Among the various types of small-scale energy collectives, microgrids have gained prominence. In these networks, multiple users share a distributed energy resource (DER) with the aim of establishing a semi-autonomous system that can operate both connected to the main power grid and in isolation.

Since DERs offer more affordable energy, but may not be able to fully meet the demand of all users, they will therefore compete with each other for these cost-effective DER energy resources, particularly during peak demand hours. Consequently, there is a need of considering fairness allocation among smart homes that share the DER in the microgrid. To the best of our knowledge, the closest study addressing this aspect is presented in [1], where fairness is measured by minimizing the discrepancy between the cost assigned to each house and the cost it would have incurred if it had been the only house in the microgrid.

In this work, we investigate a community consisting of different houses, a shared DER, and a common energy storage system. Each house has an energy demand to satisfy over a discrete planning horizon. The demand can be fulfilled either by using the DER, the battery, or by purchasing electricity from the main power grid. Excess energy can be stored in the battery or sold back to the main grid. The objective is to find a supply plan that provides a fair allocation of renewable energy while minimizing the total cost of the microgrid. We formulate the problem as a mixed-integer linear programming model, considering various fairness metrics such as the proportional allocation rule and the min-max fairness. We evaluate the obtained predictive models using real instances with up to 7 houses and a one-day time horizon with 15-minute time intervals. The data used for these instances are sourced from E4C[1] and pertains to a smart building located in France.

References

[1] Di Zhang, Songsong Liu and Lazaros G. Papageorgiou; Fair cost distribution among smart homes with microgrid, Energy Conversion and Management, 80, 498-508(2014)

[1]Interdisciplinary center Energy4Climate of Institut Polytechnique de Paris and École des Ponts ParisTech

Multidimensional analysis for the techno-economic study of the CSP plant

D. Lasluisa[1,2,4] D. Salas[3] N. Mazet[2] D. Aussel[1,2]

[1]PROMES UPR 8521, CNRS, France, nathalie.mazet@univ-perp.fr
[2]UPVD, Université de Perpignan Via Domitia, France, aussel@univ-perp.fr
[3]Universidad O'Higgins, Chile, david.salas@uoh.cl
[4]Universidad de Chile, Chile, dlasluisa@dim.uchile.cl

Keywords: Concentrated solar power plant, Thermochemical storage, SPOT market, Black-box model, Optimal design.

The profitability of a concentrated solar plant (CSP) depends on the type of storage system and its daily operation. In addition, this profitability can be evaluated using different economical criteria such as the minimum time to recover the initial investment, the total profit at the end of the project and others. Given a solar field and a power block, our aim in this work is to determine an optimal design of the storage system for the CSP. We consider two types of storage technology, namely thermochemical reactor and a Two-Tank molten salt system, two different production strategies, four different economical criteria and three price scenarios for electricity market [1, 2]. And for each such combination (called profile) our model determine the optimal design of the CSP. But, contrary to classical economical evaluation in literature, in our approach an optimal design consists of the optimal values of the physical design of the storage system but also the optimal operations of the plant (production, storage, discharge). Thanks to this model and the obtained results we are able to exhibit conditions under which thermochemical storage is more economically pertinent than Two-Tanks storage and also situations ensuring an economical equilibrium of the CSP without subsidy.

References

[1] M. Liu, N. S. Tay, S. Bell, M. Belusko, R. Jacob, G. Will, W. Saman, F. Bruno Review on concentrating solar power plants and new developments in high temperature thermal energy storage technologie, *Renewable and Sustainable Energy Reviews* 53, (2016) 1411-1432.

[2] Salas, D and Tapachès, E and Mazet, N and Aussel, D, Economical optimization of thermochemical storage in concentrated solar power plants via pre-scenario, *Energy Conversion and Management* 174, (2018). 932-954.

Mirror Descent Algorithm in Generalized Convexity

Adrien Le Franc [1] Jean-Philippe Chancelier [2] Michel De Lara[2]

[1]LAAS CNRS, [2]École des Ponts ParisTech

Keywords: generalized convexity, mirror descent, sparse optimization

In this talk, we present algorithmic perspectives based on mathematical tools of generalized Fenchel-Moreau conjugacies induced by *one-sided linear* (OSL) couplings.

We start with a generalization of the notion of Bregman divergence with OSL couplings, a first step before we also extend the mirror descent algorithm with OSL couplings [3]. This contribution is an original new step in the algorithmic toolbox of general convexity, dominated so far by the cutting angle algorithm [1].

We also comment on the possibility to apply our generalized mirror descent algorithm to sparse optimization problems that display Capra-convex properties [2]. In particular, we consider problems displaying the ℓ_0 sparse pseudonorm as an objective function. Numerical perspectives are encouraged by the knowledge of a closed-form expression for the Capra-subdifferential of ℓ_0 [4], which is requested for both the cutting angle and the mirror descent algorithm.

References

[1] A. Rubinov. Abstract convexity and global optimization. Nonconvex Optimization and its Applications, vol.44. Kluwer Academic Publishers, Dordrecht, 2000.

[2] J.-P. Chancelier and M. De Lara. Constant along primal rays conjugacies and the ℓ_0 pseudonorm, 2020. Optimization, 2022, vol. 71, no 2, p. 355-386.

[3] A. Le Franc. Subdifferentiability in convex and stochastic optimization applied to renewable power systems. PhD Thesis, École des Ponts ParisTech, 2021.

[4] A. Le Franc, J.-P. Chancelier and M. De Lara. The Capra-subdifferential of the ℓ_0 pseudonorm. Optimization, 2022.

Minimal Sparsity for Scalable Moment-SOS Relaxations of the AC-OPF Problem

Adrien Le Franc[1] Victor Magron[2]

[1]LAAS CNRS, Toulouse, France, adlefranc@laas.fr
[2]LAAS CNRS, Toulouse, France

Keywords: AC-OPF, Moment-SOS hierarchy, Semidefinite Programming

AC-OPF (Alternative Current - Optimal Power Flow) aims at minimizing the operating costs of an AC power grid. It is well-known to be a difficult optimization problem in general, as it reformulates as a nonconvex QCQP (Quadratically Constrained Quadratic Program).

The moment-SOS (Sums-Of-Squares) hierarchy has proved relevant to solve AC-OPF instances to global optimality [1]. However, obtaining the convergence of the hierarchy may requires to go beyond the first step of the involved sequence of SDP (Semidefinite Programming) relaxations (e.g. up to the third step in [1]), and thus to solve semidefinite programs whose size grows drastically at each step of the hierarchy. Thus, the application of such relaxations to large scale AC-OPF problems (with thousands of variables) remains a challenging computing task.

Large polynomial optimization problems can be tackled efficiently if they are sufficiently sparse. In this talk, we present a new sparsity pattern, that we call *minimal sparsity*, inspired by the specific structure of the AC-OPF problem [2]. We show that minimal sparsity enables the computation of second order moment-SOS relaxations on large scale AC-OPF instances with far less computing resources — i.e. RAM and time — than the standard *correlative sparsity* pattern. Experimentally, we observe that it also provides tight lower bounds to certify the global optimality of AC-OPF solutions.

References

[1] Cédric Josz, Jean Maeght, Patrick Panciatici and Jean-Charles Gilbert. Application of the moment-SOS approach to global optimization of the OPF problem. *IEEE Transactions on Power Systems*, 30(1):463–470, 2014.

[2] Adrien Le Franc, Victor Magron, Jean-Bernard Lasserre, Manuel Ruiz and Patrick Panciatici. Minimal Sparsity for Second-Order Moment-SOS Relaxations of the AC-OPF Problem. *arXiv preprint arXiv:2305.19232*, 2023.

Duality of upper bounds in stochastic dynamic programming

Vincent Leclere Bernado Freitas Paulo Da Costa

ENPC, Francevincent.leclere@cermics.enpc.fr
FGV, Brazil, bernardofpc@im.ufrj.br

For multistage stochastic programming problems with stagewise inde- pendent uncertainty, dynamic programming algorithms computes approximations of the value functions at each stage. The Stochastic Dual Dynamic Programming (SDDP) algorithm provides polyhedral lower bounds, in the spirit of the Nested Benders algorithm. On the other hand, non- statistical upper bounds took a longer time to appear. One approach uses the primal dynamic programming recursion to build inner approximations of the value functions, while another builds lower approximations for the conjugate of the value functions. Unfortunately, the dynamic programming recursion for the conjugate value functions does not decompose over uncertainties, which suggests a Lagrangian relaxation. We prove that this Lagrangian relaxation corresponds exactly to the primal upper bounds for a natural choice of multipliers

Evolutionary Algorithm Outperforms MILP Solver on the Double-Floor Single-Row Facility Layout Problem

Junqi Liu[1] [2] Benjamin Doerr[2] Zeqiang Zhang[1] Martin S. Krejca[2] Berna Ulutas [3]

[1]School of Mechanical Engineering, Southwest Jiaotong University,
junqiLiu@my.swjtu.edu.cn
zhangzq@home.swjtu.edu.cn
[2] LIX, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris,
doerr@lix.polytechnique.fr,
martin.krejca@polytechnique.edu
[3] Department of Industrial Engineering, Eskisehir Osmangazi University,
bhaktan@ogu.edu.tr

Keywords: facility planning and design, double-floor single-row facility layout problem, mixed-integer linear programming, fast evolution algorithm, real-life case

We consider facility planning and design in the single-row facility layout problem (SRFLP) [1] for double-floor spaces to minimize land and material handling costs by efficiently arranging facilities along a given transport side of double-floor spaces. Previously, this problem has been solved based on a mixed-integer linear programming formulation. Our fast (1+1) EA operates by initially constructing a solution that aligns with the characteristics of SRFLP. It then proceeds to determine mutations using a heavy-tailed distribution function. To generate new solutions, we employ local insert and swap operators and incorporate a restart strategy into the process. We demonstrate the practical value of the fast (1+1) EA for DF-SRFLP on numerical experiments (9-80 facility-scale datasets) and real-world instances, including assembly workshops (size 14) and footwear manufacturing workshops (size 54).
The experimental results show that under the numerical experiments: the solve time of the MILP solver (GUROBI 10.0.1) grows exponentially with the problem size, while (1+1) EA is relatively stable. And the large-scale, MILP solver is not able to give suitable values in a reasonable time. In a practical application involving a 14-scale assembly line shop, the 1+1 EA achieves a layout that reduces material handling costs by 24 % compared to the original layout form. Furthermore, in the context of a 54-scale shoe factory, our proposed algorithm not only successfully addresses the DF-SRFLP but also offers a valuable reference for establishing an improved layout.

References

[1] Anjos, M. F., Vieira, M. V. C. (2021). Facility Layout., *Facility Layout. Springer International Publishing.*

[2] Doerr, B., Le, H. P., Makhmara, R., Nguyen, T. D. (2017). Fast genetic algorithms,*Proceedings of the Genetic and Evolutionary Computation Conference*, 777-784.

Semiconcave Dual Dynamic Programming and Its Application to N-body Problems

Marianne Akian[1] Stéphane Gaubert[1] Shanqing LIU[2] Yang Qi[1]

[1]INRIA Saclay-Île-de France, and CMAP, École Polytechnique, IP Paris, CNRS
{Marianne.Akian, Stephane.Gaubert, Yang Qi}@inria.fr
[2]CMAP, École Polytechnique, IP Paris, CNRS and INRIA, Shanqing.Liu@polytechnique.edu.

Keywords: Multistage Optimization, Hamilton-Jacobi-Bellman Equation, Dual Dynamic Programming, Optimal Trajectory, N-body System, Tropical Numerical Method, Semiconcavity

We introduce a novel algorithm for numerically finding the value function, along with the optimal trajectory, for a class of finite horizon deterministic optimal control problems of the form

$$\begin{aligned} &\max \int_0^T \ell(x(s),u(s))ds + \phi(x(T)) \\ &\text{s t. } \begin{cases} \dot{x}(s) = f(x(s),u(s)) \ , \\ x(s) \in X,\ u(s) \in U, \forall\ s \in [0,T] \ , \end{cases} \end{aligned} \tag{1}$$

where $x(0) = x_0$ is given, and the reward function ℓ is only required to be semiconcave w.r.t. x.

We look for a tight approximation of the value function along the optimal trajectories starting from a given initial point x_0. Our approach relies on approximating the value function from above, in a given time horizon, by a minimum of quadratic "basis" functions. However, the evolutionary semigroup associated with the maximization problem is max-plus linear instead of min-plus linear. We show that we can propagate the quadratic "basis functions" by solving a dual problem, which gives a new upper approximation of the value function. The trajectory is then updated to an optimal trajectory derived from the current approximate value function.

This method is inspired by, and can be thought of as, a generalization of the Dual Dynamic Programming algorithm [1, 2], adapted to *semiconcave* problems. We show that our method converges to the global maximum under certain regularity assumptions. We present numerical benchmarks, on N-body problems.

References

[1] M.V.F. Pereira and. L.M.V.G. Pinto, Multi-stage stochastic optimization applied to energy planning, *Mathematical Programming*, 36(6):1853-1879, 1998.

[2] P. Girardeau, V. leclere and A.B. Philpott, On the convergence of decomposition methods for multistage stochastic convex programs, *Mathematics of Operations Research*, 40(1):130–145, 2015.

[3] M. Akian, J-P. Chancelier and B. Tran, Tropical dynamic programming for Lipschitz multistage stochastic programming, *arXiv preprint*, 2010.10619, 2020.

Gradient descent with a general cost

Flavien Léger[2] Pierre-Cyril Aubin-Frankowski[2]

[1]INRIA Paris, flavien.leger@inria.fr
[2]TU Wien, pierre-cyril.aubin@inria.fr

Keywords: gradient descent, alternating minimization, c-concavity, five-point property, nonnegative cross-curvature

We present a new class of gradient-type optimization methods that extends vanilla gradient descent, mirror descent, Riemannian gradient descent, and natural gradient descent. Our approach involves constructing a surrogate for the objective function in a systematic manner, based on a chosen *cost function*. This surrogate is then minimized using an alternating minimization scheme. Using optimal transport theory we establish convergence rates based on generalized notions of smoothness and convexity. We provide local versions of these two notions when the cost satisfies a condition known as nonnegative cross-curvature. In particular our framework provides the first global rates for natural gradient descent and the standard Newton's method.

References

[1] F. Léger and P.-C. Aubin-Frankowski, *Gradient descent with a general cost*, arXiv:2305.04917, 2023.

An implementable proximal-type method for computing critical points to minimization problems with a nonsmooth and nonconvex constraint

Gregorio M. Sempere[1] Welington de Oliveira[1]

[1]MINES Paris PSL (CMA), France, gregorio.martinez_sempere@minesparis.psl.eu , welington.oliveira@minesparis.psl.eu

Keywords: Nonsmooth and nonconvex optimization, Epi-convergence, Bundle methods

Modern optimization problems from stochastic optimization, machine learning, uncertainty quantification, and others often involve functions that fail to be smooth and convex. Nevertheless, many such problems possess some structure that can and should be exploited in numerical optimization. This work investigates implementable necessary optimality conditions for a broad class of nonsmooth and nonconvex optimization problems involving objective and constraint functions that can be expressed as a finite sum of pointwise maxima of finitely many weakly-concave functions. Such a setting is general enough to accommodate nonconvex reliability-based optimization problems for which mathematically sound algorithms are scarce. Application areas include energy management, electric power systems, transportation, engineering design, and manufacturing. Our main contribution is a proximal-type algorithm with convergence guarantees to critical points. Our approach does not require penalization nor a feasible initial point. Numerical experiments on some stochastic reliability-based optimization problems illustrate the practical performance of the method.

Interior Point Methods in Optimal Control

Paul Malisani[1]

[1]IFP Energies nouvelles, Applied Mathematics Department, France, paul.malisani@ifpen.fr

Keywords: optimal control, state constraint, mixed constraints, interior point methods

This presentation deals with interior point methods for optimal control problems with pure state and mixed constraints. Interior point methods in optimal control have been studied in numerous papers [2, 1, 4], however, unlike their counterpart in numerical optimization, they have not yet been fully theorized. In [1], the authors show the convergence of the adjoint state, the state, and the control for problems with control constraints of the form $a \leq u(t) \leq b$ and for nonlinear systems affine in the control variable. The proof of convergence in [1] also relies on an assumption of uniqueness of the solution and on an assumption of strong convexity of the problem. In [2], the authors show the convergence of interior point methods in optimal control, with a primal-dual implementation, for problems with control constraints and using a strong Legendre-Clebsch condition which, in this case, is equivalent to a strong convexity assumption. In [3], the authors establish sufficient conditions on the state and control penalties to ensure that the solutions of a penalized optimal control problem strictly satisfy the constraints and prove the convergence of state and control variables using an assumption of uniqueness of the optimal solution and a strong convexity assumption. More recently, in [4], the author proves the convergence of interior point methods for the primal variables (state and control), the adjoint state, and the multipliers associated with the pure state and mixed constraints. This proof of convergence is established without using a strong convexity assumption but for nonlinear systems, mixed constraints and cost affine in the control. These results generalize [4] to a general class of dynamics and mixed constraints without requiring any additional assumptions. Interestingly, interior point methods in optimal control allow pure state constraints and mixed state-control constraints to be treated in the same way, i.e. using logarithmic penalties in each case. However, the difference in nature between these two types of constraints lies in the convergence topology of the dual variables associated with them. This contribution has been submitted to ESAIM COCV in may 2023.

References

[1] J. F. Bonnans, *Using logarithmic penalties in the shooting algorithm for optimal control problems*, Optim Control Appl Methods, 24:257–278, 2003

[2] M. Weiser, *Interior Point Methods in Function Space*, SIAM Journal on Control and Optimization, 44(5):1766-1786, 2003

[3] P. Malisani, F. Chaplais, N. Petit, *An interior penalty method for optimal control problems with state and input constraints of nonlinear systems*, Optim Control Appl Methods, 37(1):3–33, 2014

[4] P. Malisani *Interior point methods in optimal control problems of affine systems: Convergence results and solving algorithms*, SIAM SICON, 2023, accepted

Quantum Contextual Optimal Transport

Nicola Mariella[1,†] Albert Akhriev[1] Jannis Born[3] Eugene Koskin[4] Marianna Rapsomaniki[3] Stefan Woerner[2] Sergiy Zhuk[1]

[1]IBM Quantum, IBM Research Europe - Dublin, [†]corresponding author nicola.mariella@ibm.com
[2]IBM Quantum, IBM Research Europe - Zurich, Switzerland
[3]IBM Research, IBM Research Europe - Zurich, Switzerland
[4]University College Dublin, Belfield, Dublin 4, Ireland

The theory of *optimal transport* (OT) [3] is a mathematical framework that focuses on finding the optimal transport plan or map that minimizes the cost of moving resources from a source distribution $\boldsymbol{\mu}$ to a target distribution $\boldsymbol{\nu}$. The cost is typically defined as a distance or a dissimilarity measure between the source and target points, and the optimal transport plan aims to minimize this cost while satisfying certain constraints. In biology, OT gained popularity in various single-cell data analysis tasks, ranging from reconstructing trajectories of cell evolution , over predicting cell responses of therapeutical interventions [2] and inferring spatial and signaling relationships between cells. In these applications, the source and target distributions correspond to measurements of different biomolecules (e.g., mRNA, proteins) at a single-cell level.

In many applications, data measurements $\boldsymbol{\mu}_i$ and $\boldsymbol{\nu}_i$ (source/target distributions) are coupled to a context variable $\mathbf{x}_i$. In such cases, one might aspire to *learn* a global transportation map that is parameterized through $\mathbf{x}_i$ and thus facilitates the *prediction* of target states $\hat{\boldsymbol{\nu}}_j$ from source states $\boldsymbol{\mu}_j$, even for an unseen context $\mathbf{x}_j$. This extension of OT, *contextual OT* [1], consists of a model that predicts the transportation map conditioned on a context variable.

We propose a quantum formulation that leverages on the following concept. Let $\overline{(\cdot)}$ denote the complex conjugate of the argument, and $(\cdot)\odot(\cdot)$ the Hadamard product between matrices. Let U be a unitary matrix (possibly depending on context $\mathbf{x}$ and learning parameters $\boldsymbol{\theta}$), then unitarity implies that $U \odot \overline{U}$ is a *doubly stochastic matrix* (DSM), that is a special case of transportation map. Also, we prove that a generic transportation map with margins $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ can be embedded into a DSM. This concept shows that there is a natural link between transportation maps and unitary operators. Moreover, the latter could be a key component for an inductive bias that is well-suited for quantum computation.

References

[1] Charlotte Bunne, Andreas Krause, et al. Supervised training of conditional monge maps. In *Advances in Neural Information Processing Systems*, 2022.

[2] Charlotte Bunne, Stefan G Stark, Gabriele Gut, Jacobo Sarabia del Castillo, Kjong-Van Lehmann, Lucas Pelkmans, Andreas Krause, and Gunnar Rätsch. Learning single-cell perturbation responses using neural optimal transport. *bioRxiv*, pages 2021–12, 2021.

[3] C. Villani. *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2008.

Mean Field Analysis of an Incentive Policy on the Regulation of Free-Floating Car-Sharing Systems

Bianca Marin Moreno[1] Christine Fricker[2] Hanene Mohamed[3]

[1]INRIA Grenoble, EDF R&D, France, bianca.marin-moreno@inria.fr
[2]INRIA Paris, France christine.fricker@inria.fr
[3]Université Paris Nanterre, France hanene.mrad@parisnanterre.fr

Keywords: Large scale analysis, mean-field, car-sharing, incentive algorithm, stochastic network

In this talk we present a load-balancing algorithm for a closed stochastic network with two zones with different demands [1]. The algorithm is motivated by an incentive algorithm for redistribution of cars in a large-scale car-sharing system. The service area is divided into two zones. When cars stay too long in the low-demand zone, users are encouraged to pick them up and return them in the high-demand zone. The zones are divided in cells called stations. The cars are the network customers. We show that the mean-field limit solution of an ODE gives the large scale distribution of the station state in both clusters for this incentive policy in a discrete Markovian framework. An equilibrium point of this ODE is characterized via the invariant measure of a random walk in the quarter-plane. The proportion of empty and saturated stations measures how the system is balanced. Finally, we present numerical experiments illustrating the impact of the incentive policy. Our study shows that the incentive policy is useful when the high-demand zone is experiencing a shortage of cars, but that saturation must be avoided, especially when the high-demand zone is small.

References

[1] B. M. Moreno, C. Fricker, H. Mohamed, A. Philippe, M. Trépanier, Mean Field Analysis of an Incentive Algorithm for a Closed Stochastic Network, *AofA 2022*

Coordinated cross-border capacity calculation through the FARAO open-source toolbox

Alexandre Marié[1] Nicolas Omont[1] Felipe Gonzalez Venegas[1]

[1]Artelys, France,

alexandre.marie@artelys.com, nicolas.omont@artelys.com, felipe.gonzalez-venegas@artelys.com

Keywords: Optimal Power Flow, Security analysis, Remedial Actions Optimisation

Cross-borders exchanges have taken a major role in European strategy to achieve climate goals. The European Commission set a target of 15% interconnections in 2030, meaning that each country should have the physical capability to export at least 15% of their production. Increasing exchanges makes short term planning more complex.

In this context, the French TSO (RTE) released an open-source toolbox FARAO [1] to perform Coordinated Capacity Calculation (CCC) and ensure the security of supply.This involves running simulations to detect network violations. This involves running simulations (load flow calculation) to detect violations on the network and then determining the preventive or curative actions to be implemented to resolve contingencies at the lowest possible cost. This second step is called Remedial Actions Optimisation (RAO) and consists in solving an Optimal Power Flow (OPF) problem.

Artelys is a consultancy expert in power systems optimization and carries out various projects around TSO operational coordination in Europe. FARAO performs the optimization of both preventive and curative remedial actions, including HVDC lines, phase-shifter transformers and counter-trading but also topological actions. It is operationally used for the exchanges between Italy and its northern neighbors as well as between France, Spain and Portugal.

Artelys will present the algorithms of the FARAO toolbox and how they are actually used to enable greater operational coordination amongst the countries.

References

[1] Fully Autonomous Remedial Actions Optimisation: *An open-source modular toolbox for power system coordination process*, https://farao-community.github.io/.

A Two-Timescale Decision-Hazard-Decision Formulation for Storage Usage Values Calculation in Energy Systems Under Uncertainty

Camila Martinez Parra[*,1] Thomas Bittar[*,2] Jean-Marc Janin[*,3] Manuel Ruiz[*,4]
Michel De Lara[†,5] Jean-Philippe Chancelier[†,6] Pierre Carpentier[‡,7]

[*]RTE R&D, France, [†]CERMICS, École des Ponts, France, [‡]UMA, ENSTA Paris, France.

Keywords: Energy system modelling, information structure, stochastic multistage optimization.

The penetration of renewable energies requires additional storages to deal with the system's intermittency. Accordingly, there is growing interest in evaluating the opportunity cost associated with stored energy (usage value), which can be obtained by solving a multistage stochastic optimization problem.

Today, to compute usage values under uncertainties, a resource adequacy problem is solved using stochastic dynamic programming assuming a *hazard-decision* information structure. This modelling assumes completely knowledge of the coming week uncertainties which is not adapted to the system operation, as the intermittency occurs at smaller timescale. Moreover, we also assume full flexibility of thermal units. But in case of unpredictable outages, it takes time to switch other thermal units on. What happens if we take this temporal rigidity into account?

We introduce a two-timescale problem [1] with a new information structure considering planning and recourse decisions: *decision-hazard-decision*. This structure is used to decompose the multistage decision-making process into a non-anticipative planning step in which switch-on decisions for thermal units are made, and a recourse step in which power modulation decisions are made once the uncertainties have been disclosed.

We address the usage value computation problem for a uninodal system composed of a cluster of thermal units, a single storage unit and a load. The uncertainties are the residual demand of the system and the availability of thermal units.

We observe that the choice of the information structure for usage values computation can change the merit order in the system, i.e. the order of the storage usage value with respect to the thermal units prices. The policy induced by the weekly fully anticipative *hazard-decision* information modelling leads to an overestimation of the thermal flexibility and as a result, a lower use of the storage. When using the policy induced by the *decision-hazard-decision* Bellman functions, we take into account some rigidity in the thermal units (specially switch-on and switch-off decisions) and the storage is more used.

References

[1] Pierre Carpentier, Jean-Philippe Chancelier, Michel de Lara, Thomas Martin, and Tristan Rigaut. Time blocks decomposition of multistage stochastic optimization problems, 2023.

[1] camila.martinezparra@rte-france.com [2] thomas.bittar@rte-france.com [3] jean-marc.janin@rte-france.com
[4] manuel.ruiz@rte-france.com [5] michel.delara@enpc.fr [6] jean-philippe.chancelier@enpc.fr
[7] pierre.carpentier@ensta-paris.fr

Kolmogorov equations on spaces of measures associated to nonlinear filtering processes

Mattia Martini[1]

[1]Université Cotê d'Azur, France, mattia.martini@unice.fr

Keywords: Backward Kolmogorov equations, Stochastic filtering, Measure-valued processes, Differential equations on spaces of measures

Backward Kolmogorov equations are partial differential equations (PDEs) of parabolic type with a given final condition. The relation between them and certain stochastic processes has been intensively investigated in both finite and infinite dimensional cases. Moreover, their comprehension represents a preliminary step in the study of the Hamilton-Jacobi-Bellman equations, which are nonlinear partial differential equations strongly connected to stochastic optimal control problems.

This talk aims to present a class of backward Kolmogorov equations on spaces of probability and positive measures, associated to measure-valued stochastic processes arising in the context of nonlinear filtering. Indeed, in the filtering framework, one can formulate two stochastic differential equations, called Zakai and Kushner-Stratonovich equations, that are satisfied by a positive measure and a probability measure-valued process respectively. Thus, one can study the associated backward Kolmogorov equations, which are partial differential equations of parabolic type on the space of measures.

In the literature, a common approach is to assume that the measure-valued processes admit a density, to exploit stochastic calculus techniques in Hilbert spaces. Here the existence of a density is not assumed and everything is done directly in the context of measures. Due to this, the obtained equations will resemble the ones usually introduced in the context of mean field games and mean field control.

The study of these PDEs is a preparatory step to address the control problems with partial observation. Indeed, these control problems can be rephrased as optimal control problems with full observation of a measure-valued state which solves a filtering equation. It follows that it is natural to formulate the associated Hamilton-Jacobi-Bellman equations on the spaces of measures, as a nonlinear version of the aforementioned backward equations.

In the talk, we will introduce the Kolmogorov equations and then present a well-posedness result for classical solutions. If it remains time, we will discuss possible future applications and developments. The talk is based on [1].

References

[1] Martini, M. *Kolmogorov equations on spaces of measures associated to nonlinear filtering processes*, (2023) Stochastic Processes and their Applications, 161, 385-423.

A Speed Restart Scheme for a Dynamics with Hessian-Driven Damping

Juan José Maulén[1] Juan Peypouquet[2]

[1]University of Chile, Chile, & University of Groningen, Netherlands, j.maulen@dim.uchile.cl
[2]University of Groningen, Netherlands, j.g.peypouquet@rug.nl

Keywords: Convex optimization, Hessian-driven damping, First-order methods, Restart

In this talk, based on [1], we analyze a speed restarting scheme for the inertial dynamics with Hessian-driven damping, introduced in [2]. We establish a linear convergence rate for the function values along the restarted trajectories. Numerical experiments suggest that the Hessian-driven damping and the restarting scheme together improve the performance of the dynamics and corresponding iterative algorithms in practice.

References

[1] Maulén, J.J., Peypouquet, J. A Speed Restart Scheme for a Dynamics with Hessian-Driven Damping. J Optim Theory Appl (2023).

[2] Attouch, H., Peypouquet, J., Redont, P.: Fast convex optimization via inertial dynamics with Hessian driven damping. J. Differ. Equ. 261(10), 5734?5783 (2016).

Column Generation Approaches for the Satellites Constellation Design Problem

Luca Mencarelli[1] Andrea Bellè[2] Julien Floquet[3]

[1]Dipartimento di Informatica, Università di Pisa, 56127 Pisa, Italy, luca.mencarelli@unipi.it
[2]Thales Research & Technology, 91120 Palaiseau, France, andrea.belle@thalesgroup.com
[3]julien.floquet@protonmail.com

Keywords: Satellites constellation design problem, MILP, column generation

The satellites constellation design problem (SCDP) represents a fundamental aerospace application [1, 2]: it consists in designinig a constellation with the minimum possible number of satellites in order to guarantee periodic observability constraints within a fixed revisit time window and with respect to a given discrete set of target points on the Earth surface. This problem can be formulated as a large dimensional mixed-integer linear program (MILP) by suitably discretising the satellites orbital parameters. Under the assumption that two satellites cannot share the same (Keplerian) orbit in the optimal constellation configuration, we can easily reformulate the SCDP by assuming the number of satellites to be equal to the number of orbits. In this way, the SCDP consists in identifying the minimum number of activated satellites. In such a framework, the number of variables, although it grows polynomially with respect to the input size, can become prohibitively large in realistic instances, making the problem computationally intractable.

In order to address this issue, we propose two column generation approaches, based on different characterisations of restricted master (RMP) and pricing (PP) problems. In the first one we relax the observability constraints linking different satellites in the RMP, and we exploit it to generate new orbits with negative associated reduced costs in the PP; in the second one, we apply a strategy based on an edge covering problem on a opportunely defined multi-graph. For both the column generation approaches, the initialisation phase needed to define an initial feasible solution is implemented by locally solving a quite small number of non-convex nonlinear programs (NLPs) through a simple multi-start algorithm (in realistic instances, very few iterations are necessary to obtain convergence).

We present computational experiments on realistic instances, in order to demonstrate the efficacy and efficiency of our novel algorithmic framework. We conclude by briefly sketching two proving strategies for showing the **NP**-hardness of the SCDP.

References

[1] Mencarelli, L., Floquet, J., Georges, F., Grenier, D., Mixed integer (non)linear approaches for the satellite constellation design problem, *Optimization and Engineering* pp. 1–22 (2022)

[2] Mencarelli, L., Floquet, J., Georges, F., Matheuristics approaches for the satellite constellation design problem, *Optimization and Engineering*, to appear.

Computing Wasserstein Barycenter via operator splitting: the Method of Averaged Marginals

D. Mimouni*[1,3], P. Malisani†[1], J. Zhu‡[2], and W. de Oliveira§[3]

[1]IFP Energies nouvelles, Dpt. Applied Mathematics
[2]IFP Energies nouvelles, Dpt. Control and Signal Processing
[3]Mines Paris, Center for Applied Mathematics

Keywords: Wassertein Barycenters, Optimal Transport, Douglas-Rachford Method

The Wasserstein barycenter (WB) is an important tool for summarizing sets of probabilities. It finds applications in applied probability, clustering, image processing, etc. When the probability supports are finite and fixed, the problem of computing a WB is formulated as a linear optimization problem whose dimensions generally exceed standard solvers' capabilities. For this reason, the WB problem is often replaced with a simpler nonlinear optimization model constructed via an entropic regularization function so that specialized algorithms can be employed to compute an approximate WB efficiently. Contrary to such a widespread inexact scheme, we propose an exact approach based on the Douglas-Rachford splitting method applied directly to the WB linear optimization problem for applications requiring accurate WB. Our algorithm, which has the interesting interpretation of being built upon averaging marginals, operates series of simple (and exact) projections that can be parallelized and even randomized, making it suitable for large-scale datasets. As a result, our method achieves good performance in terms of speed while still attaining accuracy. Furthermore, the same algorithm can be applied to compute generalized barycenters of sets of measures with different total masses by allowing for mass creation and destruction upon setting an additional parameter. Our contribution to the field lies in the development of an exact and efficient algorithm for computing barycenters, enabling its wider use in practical applications. The approach's mathematical properties are examined, and the method is benchmarked against the state-of-the-art methods on several data sets from the literature.

*daniel.mimouni@ifpen.fr
†paul.malisani@ifpen.fr
‡jiamin.zhu@ifpen.fr
§welington.oliveira@minesparis.psl.eu

Optimal Carbon Pricing Based on Dynamical Climate Model

Thomas Mongaillard[1] Samson Lasaulce[1,2] Vineeth Varma[1]
Anthony Couthures[1] Essaid Sabir[3]

[1]Université de Lorraine, CNRS, CRAN, F-54000, Nancy, France, {thomas.mongaillard, anthony.couthures, vineeth.satheeskumar-varma, samson.lasaulce}@univ-lorraine.fr

[2]Khalifa University, Abu Dhabi, UAE

[3]Department of Computer Science, UQAM, Montreal, QC H2L 2C4, Canada, sabir.es-said@uqam.ca

Keywords: carbon pricing, climate, optimization

In this work, we adapt an optimal pricing framework [1] to the case of carbon pricing by including a climate model. We consider a set of countries $\mathcal{N}$ willing to maximize their own utility function modeled by

$$\forall n \in \mathcal{N},\ u_n(e_n, P) = b_n(e_n) - P\, e_n \tag{1}$$

with e_n the CO_2 emission level of country n, P the price of carbon and b_n is the benefit function of country n. These benefit functions are assumed to be sigmoidal and represent the evolution of the gross domestic product (GDP) of the country depending on its CO_2 emissions. To fix the price of carbon, we introduce a regulatory authority willing to maximize a global utility function modeled by

$$U(\boldsymbol{e}, P) = \sum_{n \in \mathcal{N}} b_n(e_n) - D(\theta^{\mathrm{AT}}(\boldsymbol{e})) \tag{2}$$

where D is a non-decreasing function of the temperature θ^{AT} that represents the climate damages. The countries are assumed to play their best response e_n^* to the price P, so that $\forall n \in \mathcal{N}$, $e_n = e_n^*(P)$ and $U(\boldsymbol{e}, P) = U(P)$. Then, the regulatory authority can choose the price P that maximizes the global utility. The coupling between the temperature and the total CO_2 emission is given by a climate model that includes a carbon cycle model (CC) and a temperature dynamic model (TD) [2]:

$$(CC): \begin{cases} C[k+1] = A_C\, C[k] + b_C\, e[k] & \text{(3a)} \\ C^{\mathrm{AT}}[k] = d_C\, C[k] & \text{(3b)} \end{cases}$$

$$f[k] = F_{2\times \mathrm{CO}_2} \log_2 \left[\frac{C^{\mathrm{AT}}[k]}{C^{\mathrm{AT}}_{\mathrm{ref}}} \right] + F_{\mathrm{nonCO}_2}[k] \tag{5}$$

$$(DT): \begin{cases} \theta[k+1] = A_\theta\, \theta[k] + b_\theta\, f[k] & \text{(4a)} \\ \theta^{\mathrm{AT}}[k] = d_\theta\, \theta[k] & \text{(4b)} \end{cases}$$

where $e[k]$ is the CO_2 emission, $C[k]$ the vector of carbon concentration, $f[k]$ the radiative forcing and $\theta[k]$ the vector of temperature at instant t_k.

After deriving the optimal emission level for a given price, we analyze the performances of the updated sequence of such a pricing method over a given timespan compared to a fixed price.

References

[1] Pedram Samadi et al. "Optimal Real-Time Pricing Algorithm Based on Utility Maximization for Smart Grid". In: *2010 First IEEE International Conference on Smart Grid Communications*. Oct. 2010, pp. 415–420. DOI: 10.1109/SMARTGRID.2010.5622077.

[2] Simon Dietz et al. "Are Economists Getting Climate Dynamics Right and Does It Matter?" In: *Journal of the Association of Environmental and Resource Economists* 8.5 (Sept. 2021), pp. 895–921. ISSN: 2333-5955. DOI: 10.1086/713977.

Optimal Vaccination Strategies for Metropolitan Areas

Lucas Machado Moschen[1] María Soledad Aronna[2]
[1]Sorbonne University, France, lucas.moschen@etu.sorbonne-universite.fr
[2]Fundação Getulio Vargas, Brazil, soledad.aronna@fgv.br

Keywords: Constrained affine optimal control, Epidemiology, Vaccination, Commuting

This study presents a mathematical model to determine optimal vaccination strategies in metropolitan areas, considering commuting patterns [4]. The epidemiological model employed is a compartmental SEIR model, incorporating a vaccination rate as a control function for each city. The commuting patterns are incorporated using a weighted adjacency matrix and a parameter that weighs day and night periods [6]. The optimal control (OC) problem is formulated to minimize a balance between the number of hospitalizations and vaccines, including restrictions of a weekly availability cap and an application limit of vaccines per time unit [5], leading to mixed control-state and pure-state constraints, which are challenging to manage [1, 2]. Additionally, our study addresses a control-affine problem, which, while practical, poses mathematical difficulties [3]. The key findings are the bounds for the basic reproduction number, particularly in the case of a metropolitan area, and the OC problem analysis. Theoretical analysis and numerical simulations provide insights into disease dynamics and the effectiveness of control measures. The research highlights the importance of prioritizing vaccination in the capital to reduce disease spread faster, as we depicted in our numerical simulations, especially when it has higher contact rates. Future research will refine this strategy and explore theoretical aspects of constrained control-affine problems.

References

[1] Biswas, M. H. A., et al. A SEIR model for control of infectious diseases with constraints. *Mathematical Biosciences and Engineering*, 11.4, 761–784, 2014.

[2] Boccia, A., de Pinho, M. D. R., and Vinter, R. B. Optimal Control Problems with Mixed and Pure State Constraints. *SIAM Journal on Control and Optimization*, 2016.

[3] De Pinho, M. R., Kornienko, I., and Maurer, H. Optimal control of a SEIR model with mixed constraints and L1 cost. *Proceedings of the 11th Portuguese Conference on Automatic Control*. Springer. 135–145, 2015.

[4] Moschen, L. M., Aronna, M. A. Master's dissertation on Optimal vaccination strategies for metropolitan areas. 2023.

[5] Lemaitre, J. C. et al. Optimal control of the spatial allocation of COVID-19 vaccines: Italy as a case study. *PLoS computational biology*, 18.7, 2022.

[6] Nonato, L. G., Peixoto, P., Pereira, T., Sagastizábal, C., and Silva, P. J. S. Robot Dance: A mathematical optimization platform for intervention against COVID-19 in a complex network. *EURO Journal on Computational Optimization 10*, 2022.

Moving Horizon Estimation for Anesthesia dynamics

Kaouther Moussa[1,2] Bob Aubouin-Pairault[3,4] Mazen Alamir[3] Mohammad Ajami[3] Thao Dang[4]

[1]UPHF, CNRS, UMR 8201 - LAMIH, F-59313 Valenciennes, France,
[2]INSA Hauts-de-France, F-59313, Valenciennes, France,
kaouther.moussa@uphf.fr
[3]Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France,
bob.aubouin-pairault@gipsa-lab.fr, mazen.alamir@gipsa-lab.fr
mohammad.ajami@gipsa-lab.fr
[4]Univ. Grenoble Alpes, CNRS, Grenoble INP, VERIMAG, 38000 Grenoble, France
thao.dang@univ-grenoble-alpes.fr

Keywords: Moving Horizon Estimation, Anesthesia dynamics, Pharmacodynamic

During surgery, the anesthesiologists use different types of anesthetic agents in order to induce and maintain the depth of unconsciousness, the absence of movement and the absence of pain [1]. The level of hypnosis is usually measured via processed variables such as the Bispectral Index (BIS) which takes values in the interval $[0, 100]$. The range $[40, 60]$ characterizes a moderate hypnosis state and the value 50 is generally considered as suitable for surgery. In order to monitor the BIS signal, different models are used, involving pharmacokinetic (PK) and pharmacodynamic (PD) parameters. Contrary to PK parameters, the PD parameters do not have standard expressions in terms of patient information. According to [2], the uncertainty introduced by the PD parameters shows a more significant influence on the measurable outputs than the one introduced by the PK parameters.
Therefore, this presentation will focus on an extended moving horizon observer [4] to estimate both the states and the pharmacodynamic (PD) parameters of an anesthesia model, based on real data. The inputs of this model are the injection rates of Propofol and Remifentanil. The states represent the concentration of the anesthetic agents in different compartments of the human body (muscles, fat, blood) and in the effect site. The considered output is the Bispectral index (BIS) which is derived from the electroencephalogram (EEG). The observer is designed such that the parameters are estimated during the anesthesia induction phase, and then almost frozen for the rest of the surgery. The estimator is validated on real data that were extracted from the VitalDB database [3].

References

[1] Ionescu, C. M., Neckebroek, M., Ghita, M., and Copot, D. An open source patient simulator for design and evaluation of computer based multiple drug dosing control for anesthetic and hemodynamic variables. *IEEE Access 9* (2021), 8680–8694.

[2] Krieger, A., Panoskaltsis, N., Mantalaris, A., Georgiadis, M. C., and Pistikopoulos, E. N. Modeling and analysis of individualized pharmacokinetics and pharmacodynamics for volatile anesthesia. *IEEE Transactions on Biomedical Engineering 61*, 1 (2014), 25–34.

[3] Lee, H. C., Park, Y., Yoon, S. B., Yang, S. M., Park, D., and Jung, C. W. VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients. *Sci Data 9*, 1 (Jun 2022), 279.

[4] Moussa, K., Aubouin-Pairault, B., Alamir, M., and Dang, T. Data-based extended moving horizon estimation for miso anesthesia dynamics. *IEEE Control Systems Letters 7* (2023), 3054–3059.

A Machine Learning based Branch-Cut-and-Benders for Dock Assignment and Truck Scheduling Problem in Cross-Docks

Rahimeh N. Monemi [1],[2] Shahin Gelareh [3] Nelson Maculan [4]

[1]Sharkey Predictim Globe, France, r.n.monemi@gmail.com
[2]IT and Business Analytics Ltd, London, United Kingdom,
[3]Département R&T, IUT de Béthune, Université dArtois, F-62000 Béthune, France,
shahin.gelareh@gmail.com
[4]Federal University of Rio de Janeiro, COPPE-PESC, P.O. Box 68511,
Rio de Janeiro, RJ 21941-972, Brazil, nelson.maculan@gmail.com

Keywords: Cross-docking, MILP modeling, Benders decomposition, Classification techniques, Machine learning.

In this work, we study the dock-door assignment and truck scheduling problem in crossdocks with an intraday planning horizon. This timeframe allows us to view the problem as a repeating operation with a frequency of 24 h. In practice, this repeating process follows a certain distribution, which is largely sustained even if we extend the horizon. While several modeling approaches and efficient solution algorithms have been proposed for various problem variations, the utilization of decomposition techniques in exact mathematical programming methods has been the most effective. Surprisingly, none of these techniques have taken advantage of the repeating patterns inherent in the problem. We start with a recently proposed compact model that is well-designed and can be exploited in a primal (Benders) decomposition technique, although it cannot be directly used to solve a practical-sized problem. We show that its modeling deficiencies can be fixed and propose a Benders decomposition framework together with several Alternative Objective Functions (AOFs) to generate customized Benders cuts, along with other valid inequalities that can be identified and separated. A classifier is trained to identify the most efficient AOF to use at different stages of the Benders iterations, to help avoid saturation of the master problem with dominated Benders cuts. Our extensive computational experiments confirm the significant performance improvement in the comparable decomposition framework.

References

[1] Gelareh, S., Monemi, R.N., Semet, F., Goncalves, G., 2016. A branch-and-cut algorithm for the truck dock assignment problem with operational time constraints., European Journal of Operational Research, 249, 1144-1152.

[2] Neamatian Monemi, R., Gelareh, S., 2023. Dock assignment and truck scheduling problem; consideration of multiple scenarios with resource allocation constraints., Computers and Operations Research, 151, 106074. http://dx.doi.org/10.1016/j.cor.2022.106074

An ODE characterisation of Entropic Multi-Marginal Optimal Transport

Luca Nenna[1] Brendan Pass[2]

[1]Université Paris-Saclay, CNRS, LMO, 91405, Orsay, France, luca.nenna@universite-paris-saclay.fr
[2] Department of Mathematical and Statistical Sciences, 632 CAB, U. of A., Edmonton, Alberta, Canada, T6G 2G1,pass@ualberta.ca

Keywords: Optimal Transport, multi-marginal OT, Entropic Regularization

The purpose of this talk is to introduce a new numerical method to solve multi-marginal optimal transport problems with pairwise interaction costs. The complexity of multi-marginal optimal transport generally scales exponentially in the number of marginals m. We introduce a one parameter family of cost functions that interpolates between the original and a special cost function for which the problem's complexity scales linearly in m. We then show that the solution to the original problem can be recovered by solving an ordinary differential equation in the parameter η, whose initial condition corresponds to the solution for the special cost function mentioned above; we then present some simulations, using both explicit Euler and explicit higher order Runge-Kutta schemes to compute solutions to the ODE, and, as a result, the multi-marginal optimal transport problem. References like [1] can be provided.

References

[1] Nenna, Luca, and Brendan Pass. *An ODE characterisation of multi-marginal optimal transport with pairwise cost functions.* arXiv preprint arXiv:2212.12492 (2023).

Convergence rate of entropy-regularized multi-marginal optimal transport costs

Luca Nenna[1] Paul Pegon[2]

[1]LMO, Université Paris-Saclay, France, luca.nenna@universite-paris-saclay.fr
[2]Ceremade, Université Paris-Dauphine & Mokaplan, Inria Paris, France, pegon@ceremade.dauphine.fr

Keywords: optimal transport, multi-marginal optimal transport, entropic regularization, Schrödinger problem, convex analysis, Rényi dimension

We investigate the convergence rate of multi-marginal optimal transport costs that are regularized with the Boltzmann-Shannon entropy, as the noise parameter ε tends to 0. We establish lower and upper bounds on the difference with the unregularized cost of the form $C\varepsilon\log(1/\varepsilon) + O(\varepsilon)$ for some explicit dimensional constants C depending on the marginals and on the ground cost, but not on the optimal transport plans themselves. Upper bounds are obtained for Lipschitz costs or semi-concave costs (for a finer estimate), and lower bounds for C^2 costs satisfying some signature condition on the mixed second derivatives that may include degenerate costs, thus generalizing results previously obtained with Carlier and Tamanini [2], and by Eckstein and Nutz [3]. We obtain in particular matching bounds in some typical situations where the optimal plan is deterministic, like in the case of Wasserstein barycenters. This is a joint work with Luca Nenna [1].

References

[1] Luca Nenna and Paul Pegon, Convergence Rate of Entropy-Regularized Multi-Marginal Optimal Transport Costs, *arXiv*, arXiv:2307.03023, 2023.

[2] Guillaume Carlier, Paul Pegon and Luca Tamanini, Convergence Rate of General Entropic Optimal Transport Costs, *Calculus of Variations and Partial Differential Equations*, 62(4), 2023.

[3] Stephan Eckstein and Marcel Nutz, Convergence Rates for Regularized Optimal Transport via Quantization, *Mathematics of Operations Research*, Ahead of Print, July 2023.

Pickup and Delivery Problem with Cooperative Robots

Chi Thao Nguyen[1] Jean-Philippe Gayon[1] Viet Hung Nguyen[1]
Alain Quilliot[1] Anh Son Ta[2]

[1]Clermont-Auvergne University, CNRS, Mines of Saint-Etienne, Clermont Auvergne INP, LIMOS, 63000 Clermont-Ferrand, France,
chi_thao.nguyen@uca.fr, j-philippe.gayon@uca.fr, viet_hung.nguyen.uca.fr, alain.quilliot@uca.fr
[2]Hanoi University of Science and Technology, Hanoi, Vietnam, son.taanh@hust.edu.vn

Keywords: Warehouse, Traveling Salesman Problem, Pickup and Delivery Problem, Operations Research, Robots Reconfiguration

This paper explores the Pickup and Delivery Problem with Cooperative Robots (PDP-CR), a new interest that has emerged in warehouse settings with automated vehicles where multiple robots must cooperate to complete a task. i.e. when a task requires more than one robot. In PDP-CR, we have a fleet of identical robots and a set of tasks. A task has a pick-up point, a destination, and may require multiple robots. We want to complete all the tasks using these starting robots. A task is considered done if all robots needed for the task reach its pick-up point and then go to the destination simultaneously. After the robots complete a task, they may go to other tasks. The goal of PDP-CR is to minimize the makespan, i.e. the time at which the last robot returns to the depot. To our knowledge, the PDP-CR is a problem that has not been studied. However, some formulation ideas can be obtained in related problems. The Pick-up Delivery Problem (PDP) [1] deals with the transportation of goods from an origin to a destination. Synchronization in VRP [2, 3, 4] refers to the constraints that require two or more vehicles to carry out a single task. The primary goal of this paper is to develop and test several mathematical models for the PDP-CR.

References

[1] S. N. Parragh, K. F. Doerner, and R. F. Hartl, "A survey on pickup and delivery problems: Part II: Transportation between pickup and delivery locations," *Journal fÃ¼r Betriebswirtschaft*, vol. 58, pp. 81–117, June 2008.

[2] M. Drexl, "Synchronization in Vehicle RoutingâA Survey of VRPs with Multiple Synchronization Constraints," *Transportation Science*, vol. 46, pp. 297–316, Aug. 2012.

[3] B.-I. Kim, J. Koo, and J. Park, "The combined manpower-vehicle routing problem for multi-staged services," *Expert Systems with Applications*, vol. 37, pp. 8424–8431, Dec. 2010.

[4] B. L. Hollis, M. A. Forbes, and B. E. Douglas, "Vehicle routing and crew scheduling for metropolitan mail distribution at Australia Post," *European Journal of Operational Research*, vol. 173, pp. 133–150, Aug. 2006.

Proportional Fairness for Multi-Objective Optimization

Minh Hieu Nguyen[1] Mourad Baiou[1] Viet Hung Nguyen[1]

[1]LIMOS, Clermont Auvergne University, France.
{minh_hieu.nguyen, mourad.baiou, viet_hung.nguyen}@uca.fr

Keywords: Multi-Objective Optimization, Multi-Criteria Decision Making, Pareto optimal, Weighted Sum Method, Proportional Fairness

This work deals with a particular case of Multi-Objective Optimization called *Bi-Objective Discrete Optimization* (BODO), where the feasible set is discrete, and the two objectives take only positive values. Since the feasible set of a BODO problem is discrete and usually finite, it can theoretically be enumerated to identify the Pareto set, which consists of all Pareto-optimal solutions representing different trade-offs between two objectives. However, in general, this problem is algorithmically unsolvable. From a practical point of view, the Central Decision Maker (CDM) may be interested in a reduced Pareto set reflecting the own preference of the CDM, which a computationally tractable algorithm can obtain.

In this work, we propose a novel selection criterion for BODO that can efficiently guide the Weighted Sum Method to find the preferred Pareto-optimal solutions achieving some fair competition between the two objectives based on proportional fairness [1]. The latter aims to provide a compromise between the utilitarian rule - which emphasizes overall system efficiency, and the egalitarian rule - which emphasizes individual fairness. For our purpose, we introduce the concept of the *generalized Nash Fairness* solution, i.e., ρ-NF solution, for BODO where $\rho > 0$ is a factor denoting the relative importance of the first objective comparing to the second one [2], [3]. This allows the CDM to consider that ρ percent change of the first objective is comparably equivalent to a one percent change of the second one. Hence, when switching from a ρ-NF solution to any other feasible solution, the sum of the factor ρ of the proportional change in the first objective and the proportional change in the second objective is not negative.

We first introduce the concept and discuss the existence of ρ-NF solution for BODO. We then show that the ρ-NF solution set is a subset of the Pareto set. We also propose a recursive Newton-like algorithm for determining the ρ-NF solution set. Finally, an illustrative example of BODO is given.

References

[1] D. Bertsimas, V.F. Farias and N. Trichakis, The price of fairness, *Operations Research* 59-1, pp 17-31, February 2011.

[2] M.H. Nguyen, M. Baiou, V.H. Nguyen and T.Q.T. Vo, Generalized Nash Fairness solutions for Bi-Objective Minimization Problems, *Networks*, pp 1-17, August 2023.

[3] M.H. Nguyen, M. Baiou, V.H. Nguyen, Determining the generalized Nash Fairness solution set for Bi-Objective Discrete Optimization, *Submitted*, March 2023.

Complexity analysis of regularization methods for implicitly constrained least squares

Akwum Onwunta[1] Clément W. Royer[2]

[1]Lehigh University, USA ako221@lehigh.edu
[2]Université Paris Dauphine-PSL, France clement.royer@lamsade.dauphine.fr

Keywords: Complexity analysis, Implicit constraints, PDE-constrained optimization.

Optimization problems constrained by partial differential equations (PDEs) naturally arise in scientific computing, as those constraints often model physical systems or the simulation thereof. In an implicitly constrained approach, the constraints are incorporated into the objective through a reduced formulation. To this end, a numerical procedure is typically applied to solve the constraint system, and efficient numerical routines with quantifiable cost have long been developed. Meanwhile, the field of complexity in optimization, that estimates the cost of an optimization algorithm, has received significant attention in the literature, with most of the focus being on unconstrained or explicitly constrained problems. In this paper, we analyze an algorithmic framework based on quadratic regularization for implicitly constrained nonlinear least squares. By leveraging adjoint formulations, we can quantify the worst-case cost of our method to reach an approximate stationary point of the optimization problem. Our definition of such points exploits the least-squares structure of the objective, leading to an efficient implementation. Numerical experiments conducted on PDE-constrained optimization problems demonstrate the efficiency of the proposed framework.

Accelerating nonlinear programming with GPUs

François Pacaud

CERMICS, francoispacaud8@gmail.com

In this talk, we present a roadmap to implement an interior-point nonlinear solver on SIMD (Single Instructions, Multiple Data) architectures, such as GPUs. The interior-point method has become the workhorse method to solve large-scale optimization problems. By using a homotopy method, IPM obtains a smooth reformulation of the Karush-Kuhn-Tucker (KKT) conditions associated with the optimization problem. This allows us to compute the next descent direction directly by solving a KKT linear system, usually using a sparse indefinite solver in the nonconvex case. However, IPM's performance is limited by the indefinite sparse linear solver used under the hood, which has impaired its adoption on hardware accelerators such as GPUs. Luckily, we can rewrite the KKT system in a form more amenable to GPUs using different variants of the Schur complement method. Notably, we show that we can leverage the structure of the inequality and equality constraints using a condensed then reduced strategy to compress the sparse KKT system down to a dense matrix, easy to factorize. When used in conjunction with a parallel automatic differentiation backend, this strategy is effective in solving large-scale nonlinear problems on GPUs. We give extensive numerical results on large-scale optimal power flow (OPF) instances to illustrate the effectiveness of the method compared to the state-of-the-art.

Statistical Empirical Risk Minimization: The Pursuit of Generalization

Samir M. Perlaza
Centre Inria d'Université Côte d'Azur, INRIA, Sophia Antipolis, France, samir.perlaza@inria.fr

Keywords: f-Divergences, Gibbs Probability Measures, and Information Measures

In this talk, the problem of empirical risk minimization (EMR) is studied using tools from statistics, measure theory, and information theory to establish generalization guarantees for machine learning algorithms. More specifically, the ERM with regularization by f-divergences, often referred to as statistical ERM, is studied assuming that priors might be in the form of a σ-finite measure (the *reference measure*), and not necessarily a probability measure. Under this assumption, which leads to a generalization of the statistical ERM problem allowing a larger degree of flexibility for incorporating prior knowledge, numerous relevant properties are stated. Among these properties, the solution to the statistical ERM, if it exists, is shown to be a unique probability measure for several choices of the function f. In particular, it is shown that the solution is often mutually absolutely continuous with the reference measure, which introduces a strong inductive bias that dominates the evidence provided by the training data [1, 2]. Interestingly, such a solution exhibits a probably-approximately-correct guarantee for the ERM (without regularization) independently of whether such a problem possesses a solution. In the case in which $f(x) = x\log(x)$, for a fixed training dataset, the empirical risk is shown to be a sub-Gaussian random variable when the models are sampled from the solution to the statistical ERM problem. The generalization capabilities of the solution to this problem (the Gibbs algorithm) are studied via the sensitivity of the expected empirical risk to deviations from such a solution towards alternative probability measures. Finally, an interesting connection between sensitivity, generalization error, and lautum information is stablished.The talk ends by highlighting the impact of the asymmetry of f-divergences in statistical ERM. In particular, regularization by relative entropy (KL divergence) is studied in both directions. These regularizations are coined *Type-I* and *Type-II*, respectively. Finally, it is proved that Type-II regularization is equivalent to Type-I regularization with an appropriate transformation of the empirical risk function [3].

References

[1] S. M. Perlaza, G. Bisson, I. Esnaola, A. Jean-Marie, and S. Rini, "Empirical risk minimization with relative entropy regularization: Optimality and sensitivity," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Espoo, Finland, Jul. 2022, pp. 684–689.

[2] S. M. Perlaza, I. Esnaola, G. Bisson, and H. V. Poor, "On the validation of Gibbs algorithms: Training datasets, test datasets and their aggregation," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Taipei, Taiwan, Jun. 2023.

[3] F. Daunas, I. Esnaola, S. M. Perlaza, and H. V. Poor, "Analysis of the relative entropy asymmetry in the regularization of empirical risk minimization," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Taipei, Taiwan, Jun. 2023.

A Mean-Field Game Model for Energy Transition

LAVIGNE Pierre[1] LECLERE Hugo[2] PETIT Quentin[2] WARIN Xavier[2]

[1]Institut Louis Bachelier, France, pierre.lavigne@institutlouisbachelier.org
[2]EDF R&D, France, hugo.leclere@edf.fr, quentin.petit@edf.fr, xavier.warin@edf.fr

Keywords: Mean-Field Game, Numerical Simulations, Ecological Transition

We introduce a mean-field game model in which agents can invest in two types of capital: green capital and polluting capital. The productivity of each type of capital depends on a pollution variable subject to noise and common to all agents. We establish the existence of equilibria by solving a classical FBSDE. Numerical experiments are also carried out: we observe the energy transition.

Improving stochastic control by ignoring randomness

Andy Philpott[1] Dominic Keehan[1] Eddie Anderson[2]

[1]Electric Power Optimization Centre, University of Auckland, New Zealand, a.philpott@auckland.ac.nz
[1]Electric Power Optimization Centre, University of Auckland, New Zealand, dkee331@aucklanduni.ac.nz
[2]Imperial College, London, e.anderson@imperial.ac.uk

Keywords: Stochastic control, model predictive control, sample average approximation, distributional robustness

It is widely regarded as a cardinal sin in optimization under uncertainty to replace random quantities by their expectations before applying some solution technique. In some settings taking expectations is justified through a certainty equivalence principle, but it is easy to show that this never makes solutions strictly better than those obtained by solving the stochastic optimization problem. However, in nearly every practical setting, a probability distribution must be estimated from data, assumed to be a sample drawn from the distribution. Here, optimizing with an approximate distribution based on the sample can lead to strictly worse outcomes than using a model based only on the average of the sample values. We describe this phenomenon and illustrate it using a simple inventory control example.

Semidefinite programming by projective-cutting-planes

Daniel Porumbel[1]

[1]CEDRIC CS Lab, Conservatoire National des Arts et Métiers, Paris, France daniel.porumbel@cnam.fr

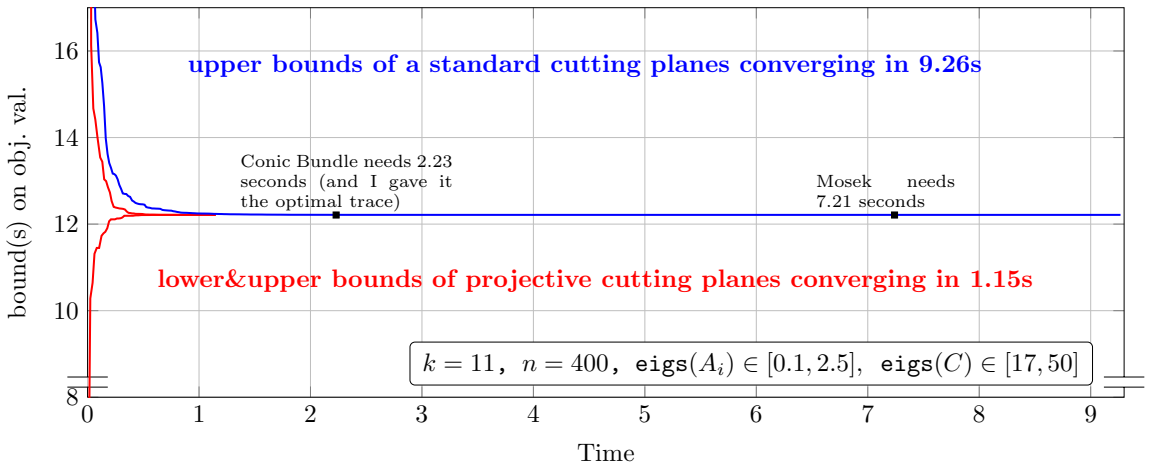
Keywords: Semidefinite programming, Projective Cutting Planes, Projection Sub-problem

We focus on the following standard (semidefinite programming) SDP problem, where $\mathcal{A}^\top \mathbf{y} = \sum_{i=1}^{k} A_i y_i$. The implication in the last constraint indicates that the SDP constraint can be expressed using an infinite number of linear cuts. We could attack the problem by progressively removing infeasible intermediate solutions D using the separation sub-problem; yet, it is well known that a standard cutting planes is not a good idea.

$$(SDP)\begin{cases} \max \quad \mathbf{b}^\top \mathbf{y} & (1a) \\ \quad s.t \quad S = C - \mathcal{A}^\top \mathbf{y} & (1b) \\ \qquad S \succeq \mathbf{0} \iff S \bullet \mathbf{s}\mathbf{s}^\top \geq 0 \ \forall \mathbf{s} \in \mathbb{R}^n & (1c) \end{cases}$$

Around ten years ago I started to publish work on the idea of upgrading the separation sub-problem to a projection (or intersection) one: given a non-feasible solution $\mathbf{d}$ in a polytope $\mathscr{P}$, what is the maximum step-length t^* so that $t^*\mathbf{d} \in \mathscr{P}$? This amounts to projecting from $\mathbf{0}$ towards $\mathbf{d}$. This idea would make no sense in above program because a projection from $\mathbf{0}$ towards some D returns ∞ if $D \succeq \mathbf{0}$ or 0 otherwise. Around five years ago I introduced Projective Cutting Planes [1] because I realized it may be possible to project from any feasible point instead of zero in some problems. I now think (SDP) is such a problem. To make Projective Cutting Planes work, we have to be able to rapidly solve $t^* = \max\{t : S + tD \succeq \mathbf{0}\}$ for a given $S \succeq \mathbf{0}$ and an arbitrary D.

This projection sub-problem is quite simple if $S \succ \mathbf{0}$. In this case, there is a unique Cholesky decomposition $S = KK^\top$ and K is non-singular. Thus, we can determine a unique E so that $D = KEK^\top$. Writing $S + t \cdot D = KI_nK^\top + t \cdot KEK^\top$, this matrix is congruent (see, *e.g.*, Prop 1.2.3 of [2]) to $I_n + t \cdot E$; thus, these matrices have same SDP status. We have to solve $t^* = \max\{t : I_n + tE \succeq \mathbf{0}\}$ which returns $t^* = -\frac{1}{\lambda_{\min}(E)}$. This projection is more difficult if S is not strictly SDP, among other reasons, because the congruence relation above no longer holds. Yet, this simplified case enabled us to solve some particular instances very rapidly:



References

[1] Daniel Porumbel, Projective Cutting-Planes, SIAM Journal on Optimization, 30(1): 1007-1032, 2020

[2] Daniel Porumbel, Demystifying the characterizations of SDP matrices in mathematical programming, unpublished expository work around 100 pages, cedric.cnam.fr/~porumbed/papers/sdp.pdf.

Lot Sizing, Job Scheduling and Energy Management

Alain Quilliot[1] Aurelien Mombelli[2]

[1]LIMOS Lab. University Clermont Auvergne, CNRS, alain.quilliot@uca.fr

Keywords: Lot Sizing, Scheduling, Energy Management

Synchronizing Lot-Sizing and Job Scheduling: *Lot Sizing* (see [3]) arises when one schedules production while avoiding strong activation costs. Part of a *Lot Sizing* strategy is to concentrate production on well-chosen periods in order to minimize those costs. One often handles resulting problems, mostly NP-Hard, through dynamic programming, which may yields FPTAS: *Fully Polynomial Time Approximation Scheme*. Now it may happen that such a process involves an interaction between the Lot Sizing player and a (several) job scheduler(s), both provided with their own agenda. Let us for instance consider a local energy producer $Feeder$ which produces energy all along N periods $i = 0, \ldots, N-1$ with varying production levels R_i and cost C_i. This energy producer interacts with a job scheduler $Eater$ who performs jobs $j = 0, \ldots, M-1$, each job requiring some energy and some time units. $Feeder$ must decide about a production vector $z = (z_i \in \{0,1\}, i = 0, \ldots, N-1)$, solution of the model:
$\{$ *Minimize* $\sum_i C_i.z_i$ such that $\sum_i R_i.z_i \geq Total\_Cons\}$,
where $Total\_Cons$ is the total amount of energy requested by $Eater$.
On the other side, $Eater$ must schedule jobs $j = 0, \ldots, M-1$, while optimizing its own criterion (for instance the makespan), while meeting temporal constraints and providing every job with energy. In order to implement this production/consumption process, $Feeder$ and $Eater$ must agree about the periods i and the dates t when $Feeder$ transfers some energy amount m to $Eater$. Such a transfer $\omega = (i, t, m)$ usually requires an additional cost, together with additional time and energy. In many case, it forbids energy production or job processing. According to this, we get a full description of our process if we know, besides $Feeder$ production vector z and $Eater$ schedule, the collection of all transfers $\omega = (i, t, m)$, that must be consistent with limited storage capacities. Resulting complex bi (multi)-level *Lot Sizing and Job Scheduling* problem with synchronization constraints typically arises in the context of (see [2])) a real time cooperation between electric vehicles and recharge facilities.

Our Contribution: We restrict ourselves to the $Feeder$ point of view and suppose that jobs $j = 0, \ldots, M-1$ are scheduled according to this order, so that the only scheduling decision is about their starting times. Then we formalize resulting **SLSS**: *Synchronized Lot Sizing/Scheduling* problem. Though the *centralized versus collaborative* issue is crucial in practice, we stick here to the centralized paradigm . We make first appear the central role played by the *transfers*, that define chains inside some partially ordered set and induce some *no-antichain* constraints inside MILP: *Mixed Integer Linear Programming* formulations. This leads us to a first algorithmic handling of **SLSS** through *branch and cut*. Next we enhance this MILP model with additional valid constraints, restricted to the decision variables, and get rid of non combinatorial variables while relying on Benders decomposition. Finally we reformulate **SLSS** as a path search problem inside some *transfer* network. This allows us to prove that **SLSS** is pseudo-polynomial and get a **PTAS** scheme. We conclude by discussing the collaborative issue.

References

[1] Biel K., Glock G. H.: *Systematic literature review of decision support models for energy efficient decision planning*: Computer and Industrial Engineering 101, p 243-259, (2016).

[2] Clark A., Almada Lobo B., Almeder C.: *Lot Sizing and Scheduling: Industrial extensions and research opportunities*; International Journal of Production Research 49 (9), p 2457-2461, (2011).

Funding: This work was sponsored by PGMO fundation.

Electric Vehicle Coordinated Charging through Hierarchical Optimization models

Luis Rojo-González[1,2] Miguel F. Anjos[4] Luce Brotcorne[3] Michel Gendreau[1]

[1]Centre Interuniversitaire de Recherche sur les Réseaux dâEntreprise, la Logistique et le Transport (CIRRELT) and Département de Mathématiques et de Génie industriel, Polytechnique Montreál, Montréal, H3C 3A7, Quebec, Canada

[2]Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France

[3]INRIA Lille Nord-Europe, 40 Avenue Halley, Lille 59000, France

[4]School of Mathematics, University of Edinburgh, United Kingdom

Keywords: Coordinated charging, Electric vehicles, Hierarchical optimization, Taxonomy.

The adoption of electric vehicles (EVs) has arisen as a good opportunity to reduce the transportation sector's contribution to global warming. The latter is given the fact that EVs do not produce greenhouse gases (GHGs) either when driving or idle. As a consequence, it is estimated that roughly 100 million EVs will be under use by 2030 around the world. However, if this becomes a reality, the energy generation must increase and, henceforth, imply -heavy- extra load and power variations to the grid. In order to tackle these negative effects, the so-called *coordinated charging* has been coined and recently studied to allocate power and supply energy such that power availability and infrastructure constraints are not violated. Moreover, the inclusion of renewable energy sources (RESs) and energy storage systems (ESSs) appear as the grid's most important allies. Accordingly, several modeling techniques have been adopted regarding the EV profiles and grid characteristics. Regardless of the model, it aims to control the energy supply and therefore restrict the maximum power, increasing the dwelling-time.

In general, the dwell-time flexibility is replaced by fixed but longer dwelling-times, e.g., at homes (night) and at workplaces (working hours), representing an important opportunity for new approaches as well as properly supply power since fully-charged batteries are not required at a single session. As such, the so-called vehicle-to-grid (V2G, a.k.a bi-directional power flow) has gained popularity giving economic benefits to EV owners by representing a distributed energy resource (DER).

There exists a rich literature regarding coordinated charging and the aforementioned elements, however, stakeholders' inherent relationships have just recently been considered. This work precisely overviews studies addressing the coordinated charging problem through hierarchical optimization models, along with the adopted resolution methods, and proposes a novel methodology to consolidate the existing studies. To this end, problems found in the literature are firstly characterized by identifying the elements involved in the phenomenon, such that a novel taxonomy is proposed to deal with the huge number of possible combinations a study may address.

Thus, the goal of this work is three-fold: i) to trace the evolution of models exploiting the hierarchical relationship between stakeholders in the problem under study, ii) to consolidate the existing literature, and iii) to set an identification method in order to compare state-of-the-art studies.

Fast, Differentiable and Sparse Top-k: a Convex Analysis Perspective

Michael E. Sander[1] Joan Puigcerver[2] Josip Djolonga[2] Gabriel Peyre[1] Mathieu Blondel[2]

[1]Ecole Normale Superieure et CNRS, Paris, France CNRS michael.sander@ens.fr
[2]Google Deepmind mblondel@google.com

Keywords: differentiable programming, top-k, permutahedron, isotonic optimization, smoothing)

The top-k operator returns a sparse vector, where the non-zero values correspond to the k largest values of the input. Unfortunately, because it is a discontinuous function, it is difficult to incorporate in neural networks trained end-to-end with backpropagation. Recent works have considered differentiable relaxations, based either on regularization or perturbation techniques. However, to date, no approach is fully differentiable *and* sparse. In this talk, we propose new differentiable *and* sparse top-k operators. We view the top-k operator as a linear program over the permutahedron, the convex hull of permutations. We then introduce a p-norm regularization term to smooth out the operator, and show that its computation can be reduced to isotonic optimization. Our framework is significantly more general than the existing one and allows for example to express top-k operators that select values *in magnitude*. On the algorithmic side, in addition to pool adjacent violator (PAV) algorithms, we propose a new GPU/TPU-friendly Dykstra algorithm to solve isotonic optimization problems. We successfully use our operators to prune weights in neural networks, to fine-tune vision transformers, and as a router in sparse mixture of experts.

References

[1] Sander, M. E., Puigcerver, J., Djolonga, J., Peyre, G., and Blondel, M. (2023, July). Fast, differentiable and sparse top-k: a convex analysis perspective. In International Conference on Machine Learning (pp. 29919-29936). PMLR.

Learning the Follower's Objective Function in Sequential Bilevel Games

Ioana Molan, Martin Schmidt, Johannes Thürauf

Trier University, Department of Mathematics, Universitätsring 15, 54296 Trier, Germany
molan@uni-trier.de,
martin.schmidt@uni-trier.de,
johannes.thuerauf@uni-trier.de

Keywords: Bilevel optimization, Inverse optimization, Unknown follower data, Multiplicative weight update, Inverse KKT method, Sequential games

We consider bilevel optimization problems in which the leader has no or only partial knowledge about the objective function of the follower. The studied setting is a sequential one in which the bilevel game is played repeatedly. This allows the leader to learn the objective function of the follower over time. We focus on two methods: a multiplicative weight update (MWU) method and one based on the lower-level's KKT conditions that are used in the fashion of inverse optimization. The MWU method requires less assumptions but the convergence guarantee is also only on the objective function values, whereas the inverse KKT method requires stronger assumptions but actually allows to learn the objective function itself. The applicability of the proposed methods is shown using two case studies. First, we study a repeatedly played continuous knapsack interdiction problem and, second, a sequential bilevel pricing game in which the leader needs to learn the utility function of the follower.

Learning Sparsified Networks in Column Generation: Applications to Multi-Commodity Network Flow Problems

Alexandre Schulz[1] Mathieu Lacroix[1] Joseph Le Roux[1] Sébastien Martin[2] Roberto Wolfler Calvo[1]

[1]Université Sorbonne Paris Nord, France,
schulz@lipn.univ-paris13.fr,lacroix@lipn.univ-paris13.fr, leroux@lipn.univ-paris13.fr, roberto.wolfler@lipn.univ-paris13.fr

[2]Huawei Technologies Research and Development, France, sebastien.martin@huawei.com

Keywords: Column Generation, Multi-Commodity Network Flow, Supervised Learning, Graph Convolution Networks

Column Generation (CG) is a well known method suited to solving Linear Optimisation problems that have an exponential number of variables such as those stemming from Danzig-Wolfe decomposition. CG consists in solving the Restricted Master problem (RMP); the original problem considering only a subset of variables and iteratively adding promising variables by solving a pricing problem. The algorithm stops as soon as no such variables exist.

CG is notably used for solving the Multi-Commodity Network Flow problem (MCNF) [1]. The latter consists, given a capacitated network and a set of demands, in routing the demands from their source to their target at minimum cost while fulfilling the network capacity requirements. In the MCNF path-flow formulation, a variable represents a path that can be used to route a demand, and new variables are found by solving a shortest-path problem for each demand, which can be computationally expensive for large instances. In telecommunication networks MCNFs are repeatedly solved over the same network with demands varying from instance to instance. We propose to leverage this information to devise an effective heuristic. Our approach relies on training a graph neural network [2] to identify regions of interest of the graph for each demand, reducing drastically its size and leading to an improved performance of the pricing algorithm. Improvements are illustrated by applying the aforementioned method on a set of realistic MCNFs, showing significant reduction in computation time with only slight degradation of the objective.

References

[1] C. Barnhart, C. A. Hane, P. H. Vance, Using Branch-and-Price-and-Cut to Solve Origin-Destination Integer Multicommodity Flow Problems, *Operations Research*, 48:318-326, 2000.

[2] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, Graph neural networks: A review of methods and applications, *AI Open*, 1:57-81, 2020.

Differential Privacy via Distributionally Robust Optimization

Aras Selvi[1] Huikang Liu[2] Wolfram Wiesemann[1]

[1]Imperial College Business School, United Kingdom, a.selvi19@imperial.ac.uk, ww@imperial.ac.uk
[2]Shanghai University of Finance and Economics, China, liuhuikang@sufe.edu.cn

Keywords: Differential Privacy; Privacy-Accuracy Trade-Off; Distributionally Robust Optimization

In recent years, differential privacy has emerged as the *de facto* standard for sharing statistics of datasets while limiting the disclosure of private information about the involved individuals. This is achieved by randomly perturbing the statistics to be published, which in turn leads to a privacy-accuracy trade-off: larger perturbations provide stronger privacy guarantees, but they result in less accurate statistics that offer lower utility to the recipients. Of particular interest are therefore optimal mechanisms that provide the highest accuracy for a pre-selected level of privacy. To date, work in this area has focused on specifying families of perturbations *a priori* and subsequently proving their asymptotic and/or best-in-class optimality.

In this work, we develop a class of mechanisms that enjoy non-asymptotic and unconditional optimality guarantees. To this end, we formulate the mechanism design problem as an infinite-dimensional distributionally robust optimization problem. We show that the problem affords a strong dual, and we exploit this duality to develop converging hierarchies of finite-dimensional upper and lower bounding problems. Our upper (primal) bounds correspond to implementable perturbations whose suboptimality can be bounded by our lower (dual) bounds. Both bounding problems can be solved within seconds via cutting plane techniques that exploit the inherent problem structure. Our numerical experiments demonstrate that our perturbations can outperform the previously best results from the literature on artificial as well as standard benchmark problems.

Wasserstein Barycenter-based Evolutionary Algorithm for the optimization of sets of points

Babacar Sow[1] Rodolphe Le Riche[1] Julien Pelamatti[2] Merlin Keller[2] Sanaa Zannane[2]

[1]Mines Saint-Etienne and CNRS LIMOS, babacar.sow@emse.fr, rodolphe.leriche@emse.fr
[2]EDF julien.pelamatti@edf.fr, merlin.keller@edf.fr, zannane.sanaa@edfenergy.com

Keywords: Wasserstein barycenter, evolutionary optimization, sets of points optimization

We consider the problem of optimizing black-box functions having sets of points as inputs (also referred to as clouds of points). Functions defined over sets are of common practical use and can be encountered when modeling turbine positions in wind farms, designs of experiments and designs of sensors or actuators networks, among others.
In this work, we propose an evolutionary algorithm for functions defined over clouds of points. The clouds of points, with a possibly varying size, are represented as discrete uniform measures. In order to characterize the cross-over and mutation operators of the evolutionary algorithm, we rely on the Wasserstein barycenter [2].
The Wasserstein barycenter being a contracting operator, we introduce two types of mutations, namely the *Boundary Mutation* and the *Random Mutation*. The Boundary Mutation is devised to correct the contracting effect of the crossover. The Random Mutation plays its standard role of an unbiased and ergodic perturbation operator to counter premature convergence.

The effects of the proposed operators on the evolutionary algorithm performance are studied on three families of test functions: wind farm proxies, the inertia of the points set and a Maximin design criterion. We implement the Wasserstein barycenter computation using the POT library [1]. These numerical tests allow us to asses the effect of the barycenter weights. Indeed, while for the crossover we adopt equal weights, several strategies to choose the weights (fixed versus random) are compared for the mutation. A second degree of freedom that is analyzed is the algorithmic arrangement of the calls to the operators. In this work, we consider and compare sequential and random calls to the various mutations. The performance of the algorithms involving Wasserstein barycenters are compared to a standard evolutionary search and a fully random search.
The obtained results show the suitability of evolutionary operators involving optimal transport for the optimization of functions defined over sets of points, as near global optima are efficiently found for all three classes of problems (wind farm, inertia, design of experiment).

References

[1] Flamary, Rémi and Courty, Nicolas and Gramfort, Alexandre and Alaya, Mokhtar Z and Boisbunon, Aurélie and Chambon, Stanislas and Chapel, Laetitia and Corenflos, Adrien and Fatras, Kilian and Fournier, Nemo and others,Pot: Python optimal transport, *Journal of Machine Learning Research*, 22(78):1–7, 2021.

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

A regularized variance-reduced modified extragradient method for stochastic hierarchical games

Shisheng Cui[1] Uday V. Shanbhag [1] Mathias Staudigl[2]

[1]Department of Industrial and Manufacturing Engineering
Pennsylvania State University, University Park, USA, udaybag@psu.edu
[2]University of Mannheim Department of Mathematics
B6, 68159 Mannheim, Germany, mathias.staudigl@uni-mannheim.de

Keywords: Hierarchical Games, Stochastic Variational Inequalities, Variance Reduction

The theory of learning in games has so far focused mainly on games with simultaneous moves. Recently, researchers in machine learning have started investigating learning dynamics in games involving hierarchical decision-making. We consider an N-player hierarchical game in which the ith player's objective comprises of an expectation-valued term, parametrized by rival decisions, and a hierarchical term. Such a framework allows for capturing a broad range of stochastic hierarchical optimization problems, Stackelberg equilibrium problems, and leader-follower games. We develop an iteratively regularized and smoothed variance-reduced modified extragradient framework for learning hierarchical equilibria in a stochastic setting. We equip our analysis with rate statements, complexity guarantees, and almost-sure convergence claims. We then extend these statements to settings where the lower-level problem is solved inexactly and provide the corresponding rate and complexity statements.

Hyperbolic deep reinforcement learning for repeated exact combinatorial optimization

Paul Strang[1,2] Zacharie Ales[2] Come Bissuel[1] Olivier Juan[1] Safia Kedad-Sidhoum[2]

[1]EDF R&D, France, name.surname@edf.fr

[2]Conservatoire national des arts et metiers, France, name.surname@cnam.fr

Keywords: Mixed-integer programming, Branch-and-bound, Reinforcement learning, Hyperbolic geometry

Mixed-integer linear programs, or MILPs, are combinatorial optimization problems defined such as $P : \{\min c^T x \mid Ax \leq b \ ; \ x \in \mathbb{N}^{|\mathcal{I}|} \times \mathbb{R}^{n-|\mathcal{I}|}\}$ with $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, n the number of variables, m the number of linear constraints and $\mathcal{I}$ the indices of integer variables. MILPs are thus linear programs (LPs) with additional integrity constraints on a subset of variables. These integrity constraints make the problem non-convex, in fact, it is $\mathcal{NP}$-hard in the general case. We are interested in repeated problems of fixed dimensions $\{P_i = (A_i, b_i, c_i)\}_{i \in N}$, which are understood as realizations of a random variable $\mathcal{P}$ following an unknown distribution $\mathcal{D} : \Omega \to \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n$.

Mixed-integer programming solvers developed over the last decades have relied on the Branch and Bound (B&B) algorithm [Wol21] to efficiently explore the space of solutions while guaranteeing the optimality of the returned solution. These solvers are based on complex heuristics fine-tuned by experts on large MIP datasets to obtain the best average performance. In the context of real-world applications, in which similar instances with slightly varying inputs are solved on a regular basis, there is a huge incentive to reduce the solving time by learning efficient tailor-made heuristics.

Directly extending the work of Marc Etheve [Eth21], this thesis proposes to discover new B&B heuristics outperforming existing solvers using reinforcement learning (RL) algorithms. In fact, many contributions have sought to reformulate the task of learning optimal branching policies as a tree Markovian decision process (MDP) [Eth+20; Sun+22; Sca+22; PLB22]. However, tree MDPs represent a challenging setup for RL algorithms: credit assignement problem, large state-action spaces yielding complex exploration and partial observability are as many predicament hindering RL agents' performance on high dimensional MIP instances. We argue that learning better state representations acknowledging the tree structure of successive tree MDP states is key to mitigate these challenges.

Contributions adapting models from Riemannian geometry and hyperbolic spaces to traditional deep learning architectures have florished over the last five years [Pen+21]. Owing to their high capacity to model data exhibiting hierarchical (tree-like) structures, hyperbolic neural networks have been applied to improve performance in computer vision, natural language processing and other tasks involving graph embeddings. In turn, we propose to harness the inductive bias of hyperbolic neural networks [GBH18; SMH21] to learn better representations of tree Markovian decision processes.

References

[GBH18] Octavian-Eugen Ganea, Gary BÃ©cigneul, and Thomas Hofmann. "Hyperbolic Neural Networks". In: (2018).

[Eth+20] Marc Etheve et al. "Reinforcement Learning for Variable Selection in a Branch and Bound Algorithm". In: *Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. Ed. by Emmanuel Hebrard and Nysret Musliu. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, pp. 176–185. ISBN: 978-3-030-58942-4. DOI: 10.1007/978-3-030-58942-4_12.

[Eth21] Marc Etheve. "Solving repeated optimization problems by Machine Learning". PhD thesis. HESAM UniversitÃ©, Dec. 3, 2021. URL: https://theses.hal.science/tel-03675471 (visited on 12/08/2022).

[Pen+21] Wei Peng et al. *Hyperbolic Deep Neural Networks: A Survey*. Feb. 17, 2021. DOI: 10.48550/arXiv.2101.04562. arXiv: 2101.04562[cs]. URL: http://arxiv.org/abs/2101.04562 (visited on 05/30/2023).

[SMH21] Ryohei Shimizu, Yusuke Mukuta, and Tatsuya Harada. *Hyperbolic Neural Networks++*. Mar. 17, 2021. DOI: 10.48550/arXiv.2006.08210. arXiv: 2006.08210[cs,stat]. URL: http://arxiv.org/abs/2006.08210 (visited on 07/25/2023).

[Wol21] Laurence A. Wolsey. *Integer programming*. Second edition. Hoboken, NJ: Wiley, 2021. ISBN: 978-1-119-60653-6.

[PLB22] Christopher W. F. Parsonson, Alexandre Laterre, and Thomas D. Barrett. *Reinforcement Learning for Branch-and-Bound Optimisation using Retrospective Trajectories*. Dec. 5, 2022. DOI: 10.48550/arXiv.2205.14345. arXiv: 2205.14345[cs]. URL: http://arxiv.org/abs/2205.14345 (visited on 02/25/2023).

[Sca+22] Lara Scavuzzo et al. *Learning to branch with Tree MDPs*. Oct. 13, 2022. DOI: 10.48550/arXiv.2205.11107. arXiv: 2205.11107[cs,math]. URL: http://arxiv.org/abs/2205.11107 (visited on 12/21/2022).

[Sun+22] Haoran Sun et al. "Improving Learning to Branch via Reinforcement Learning". In: Learning Meets Combinatorial Algorithms at NeurIPS2020. July 6, 2022. URL: https://openreview.net/forum?id=z4D7-PTxTb (visited on 12/14/2022).

Mean-field games among teams

Jayakumar Subramanian[1] Balaji Krishnamurthy[1] Akshat Kumar[2] Aditya Mahajan[3]

[1] Adobe Inc., Noida, Uttar Pradesh, India, jasubram@adobe.com, kbalaji@adobe.com
[2] Singapore Management University, Singapore, akshatkumar@smu.edu.sg
[3] McGill University, Montreal, Canada, aditya.mahajan@mcgill.ca

Keywords: Mean-field games, games among teams, approximate information state, Markov perfect equilibrium, approximate Markov perfect equilibrium

In this paper, we present a model of a game among teams. Each team consists of a homogeneous population of agents. Agents within a team are cooperative while the teams compete with other teams. The dynamics and the costs are coupled through the empirical distribution (or the mean field) of the state of agents in each team. This mean-field is assumed to be observed by all agents. Agents have asymmetric information (also called a non-classical information structure). We propose a mean-field based refinement of the Team-Nash equilibrium of the game, which we call mean-field Markov perfect equilibrium (MF-MPE). We identify a dynamic programming decomposition to characterize MF-MPE. We then consider the case where each team has a large number of players and present a mean-field approximation which approximates the game among large-population teams as a game among infinite-population teams. We show that MF-MPE of the game among teams of infinite population is easier to compute and is an ε-approximate MF-MPE of the game among teams of finite population.

Difference-of-Convex Approach to Chance-Constrained Optimal Power Flow modelling the DSO Power Modulation Lever for Distribution Networks

Ksenia Syrtseva[1,2] Paul Javal[1,2*] Welington de Oliveira[2] Sophie Demassey[2]
Hugo Morais[3] Bhargav Swaminathan[1]

[1]EDF R&D, Boulevard Gaspard Monge, Paris, France,
ksenia-k.syrtseva@edf.fr, paul.javal@enedis.fr,
bhargav.swaminathan@edf.fr
[2]Mines Paris, Université PSL, Centre de Mathématiques Appliquées (CMA), Sophia Antipolis, France,
ksenia.syrtseva@minesparis.psl.eu, paul.javal@enedis.fr
welington.oliveira@minesparis.psl.eu,
sophie.demassey@minesparis.psl.eu,
[3]Department of Electrical and Computer Engineering, INESC-ID-IST, University of Lisbon, Portugal
hugo.morais@tecnico.ulisboa.pt

Keywords: Chance-constrained optimal power flow; Power modulation and curtailment; Difference-of-Convex; Uncertainty

The increasing expansion of renewable energy sources leads to the growth of uncertainty in the distribution network operation. Short-term operational planning performed by distribution system operators should evolve to address those new operating conditions, in particular to allow the efficient utilization of different flexibility levers. In this work, the use of a chance-constrained Alternating Current Optimal Power Flow (AC-OPF) is proposed to model the operational planning problem, considering the activation of several levers such as power modulation and power curtailment. The correlation between the renewable generation profiles and loads is considered via a joint probability constraint approximated with scenarios. The main novelty of the present manuscript is the adoption of a Difference-of-Convex approach that allows to solve the obtained optimization problem without convexification or linearization of the core OPF equations. Furthermore, the approach yields a natural and embarrassingly parallelizable scenario decomposition. The method starts with a reformulation of the model as a Difference-of-Convex optimization problem, and then a proximal bundle method algorithm is applied to solve it. The proposed methodology is tested in a 33 bus distribution network with 11 different values for the safety level defining the probability constraint, ranging from 0.75 to 1.

*Now at Enedis, Direction Technique, Departement Conduite, 26 rue de la Villette 69003, Lyon

Convergence of a Two-Player Version of Macqueen's k-means Algorithm

Stéphan Sémirat[1]

[1]Université Grenoble Alpes, GAEL, France stephan.semirat@univ-grenoble-alpes.fr

Keywords: k-means algorithm, sender-receiver games, single-crossing

We study pure perfect Bayesian equilibria (PBE) in sender-receiver games with finitely many types for the sender. PBE in sender-receiver games are characterized by partitions of the set of sender's types, in which no type prefers to be pooled with types it is not currently pooled with, given the decisions of the receiver, optimally chosen at each cell of the partition. We propose a family of algorithms that converge to a PBE, in a standard class of sender-receiver games, reminiscent of Macqueen's version of the k-means algorithm [1].

In Macqueen's algorithm, at every step: (i) a single data is moved from its cell to a cell with a closer mean, and (ii) cell-contingent means are readjusted accordingly. In our sender-receiver adaptation, the sender is in charge of moving a type from its cell to a cell it prefers (i), while the receiver readjusts the cell-contingent decisions accordingly (ii). Thus, a stopping partition Π amounts to a PBE.

Since players may have different objective functions, the simple convergence argument invoked in the k-means context (monotonicity of stages (i) and (ii) with regard to a single objective function) no more holds. Instead, we employ a combinatorial argument to achieve convergence, for every initial partition and every path taken by the algorithm.

We do this for the family of utility functions $(\theta, x) \mapsto U_i^\theta(x)$ satisfying the following conditions, standard in the literature on sender-receiver games. For every player $i \in \{S, R\}$, for every type $\theta \in \Theta$, Θ finite and ordered, decision function $x \mapsto U_i^\theta(x)$ is twice continuously differentiable and strictly concave on $\mathbb{R}$, with a unique maximizing argument $x_i^*(\theta)$ such that $x_S^*(\theta) \geq x_R^*(\theta)$ (upward bias of the sender). For every $i \in \{S, R\}$, for every $\theta_2 > \theta_1$ and every $x_2 > x_1$, $U_i^{\theta_1}(x_2) - U_i^{\theta_1}(x_1) \geq 0 \implies U_i^{\theta_2}(x_2) - U_i^{\theta_2}(x_1) > 0$ (single-crossing).

The family covers, but goes much beyond, the usual univariate k-means Euclidean case, in which $U_S^\theta(x) = U_R^\theta(x) = -(\theta - x)^2$. However, our argument crucially relies on the order on types and actions.

References

[1] J. Macqueen, Some Methods for Classification and Analysis of Multivariate Observations, *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 281–297,1967

Pricing strategy for dual-channel retailing using a stochastic attraction-demand model

Minh Tam Tran[1] Yacine Rekik[2] Khaled Hadj-Hamou[1]

[1]INSA Lyon, France, minh-tam.tran@insa-lyon.fr, khaled.hadj-hamou@insa-lyon.fr
[2]ESCP Business School, France, yrekik@escp.com

Keywords: inventory, pricing, dual-channel, attraction model

Nowadays, many retailers sell their goods to customers online and through brick-and-mortar stores, creating dual-channel supply chains. Determining the optimal selling prices under customer behavior is a crucial issue in dual-channel supply chains. In this study, we propose and investigate a dual-channel pricing model, with the customers' behavior on channel choice. Motivated by market-share models, we assume a demand model involving attraction between channels: online and physical stores [1]. We consider a stochastic assumption for the demand of the potential market and price-based interaction between the demand of the two channels. In contrast to previous works, within the attraction model, the channel's stochastic demand involves a nonlinear function of prices. We permit different reactions of customers when the physical store is stock-out. The contribution of the paper is twofold: First, the results highlight that the joint concavity of the retailer's expected profit function with respect to selling prices is analytically complex to verify. We introduce a novel approach to show the existence of optimal global prices with non-linear demand and non-linear objective function. The second contribution is related to the applicability in the practice of the model. We run the model with various operational settings, and we provide managerial insights on the best pricing strategy to apply when the physical store capacity is limited.

References

[1] Gallego, G., Huh, W. T., Kang, W., & Phillips, R. (2006). Price competition with the attraction demand model: Existence of unique equilibrium and its stability. *Manufacturing & Service Operations Management*, 8(4), 359-375.

Parallel Approaches to Large-scale Stochastic Energy System Design Problems

Wim van Ackooij[1] Antonio Frangioni[2] Rafael Durbano Lobato[3]

[1]EDF R&D "Osiris", France, wim.van-ackooij@edf.fr

[2]Dipartimento di Informatica, Università di Pisa, Italia, frangio@di.unipi.it

[3]Chamber of Electric Energy Commercialization of Brazil, Brazil, rafael.lobato@gmail.com

Keywords: Large-scale optimization, Energy system design, Stochastic optimization, Parallel optimization, Lagrangian decomposition, Stochastic Dual Dynamic Programming, SMS++ software framework

We describe our latest computational experience in solving large-scale energy network design problems. This is based on a three-level decomposition approach, where a Benders-style upper level is used to manage the (continuous) design variables, the middle level is a SDDP approach to compute the long-term energy system cost (with given decisions) taking into account the yearly effect of energy storages, and the lower-level is the Lagrangian Dual (whose automatic convexification properties provide the necessary dual information to SDDP) of the weekly Unit Commitment problem. Implementing such a complex approach to an extremely-large-scale problem on significant parallel hardware (hundredths of cores, above a Terabyte of RAM) is only made possible by the use of the advanced `SMS++` [1] modelling framework. We report on the several crucial aspects of the implementation, such as how the available computing power is best subdivided between the different algorithmic levels (different scenarios in SDDP, different oracles in the Lagrangian Dual, ...), the impact of the different modelling choices in the Unit Commitment, the use of Lagrangian Dual vs. a continuous relaxation, and others. Our results show that obtaining solutions of problems of hitherto unheard-of size and complexity is possible with the combined use of sophisticated solution approaches, powerful computing hardware and advanced software techniques.

References

[1] `https://gitlab.com/smspp/smspp-project`

Per-Class Algorithm Selection for Black-Box Optimisation

Koen van der Blom[1] Carola Doerr[1]

[1]Sorbonne Université, CNRS, LIP6, France, koen.vdblom@lip6.fr, carola.doerr@lip6.fr

Keywords: Algorithm selection, black-box optimisation, feature-free algorithm selection, per-class algorithm selection.

Algorithm selection works on the premise that no single algorithm is the best on each individual instance of a problem. With this in mind, the aim is to choose the most suitable algorithm for a problem, out of a set of available algorithms. Most algorithm selection approaches use precomputed features to predict the most suitable algorithm. However, this both requires experts to design high-quality features, and incurs a cost to compute these features for each problem instance before being able to make a prediction.

Feature-free algorithm selection approaches use models like artificial neural networks to predict the most suitable algorithm, with the raw problem instance description as input. Although this avoids the need for expert designed features and their computation, a specific description for each problem instance is still needed.

In black-box optimisation neither features, nor a raw problem instance description may be available a priori. Information about the specific problem instance may only be obtained by evaluating possible solutions to it, which may be computationally expensive.

Recognising that we cannot make predictions without any cost on a per-instance basis, we take a step back and propose an algorithm selector that makes predictions on a per-class basis. Our proposed method predicts the most suitable algorithm for a class of continuous black-box optimisation problems, using only the basic properties of this class. Specifically, we take into account only the number of decision variables, and the available evaluation budget. This requires neither instance-specific information, nor any expensive computation when applying the selector.

Optimizing Variational Circuits for Higher-Order Binary Optimization

Zoé Verchère[1,2], Sourour Elloumi[1,2], Andrea Simonetto[1]

[1]Unité de Mathématiques Appliquées, ENSTA Paris,
Institut Polytechnique de Paris, 91120 Palaiseau, France andrea.simonetto@ensta-paris.fr
[2]CEDRIC, Conservatoire National des Arts et Métiers, 75003 Paris, France

Keywords: Quantum computing, combinatorial optimisation, circuit compilation

Variational quantum algorithms have been advocated as promising candidates to solve combinatorial optimization problems on near-term quantum computers. Their methodology involves transforming the optimization problem into a quadratic unconstrained binary optimization (QUBO) problem. While this transformation offers flexibility and a ready-to-implement circuit involving only two-qubit gates, it has been shown to be less than optimal in the number of employed qubits and circuit depth, especially for polynomial optimization. On the other hand, strategies based on higher-order binary optimization (HOBO) could save qubits, but they would introduce additional circuit layers, given the presence of higher-than-two-qubit gates. In this talk, we present HOBO problems and propose new approaches to encode their Hamiltonian into a ready-to-implement circuit involving only two-qubit gates. Our methodology relies on formulating the circuit design as a combinatorial optimization problem, in which we seek to minimize circuit depth. We also propose handy simplifications and heuristics that can solve the circuit design problem in polynomial time. We evaluate our approaches by comparing them with the state of the art, showcasing clear gains in terms of circuit depth.

References

[1] IBM Quantum, "Qiskit: An Open-source Framework for Quantum Computing," 2019.

[2] A. Glos, A. Krawiec, and Z. Zimborás, "Space-efficient binary optimization for variational quantum computing," *npj Quantum Information*, vol. 8, no. 1, pp. 1–8, 2022.

Table 1: Sample results for the considered approach, showcasing the benefit of our methodologies. We have imposed a 1200 s time limit. In the case of −, the solver could not improve on the warm start solution or prove its optimality in the time limit.

				Qiskit [1]		Gray code [2]		Our 1st Algo.		Our 2nd Algo.		Templates, our 3rd Algo.	
Instance	n	$\|P\|$	D	Depth	CPU [s]	Depth	CPU [s]	Depth	CPU [s]	Depth	CPU [s]	Depth	CPU [s]
qubo1	4	5	2	17	0.6	35	<0.1	**8**	481	8	0.3	12	<0.1
qubo2	6	10	2	29	0.6	70	<0.1	16	1200	**11**	6.9	20	<0.1
monomial3	3	1	3	12	0.5	15	<0.1	**8**	14	**8**	0.1	**8**	<0.1
poly3-1	4	2	3	20	0.6	30	<0.1	**9**	1200	10	1.6	16	<0.1
poly3-2	5	4	3	41	0.6	60	<0.1	26	1200	**16**	8.0	32	<0.1
monomial4	4	1	4	26	0.6	31	<0.1	**16**	1200	**16**	2.4	16	<0.1
poly4-1	5	2	4	45	0.6	62	<0.1	26	1200	**18**	1200	32	<0.1
poly4-2	6	4	4	89	0.8	124	<0.1	-	-	**47**	1200	64	<0.1
monomial5	5	1	5	56	0.6	63	<0.1	-	-	-	-	**32**	<0.1
poly5-1	6	2	5	101	0.7	126	<0.1	-	-	**57**	1200	64	<0.1
poly5-2	7	4	5	206	1.1	252	<0.1	-	-	-	-	**128**	<0.1
monomial6	6	1	6	118	1.7	127	<0.1	-	-	-	-	**64**	<0.1
poly6-1	7	2	6	228	1.3	254	<0.1	-	-	-	-	**128**	<0.1
poly6-2	8	4	6	438	1.9	508	<0.1	-	-	-	-	**256**	<0.1

A Branch-and-Cut algorithm for the Balanced Traveling Salesman Problem

Thi Quynh Trang Vo[1] Mourad Baiou[1] Viet Hung Nguyen[1]

[1]LIMOS, Université Clermont Auvergne, France.

thi_quynh_trang.vo@uca.fr, mourad.baiou@uca.fr, viet_hung.nguyen@uca.fr

Keywords: Traveling Salesman Problem Balanced optimization Mixed-Integer Programming Branch-and-Cut

The balanced traveling salesman problem (BTSP) is a variant of the traveling salesman problem, in which one seeks a tour that minimizes the difference between the largest and smallest edge costs in the tour. The BTSP, which is obviously NP-hard, was first investigated by Larusic and Punnen in 2011 [1]. They proposed several heuristics based on the double-threshold framework, which converge to good-quality solutions though not always optimal. In this paper, we design a special-purpose branch-and-cut algorithm for exactly solving the BTSP. In contrast with the classical TSP, due to the BTSP's objective function, the efficiency of algorithms for solving the BTSP depends heavily on determining correctly the largest and smallest edge costs in the tour. In the proposed branch-and-cut algorithm, we develop several mechanisms based on local cutting planes, edge elimination, and variable fixing to locate those edge costs more precisely. Other critical ingredients in our method are algorithms for initializing lower and upper bounds on the optimal value of the BTSP, which serve as warm starts for the branch-and-cut algorithm. Experiments on the same testbed of TSPLIB instances show that our algorithm can solve 63 out of 65 instances to proven optimality.

References

[1] John Larusic and Abraham P Punnen. The balanced traveling salesmanproblem. *Computers & Operations Research*, 38(5):868–875, 2011

Fast Biased Random Key Genetic Algorithm in Target Set Selection

Hoang Nguyen Vu[1] Martin Krejca[2] Benjamin Doerr[2]

[1]École Polytechnique, IP Paris, France nguyen.vu@polytechnique.edu
[2]LIX, École Polytechnique, IP Paris, France firstname.lastname@polytechnique.edu

Keywords: Evolutionary algorithms, combinatorial optimization, biased random-key genetic algorithms.

The *biased random key genetic algorithm* (BRKGA) [3] is a genetic algorithm used to find optimal or suboptimal solutions to hard combinatorial optimization problems. It has recently proven highly effective in addressing the *J-MIN target set selection* problem [4], a combinatorial optimization problem based on graph diffusion models, as shown in [5]. While the BRKGA performs well on the *target set selection* problem, it demands meticulous tuning of several parameters which is troublesome and complexifies the problem. To simplify this, we adopt an approach recommended in [1, 2] to randomly select parameters from a power-law distribution. This approach has been recently proven mathematically to be favourable in certain problems.

In this talk, we apply this approach to propose a *fast biased random key genetic algorithm* (fastBRKGA) and also a greedy heuristic *reverse Maximum-Degree Heuristic* for solving the TSS problem in large-scale social networks. Our experimental investigation finds that this new algorithm yields comparable results to traditional parameter tuning, and the new greedy heuristic improves the result significantly.

This work is part of the PGMO-funded project *Mathematical Analysis of Complex Randomized Search Heuristics* (PI: Benjamin Doerr).

References

[1] D. Antipov, M. Buzdalov, and B. Doerr. Lazy parameter tuning and control: Choosing all parameters randomly from a power-law distribution. In *Genetic and Evolutionary Computation Conference, GECCO 2021*, pages 1115–1123. ACM, 2021.

[2] B. Doerr, H. P. Le, R. Makhmara, and T. D. Nguyen. Fast genetic algorithms. In *Genetic and Evolutionary Computation Conference, GECCO 2017*, pages 777–784. ACM, 2017.

[3] J. F. Goncalves and M. G. C. Resende. Biased random-key genetic algorithms for combinatorial optimization. *Journal of Heuristics*, 6:487–525, 2011.

[4] C. Long and R. C.-W. Wong. Minimizing seed set for viral marketing. In *International Conference on Data Mining, ICDM 2011*, pages 427–436, 2011.

[5] A. L. Serrano and C. Blum. A biased random key genetic algorithm applied to target set selection in viral marketing. In *Genetic and Evolutionary Computation Conference, GECCO 2022*, pages 241–250. ACM, 2022.

The Smoothed Duality Gap as a Stopping Criterion

I. Walwil[1] O. Fercoq[2]

[1]Télécom Paris, France, iyad.walwil@telecom-paris.fr
[2]Télécom Paris, France, olivier.fercoq@telecom-paris.fr

Keywords: Convex optimization, Stopping criterion, Optimality gap and Feasibility error, Karush-Kuhn-Tucker, Smoothed Duality Gap

We optimize the running time of the primal-dual algorithms by optimizing their stopping criteria for solving convex optimization problems under affine equality constraints. We study the relations between several stopping criteria and show under which conditions they are accurate to detect optimal solutions, which means terminating the algorithm earlier with fewer iterations. Basically, we examine the accuracy of three stopping criteria: the theoretically, uncomputable in practice *"Optimality gap and Feasibility error"*, the widely-used *"Karush-Kuhn-Tucker error"* [3], and the new, widely-applicable but less well-studied *"Smoothed Duality Gap"* [5]. Assuming metric sub-regularity [1], and quadratic error bound of the smoothed gap [1], we establish that both the KKT error and smoothed duality gap provide practical upper bounds for the optimality gap, and approximate it effectively. Furthermore, we establish comparability between the KKT error and smoothed duality gap under certain conditions. Numerical experiments on quadratic programs, basis pursuit, and quadratic programs with non-negative weights corroborate these findings and show the superior stability of the smoothed duality gap over the KKT error.

References

[1] O. Fercoq. Quadratic error bound of the smoothed gap and the restarted averaged primal-dual hybrid gradient. working paper or preprint, June 2022.

[2] H. Lu and J. Yang. On the infimal sub-differential size of primal-dual hybrid gradient method and beyond, 2023.

[3] R. Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1970.

[4] L. V. Stephen Boyd. *Convex Optimization*. Cambridge University Press, 2004.

[5] Q. Tran-Dinh, O. Fercoq, and V. Cevher. A smooth primal-dual optimization framework for nonsmooth composite convex minimization, 2018.

Randomized Assortment Optimization

Zhengchao Wang[1] Heikki Peura[1] Wolfram Wiesemann[1]

[1]Imperial College Business School, UK,
z.wang19@imperial.ac.uk, h.peura@imperial.ac.uk, ww@imperial.ac.uk

Keywords: Randomization, Robust, Assortment Optimization, Multinomial Logit Model, Markov Chain Model, Preference Ranking Model

When a firm selects an assortment of products to offer to customers, it uses a choice model to anticipate their probability of purchasing each product. In practice, the estimation of these models is subject to statistical errors, which may lead to significantly suboptimal assortment decisions. Recent work has addressed this issue using robust optimization, where the true parameter values are assumed unknown and the firm chooses an assortment that maximizes its worst-case expected revenues over an uncertainty set of likely parameter values, thus mitigating estimation errors. In this paper, we introduce the concept of *randomization* into the robust assortment optimization literature. We show that the standard approach of deterministically selecting a single assortment to offer is not always optimal in the robust assortment optimization problem. Instead, the firm can improve its worst-case expected revenues by selecting an assortment randomly according to a prudently designed probability distribution. We demonstrate this potential benefit of randomization both theoretically in an abstract problem formulation as well as empirically across three popular choice models: the multinomial logit model, the Markov chain model, and the preference ranking model. We show how an optimal randomization strategy can be determined exactly and heuristically. Besides the superior in-sample performance of randomized assortments, we demonstrate improved out-of-sample performance in a data-driven setting that combines estimation with optimization. Our results suggest that more general versions of the assortment optimization problem—incorporating business constraints, more flexible choice models and/or more general uncertainty sets—tend to be more receptive to the benefits of randomization.

Decentralized Smart Charging of Large-Scale EV Fleets using Adaptive Multi-Agent Multi-Armed Bandits

Sharyal Zafar[1] Raphaël Féraud[2] Anne Blavette[3] Guy Camilleri[4] Hamid Ben Ahmed[5]

[1]SATIE Lab, ENS Rennes, France, sharyal.zafar@ens-rennes.fr
[2]Orange Labs, Lannion, France, raphael.feraud@orange.com
[3]CNRS & SATIE Lab, ENS Rennes, France, anne.blavette@ens-rennes.fr
[4]IRIT Lab, UPS Toulouse, France, guy.camilleri@irit.fr
[5]SATIE Lab, ENS Rennes, France, benahmed@ens-rennes.fr

Keywords: Active distribution networks, Electric vehicles, Multi-agent reinforcement learning, Combinatorial multi-armed bandits, Decentralized optimization

As solar photovoltaics (PVs) and electric vehicles (EVs) integrate into electrical grids, challenges like network congestion and peak demand emerge. Traditional solutions require costly infrastructure investments. However, smart EV charging offers an elegant alternative. Existing solutions include centralized, hierarchical, and decentralized systems. However, scalability and real-time operation may pose challenges for centralized and hierarchical architectures. To address these issues comprehensively, we propose a fully decentralized smart charging system, ensuring scalability, real-time operation, and data privacy.

In recent years, decentralized smart charging solutions utilizing standard reinforcement learning (RL) algorithms have gained prominence. However, the integration of multi-level constraints and objectives, encompassing both prosumer-specific local considerations as well as broader network-global considerations, remains a complex challenge. This complexity arises from the absence of predefined reward signals in the majority of smart-grid applications. Furthermore, standard RL algorithms employing deep learning-based approximations may exhibit slower convergence when compared to the multi-armed bandit class of RL algorithms, especially in scenarios where a perfect oracle is unavailable. Thus, we introduce a novel adaptive multi-agent system employing combinatorial 2-armed bandits with Thompson Sampling to minimize EV charging costs, considering variable electricity prices and PV energy production uncertainty, while ensuring fairness among prosumers. Validation includes a large-scale simulation with over 10,000 EVs, comparing our approach to "dumb" charging and centralized mixed-integer linear programming optimization (computed only at a small-scale). Evaluation criteria encompass cost optimization, constraint satisfaction, fairness, and solution speed. Our results demonstrate the effectiveness of our decentralized approach [1].

References

[1] S. Zafar, R. Féraud, A. Blavette, G. Camilleri, and H. Ben Ahmed, Decentralized Smart Charging of Large-Scale EVs using Adaptive Multi-Agent Multi-Armed Bandits, *CIRED*, 2023, arXiv preprint, arXiv:2307.10704. 2023 Jul 20.

Minimizing total completion time and makespan for a multi-scenario bi-criteria parallel machine scheduling problem

Zhang Xiechen[1] Angel Eric[1] Chu Feng[1] Regnault Damien[1]

[1]Laboratory IBISC, Univ-Evry, Universite Paris-Saclay, Paris, France
xiechen.zhang@universite-paris-saclay.fr,
{eric.angel, feng.chu, damien.regnault}@univ-evry.fr

Keywords: Parallel machine Scheduling, Multi-objective optimization, Uncertainty, Makespan, Total completion time, Approximation algorithms

Multi-criteria scheduling problems face enormous challenges under uncertainty in academic and industrial fields. Despite the significant interest in practice, this area is still underinvestigated theoretically. In this study, we investigate two new parallel machine scheduling problems to minimize makespan and total completion time, respectively, in which the number of scenarios can be infinite. The complexities of the problems are analyzed. An exact and a $(1+\epsilon)$-approximation algorithms are developed respectively for total completion time and makespan minimization.

Then, we address a novel bi-objective parallel machine scheduling problem with infinite scenarios. Based on the obtained results for the above single objective problem, the complexity of the problem is discussed, and two approximation algorithms are designed to find approximate Pareto sets with distinct performance guarantees for a Pareto optimal set for multi-objective optimization problems under uncertainty defined in the literature [1]. We also consider some special cases and show that the problem can be transformed into a bi-objective scheduling problem without scenarios for those cases.

References

[1] Engau, Alexander, and Devon Sigler. "Pareto solutions in multicriteria optimization under uncertainty." *European Journal of Operational Research* 281.2 (2020): 357-368.

Advanced Optimization Techniques for Grid Design and Real-Time Operations in Power Networks

Alessandro Zocca Erica van der Sar Leon Lan Sandjai Bhulai

Vrije Universiteit Amsterdam, a.zocca@vu.nl

In transmission power networks, power flows and network topology are deeply intertwined due to power flow physics. With increasing energy demands and the rapid integration of volatile renewable energy sources like wind and solar, ensuring safe and reliable operation of these networks has grown increasingly complex. A crucial yet underexplored avenue for enhancing network robustness lies in the strategic usage of line-switching and bus-switching actions. These actions enable remote activation or deactivation of transmission lines and local adjustments to network topology but come with the challenge of identifying optimal configurations within an exponentially large search space.

To navigate this complex landscape, advanced mathematical tools are needed to assist decision-makers in the real-time identification of robust and resilient network configurations. In this presentation, I will outline two state-of-the-art approaches: one geared towards grid design and planning, and the other aimed at real-time grid operations and control. The first approach utilizes a mixed-integer linear programming (MILP) framework, drawing insights from recent tree partition literature. Its twofold objective is to alleviate network congestion while at the same time creating a more hierarchical structure that enhances the system's resilience against cascading failures. The second approach focuses on scalability and employs multi-agent reinforcement learning (MARL) techniques to effectively navigate the combinatorial action space. This new model-free approach is inspired by the 2019 "Learning to Run a Power Network" challenge, initiated by the French network operator RTE to get researchers from different fields involved with the task of using reinforcement learning in the context of power grid control. Together, these approaches offer complementary pathways for exploiting the intrinsic flexibility of modern electrical infrastructures, thereby facilitating a transition towards more sustainable and resilient power systems.