Book of abstracts

PGMO DAYS 2025

November 18th and 19th, 2025

EDF Labs Paris-Saclay

FONDATION MATHÉMATIQUE JACQUES HADAMARD · EDF · INSTITUT POLYTECHNIQUE DE PARIS · université PARIS-SACLAY · Inria · CNRS · Société Mathématique de France · SMAI · SFdS SOCIETE FRANÇAISE DE STATISTIQUE · ROADEF · RT Réseau thématique Optimisation

Preface

This volume contains the extended abstracts of the talks presented at the conference PGMO Days 2025 held on November 18th – 19th, 2025 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, and SMF. We are grateful to the plenary speakers and to the organizers of the invited sessions.

November 10, 2025 The organizers of the PGMO Days
Palaiseau

Organization Committee

Scientific organizers	
Wim van Ackooij	EDF
Claudia d'Ambrosio	CNRS LIX, École polytechnique
Catherine Bonnet	Inria and LSS/Centrale-Supelec
Antonin Chambolle	CEREMADE, CNRS, PSL
Sourour Elloumi	UMA, Ensta Paris and CEDRIC, Cnam *(coordinator of PGMO)*
Olivier Fercoq	Télecom Paris
Stéphane Gaubert	INRIA and CMAP, Ecole Polytechnique *(deputy coordinator of PGMO)*
Vianney Perchet	ENSAE Paris *(scientific committee chair)*
Local organization	
Magali Le Chaponnier	Fondation Mathématique Jacques Hadamard
Nadège Menasse	EDF
Mélanie Mihoc	EDF

Scientific Program Committee

Grégoire Allaire	CMAP, Ecole Polytechnique
Christian Artigues	LAAS, CNRS Toulouse
Walid Ben Ameur	SAMOVAR, Télécom SudParis
Claire Boyer	LMO, Université Paris-Saclay
Luce Brotcorne	INOCS, INRIA Lille
Guillaume Carlier	CEREMADE, Université Paris-Dauphine
Antonin Chambolle	CEREMADE, CNRS, PSL
Johanne Cohen	CNRS, LISN, Université Paris-Saclay
Michel De Lara	ENPC, CERMICS
Sophie Demassey	CMA, MINES ParisTech
Frédéric Jean	UMA, ENSTA
Safia Kedad-Sidhoum	CEDRIC, CNAM
Franck Iutzeler	IMT Université de Toulouse
Edouard Pauwels	Toulouse School of Economics
Vianney Perchet	CREST, ENSAE Paris, Chairman
Olivier Spanjaard	LIP6 CNRS, Sorbonne Université

Organizers of Invited Sessions

Miguel F. Anjos: *Green energy applications of optimization*
Pierre-Cyril Aubin and Michel De Lara: *Beyond Euclidean Convexity: Methods and Algorithms*
Térence Bayen: Optimal Control and its Applications
Matias R. Bender and Elias Tsigaridas: Nonlinear Algebra and its Application
Anne Blavette: Fast and scalable computation for power systems: accelerated security assessment and decentralised management of large-scale smart grids
Olivier Bokanowski and Carlos Esteve-Yagüe: Neural Network approximations and related techniques for high-dimensional PDE
Guillaume Carlier, Luca Nenna, and Paul Pegon: New trends in Optimal Transport and Applications
Danila Cherkashin, Alexey Gordeev, and Yana Teplitskaya: Recent advances in length-min problems In Euclidean spaces
Sophie Demassey and Safia Kedad-Sidhoum: Discrete Optimization and Energy
Guillaume Garrigos: Advances in the convergence analysis of gradient methods
Frédéric Jean: Optimal geometric control
Dylan Laplace Mermoud: Quantum computing and combinatorial optimization
Juan Peypouquet: Dynamic approaches to saddle point problems
Titus Pinta: Optimization on Non-linear and Metric spaces
Mathias Staudigl: Recent advances in Multi Objective Optimization: Theory meets Practice

PGMO 2025 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 defended their PhD in France, during the previous civil year.

The two 2025 PGMO PhD prizes were awarded to

- ***Armand Gissler**, , for his PhD at Ecole polytechnique, under the supervision of Anne Auger on "Linear convergence of evolution strategies with covariance matrix adaptation"*
- ***Maher Mallem**, for his PhD at LIP6 (Sorbonne Université), under the supervision of Claire Hanen on "Complexité Paramétrée et Nouveaux Schémas Enumeratifs Efficaces pour le RCPSP"*

The 2025 PhD prize committee was chaired by Walid Ben-Ameur (SAMOVAR, Télécom SudParis). It was composed of the following researchers:

- *Catherine Bonnet (L2S, Inria)*
- *Elsa Cazelles (CNRS, IRIT Toulouse)*
- *Aris Daniilidis (VADOR, TU Vienne)*
- *Charles Dossal (INSA Toulouse)*
- *Nacima Labadie (ICD-LOSI, Université Technologique de Troyes)*
- *Céline Lévy-Leduc (LPSM, Université Paris Cité)*
- *Panayotis Mertikopoulos (CNRS LIG, Université Grenoble Alpes)*
- *Vincent T'kindt (LIFAT, Université de Tours)*

Table of Contents

Training Neural Networks at Any Scale

Volkan Cevher[1]

[1] Laboratory for Information and Inference Systems, EPFL, Switzerland volkan.cevher@epfl.ch

Keywords: Deep Learning, Large scale, Scaling

At the heart of deep learning's transformative impact lies the concept of scale–encompassing both data and computational resources, as well as their interaction with neural network architectures. Scale, however, presents critical challenges, such as increased instability during training and prohibitively expensive model-specific tuning. Given the substantial resources required to train such models, formulating high-confidence scaling hypotheses backed by rigorous theoretical research has become paramount.

To bridge theory and practice, the talk explores a key mathematical ingredient of scaling in tandem with scaling theory: the numerical solution algorithms commonly employed in deep learning, spanning domains from vision to language models. We unify these algorithms under a common master template, making their foundational principles transparent. In doing so, we reveal the interplay between adaptation to smoothness structures via online learning and the exploitation of optimization geometry through non-Euclidean norms. Our exposition moves beyond simply building larger models–it emphasizes strategic scaling, offering insights that promise to advance the field while economizing on resources.

Exact approaches for multi-objective binary quadratic optimization

Marianna De Santis[1]

[1] DINFO - University of Florence, Firenze, Italy marianna.desantis@unifi.it

Keywords: Multi-Objective optimization, Binary quadratic programming, Convex reformulation

Multi-objective mixed-integer nonlinear programming (MOMINLP) provides a powerful framework for modeling complex real-world decision problems. Practical applications often involve multiple, conflicting objectives and require binary or general integer variables to capture logical constraints or discrete decisions. Solving MOMINLP problems entails identifying the set of efficient solutions - i.e., solutions where no objective can be improved without deteriorating at least one other.

An efficient integer assignment refers to a fixing of the integer variables such that there exists at least one efficient solution for the corresponding continuous subproblem. In many cases, correctly solving MOMINLPs requires exploring a large number of such assignments - potentially all integer-feasible ones - rendering full enumeration unavoidable. This significantly complicates the solution process compared to the single-objective case and poses substantial challenges for algorithm development.

In this talk, we review essential tools recently proposed for developing branch-and-bound methods for multi-objective mixed-integer nonlinear optimization, and we explore approaches specifically tailored to problems with quadratic objective functions. Building on the well-established technique of quadratic convex reformulation - originally developed for single-objective binary quadratic programs - we extend this methodology to the multi-objective setting. We propose a branch-and-bound algorithm where lower bound sets are obtained from appropriately defined convex quadratic subproblems. Computational experiments on multiobjective k-item Quadratic Knapsack and multiobjective Max-Cut instances demonstrate the effectiveness of our approach.

Risk-Averse PDE-Constrained Optimization: Theory, Algorithms, and Statistical Foundations

Thomas M. Surowiec[1]

[1] Simula Research Laboratory, Oslo, Norway thomasms@simula.no

Keywords: Risk-averse optimization, PDE, Uncertainty

Over the past few decades, risk-averse optimization has become an important framework for decision-making under uncertainty in systems governed by partial differential equations (PDEs). Uncertainty in the PDE parameters creates both theoretical and computational challenges. Addressing these challenges requires techniques from functional analysis, numerical optimization, and statistics. In this plenary, I will survey recent developments in risk-averse PDE-constrained optimization, with an emphasis on theory and algorithms as well as stability and asymptotic analysis.

The talk will address three central themes. First, we develop rigorous formulations that incorporate risk measures into PDE-constrained settings to capture the behavior of risk-averse decision makers. Second, we consider algorithmic strategies for handling nonsmooth risk measures that make large-scale computations both tractable and robust. Third, we turn our attention to results on stability and asymptotic behavior that characterize solution sensitivity to perturbations in data, discretization, and sampling, and provide statistical guarantees for risk-averse models.

Together, these directions show how risk-averse optimization enriches the mathematical landscape of PDE-constrained problems and offers a principled framework for reliable decision-making in the presence of uncertainty.

Learning to Optimize

Pascal Van Hentenryck[1]

[1] Georgia Institute of Technology, Atlanta, United States pvh@gatech.edu

Keywords: Machine Learning, Optimization, Industrial problems

In many industry settings, the same optimization problem is solved repeatedly for instances taken from a distribution that can be learned or forecasted. Indeed, such parametric optimization problems are ubiquitous in applications over complex infrastructures such as electrical power grids, supply chains, manufacturing, and transportation networks. The scale and complexity of these applications have grown significantly in recent years, challenging traditional optimization approaches. This talk studies how to speed up these parametric optimization problems to meet real-time constraints present in many applications. covering concepts such as primal and dual optimization proxies, learning to optimize, contextual optimization, and decision-focused learning. The methodologies are highlighted on industrial problems in grid optimization, end-to-end supply chains, logistics, and transporation systems.

Linear Convergence of Evolution Strategies with Covariance Matrix Adaptation

Armand Gissler[1] Anne Auger[2] Nikolaus Hansen[2]

[1]SIERRA, Inria & ENS, Paris, France, firstname.lastname@inria.fr
[2]RANDOPT, Inria & École polytechnique, Palaiseau, France, firstname.lastname@inria.fr

Keywords: CMA-ES, Linear convergence, Markov chains

In the fields of black-box optimization, CMA-ES is the state-of-the-art algorithm among the evolution strategies. Although it has found many applications for more than 20 years, a proof of its convergence remained an open question. In this thesis [1], we provide theoretical guarantees of linear convergence of CMA-ES when minimizing an ellipsoidal objective function. Moreover, we prove that second-order information of a convex-quadratic function is learnt since the covariance matrix approximates the inverse Hessian matrix. To establish our results, we normalize the states of the algorithm and define a Markov chain when the objective function is scaling-invariant. The stability of this Markov chain proves then the convergence of CMA-ES.

References

[1] A. Gissler, Linear convergence of evolution strategies with covariance matrix adaptation, *Institut Polytechnique de Paris*, 2024, <NNT : 2024IPPAX115>, <tel-04935192>.

Parameterized Complexity and New Efficient Enumerative Schemes for RCPSP

Maher Mallem[1]

[1]Inria, CNRS, ENS de Lyon, Université Claude Bernard Lyon 1, LIP, UMR 5668,
maher.mallem@ens-lyon.fr

Keywords: Algorithm design, scheduling, parameterized complexity

Most scheduling problems are strongly NP-hard and have required intricate heuristics to be solved in practice. While yielding results short term, such heuristics rarely help to identify what fundamentally makes these problems difficult. In contrast, over the past ten years, parameterized complexity has become an increasingly popular framework to study the computational complexity of scheduling problems and find new efficient algorithms for them [2]. Such analysis typically highlights similarities between seemingly different problems and facilitates strategy transfers. This also proves useful whenever one would like to enrich a problem - for example when adding precedence constraints or going from equal-length jobs to jobs of arbitrary time length.

In this talk, we illustrate this framework by studying the parameterized complexity of several subproblems of the Resource-Constrained Project Scheduling Problem (RCPSP), possibly enhanced with job time windows and precedence delays. With precedence delays we consider parameter ℓ_{max} - i.e. the maximum delay value appearing in the input. We show that scheduling unit-time jobs with precedence delays on a single machine is hard even with small ℓ_{max}. This suggests that another property of the problem has to be bounded in order to deal with precedence delays. This motivates the integration of job time windows to problems featuring precedence delays in order to broaden available parameter choices. We consider two parameters which have shown recent success in the literature ; namely the slack - i.e. the maximum difference between the time window length of a job and its processing time - and the pathwidth - i.e. the maximum number of overlapping job time windows at any time. We obtain several fixed-parameter tractable algorithms and set multiple hardness results with to respect to these parameters, taken separately or in combination with parameter ℓ_{max}. The presented results are available in Sections 3 to 5 of the speaker's Ph.D. thesis [1].

References

[1] Maher Mallem. *Parameterized Complexity and new Efficient Enumerative Schemes for RCPSP*. PhD thesis, Sorbonne Université, 2024.

[2] Matthias Mnich and René van Bevern. Parameterized complexity of machine scheduling: 15 open problems. *Computers & Operations Research*, 100:254–261, December 2018.

A Parametric Optimization Approach to Comparison Functions

Assalé Adjé[1]

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

Keywords: Parametric Optimization, Continuity of Optimal Value, Comparison Functions, Stability

Lyapunov functions serve as constructive indicators of the stability of dynamical systems. Recall that for a given dynamical system, Lyapunov functions are (i) positive definite functions; and (ii) strictly decreasing along the trajectories in time. Then, we can formulate these two conditions in terms of algebraic inequalities. In a few advantageous situations (when the dynamics is linear, polynomial, piecewise polynomial, etc.), we can even implement them into optimization solvers. However, it is not clear whether these two conditions imply stability. This becomes clear when introducing *comparison functions* (see, e.g., [2]), which aim to compare functions with norms. More precisely, Hahn [1] has shown that, for a given continuous function h on $\mathbb{R}^d$, (i) h is positive definite on X containing 0 in its interior; if and only if (ii) there exist $\alpha_1, \alpha_2 \in \mathcal{K}$ such that for all x in a closed ball centered at 0 included in X:

$$\alpha_1(\|x\|) \leq h(x) \leq \alpha_2(\|x\|),$$

increasing,

$$s \mapsto \inf_{s \leq g(x)} f(x) \quad \text{and} \quad s \mapsto \sup_{g(x) \leq s} f(x) \text{ where } g = \|\cdot\|$$

In this presentation, we propose an in-depth examination of the membership of α_1, α_2 to $\mathcal{K}$ (or $\mathcal{K}_\infty$ for a general function g. This analysis relies on parametric optimization tools. The study of the lower and upper continuity of $\mathbb{R} \ni s \mapsto \sup_{g(x) \leq s} f(x)$ has already been explored in the literature and can be treated from Berge's maximum theorem. Positive definiteness and coercivity are not classical questions in parametric optimization theory. Continuity questions about $\mathbb{R} \ni s \mapsto \inf_{s \leq g(x)} f(x)$ are not directly addressed. Some results involving the continuity of parametric minimization problems require the convexity of the constraints set, and thus do not apply in our case. In this presentation, we also discuss the potential advantages of this general comparison approach in terms of approximating optimization problems.

References

[1] W. Hahn. *Stability of Motion*, volume 138. Springer, 1967.

[2] C. M. Kellett. A Compendium of Comparison Function results. *Math. Control. Signals Syst.*, 26(3):339–374, 2014.

Dynamic Algorithm Configuration for Next-Generation Optimization

Imène Ait Abderrahim[1,2]

[1]Sorbonne University, CNRS, LIP6, F-75005 Paris, France imene.ait-abderrahim@lip6.fr
[2]Khemis Meliana University, Algeria i.aitabderrahim@univ-dbkm.dz,

Keywords: Dynamic algorithm configuration, Optimization, Self-adapting algorithms

The process of tuning parameters is called *algorithm configuration* (AC), and has been widely studied [4]. In recent years, algorithm configuration has become a critical component in advancing the performance of optimization, machine learning, and AI systems. Traditionally, algorithms rely on static, one-size-fits-all configurations, which often fail to adapt to diverse problem landscapes or dynamic environments. This limitation has motivated research into dynamic algorithm configuration (DAC), where parameters, heuristics, or even whole algorithms are adapted online based on problem-specific features, search progress, or environmental feedback. Many works have shown that adjusting parameters during search can greatly improve performance [2, 1]. One application of DAC is my ongoing work that consists on applying the hyperconfigurable algorithm search for the combinatorial problem of multi port continuous berth allocation problem (MPCBAP) [3]. In this work, the following contributions are made: (1) we propose a dynamically configured ALNS method that we call hyper-configurable ALNS (HCALNS) and (2) we apply and evaluate the HCALNS algorithm on the MPCBAP and present its performance on a benchmark of instances from the literature. The results show that the HCALNS out- performed the ALNS algorithm and was able to define a new state-of-the-art for the large instances of MPCBAP. This demonstrates the effectiveness of using self-adjusted parameters to solve the problem. This adaptability leads to better performance, especially in complex or evolving scenarios, as the algorithm finetunes its behavior in real-time.

References

[1] S. Adriaensen, A. Biedenkapp, G. Shala, N. Awad, T. Eimer, M. Lindauer, and F. Hutter. Automated dynamic algorithm configuration. *Journal of Artificial Intelligence Research*, 75:1633–1699, 2022.

[2] R. Battiti and G. Tecchiolli. The reactive tabu search. *ORSA journal on computing*, 6(2):126–140, 1994.

[3] B. Martin-Iradi, D. Pacino, and S. Ropke. An adaptive large neighborhood search heuristic for the multi-port continuous berth allocation problem. *European Journal of Operational Research*, 2024.

[4] E. Schede, J. Brandt, A. Tornede, M. Wever, V. Bengs, E. Hüllermeier, and K. Tierney. A survey of methods for automated algorithm configuration. *Journal of Artificial Intelligence Research*, 75:425–487, 2022.

Optimal Operation of a Battery System Considering Uncertainties on Energy Prices: a Least-Square Monte Carlo Approach

Asja Alic[1*], Vincenzo Trovato[1], Adrien Seguret[2]

[1] Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy – *asja.alic@unitn.it
[2] OSIRIS Department, EDF Lab, Palaiseau, France

Keywords: battery system, cyclic degradation, Longstaff-Schwartz, stochastic optimization, energy prices uncertainties

Battery Energy Storage Systems (BESS) are expected to play a key role in the fight towards climate change. Nevertheless, despite their flexibility and decreasing investment costs, BESS still face high revenue risks. On the one hand, these assets are characterized by inevitable capacity degradation phenomena pursuant to frequent cyclic actions. If not properly considered, this loss in energy capacity may impede the assets to effectively operate in the energy markets and ultimately effect their financial viability. On the other hand, BESS that rely on merchant revenues (e.g. arbitrage in the Day-Ahead energy markets) are also subject to market uncertainties, which introduce further challenges to their optimal operation and scheduling.

Within this context, this work extends a deterministic scheduling model –developed by EDF researches - to a stochastic framework in order to account for uncertainties on energy prices. The original model, formulated with a Dynamic Programming–Mixed Integer Linear Programming approach, associates a "cost" $c \in \mathcal{C}$ for each optimal charging/discharging cycle to maximize the revenues while accounting for non-linear the cyclic degradation.

When accounting for uncertainties on the Day-Ahead energy prices, it is possible to formulate the *Bellman Equation* of the optimization problem at a generic time-step t, for a generic value of the energy capacity E_t and at energy price profile λ [€/MWh] as in (1):

$$V(t, E_t, \lambda) = \max_{c \in \mathcal{C}} \{R(t, E_t, p_t^*(c, \lambda), \lambda) + \mathbb{E}[V(t+1, f(t, E_t, p_t^*(c), \lambda_{t+1}) | \lambda_t = \lambda)]\} \quad (1)$$

Subject to:

$$f\big(t, E_t, p_t^*(c)\big) = E_t \cdot e^{-k \cdot N_t^{eq}} \quad (2) \qquad \lambda_{t+1} \sim Pr_{t+1}(\cdot \,|\lambda_t) \quad (3)$$

$$p_t^*(c_t) = arg \max_{p_t \in \mathcal{P}} \{\lambda_t \cdot (p_t^d - p_t^c) \cdot \Delta t - c \cdot N_t^{eq}\} \quad (4) \qquad N_t^{eq} = (p_t^d + p_t^c) \cdot \Delta t / (2 \cdot \hat{E}) \quad (5)$$

In (1), $R(\cdot)$ [€] indicates the immediate arbitrage revenues. Notably, $\mathbb{E}[\cdot \mid \lambda_t = \lambda]$ denotes the *Conditional Expected Value*, where λ_t is a stochastic process modelling the energy price, while $f\big(t, E_t, p_t^*(c)\big)$ indicates the degraded energy capacity E_{t+1} at the next stage pursuant to cyclic actions. Finally, $Pr_{t+1}(\cdot \,|\lambda_t)$ in (3) is the conditional probability measure of the random variable λ_{t+1} while the variable $p_t^*(c)$ [MW] in (1) is the optimal control variable obtained from the resolution of the problem in (4) subject to standard operational constraints (e.g. energy balance, power limits) and to (5).

In order to compute $\mathbb{E}[\cdot]$, the Longstaff-Schwartz method (known also as Least-Square Monte Carlo approach) [1] is adopted. Specifically, $\mathbb{E}[\cdot]$ can be approximated with a linear function of $R[t](\lambda)$, which indicates the maximum revenues that a BESS without degradation can obtain at time-step t at a given price λ.

Simulations have been conducted over a representative planning horizon of ten years with monthly stages and an hourly resolution. The energy prices profiles have been obtained from a One-Factor model implemented by researches from EDF and consider the Italian Day-Ahead energy prices.

Acknowledgements: Alic Asja is supported by the Italian National PhD in Photovoltaics. This research benefited from the support of FMJH Program PGMO n. 2024-0013.

[1] F.A. Longstaff and E.S. Schwartz, "Valuating American Option by Simulation: A Simple Least-Square Approach," *Review on Financial Studies,* vol. 14, no. 1, pp. 113-147, 2001.

[2] A. Boogert and C. de Jong, "Gas Storage Valuation Uing a Monte Carlo Method," *The Journal of Derivatives,* 2008.

Computing the Congestion Phases of Dynamical Systems with Priorities and Application to Emergency Departments

Xavier Allamigeon[1] Pascal Capetillo[1] Stéphane Gaubert[1]
[1]CMAP, Inria (TROPICAL team), France
xavier.allamigeon@inria.fr, pascal.capetillo@inria.fr, stephane.gaubert@inria.fr

Keywords: Performance evaluation, Emergency departments, Piecewiselinear Dynamics, Petri nets with priorities, Polyhedral Computation.

Medical emergency departments are complex systems in which patients must be treated according to priority rules based on the severity of their condition. We develop a model of emergency departments using Petri nets with priorities, described by nonmonotone piecewise linear dynamical systems. The collection of stationary solutions of such systems forms a "phase diagram", in which each phase corresponds to a subset of bottleneck resources (like senior doctors, interns, nurses, consultation rooms, etc.). Since the number of phases is generally exponential in the number of resources, developing automated methods is essential to tackle realistic models. We develop a general method to compute congestion diagrams. A key ingredient is a polynomial time algorithm to test whether a given "policy" (configuration of bottleneck tasks) is achievable by a choice of resources. This is done by reduction to a feasibility problem for an unusual class of lexicographic polyhedra. Furthermore, we show that each policy uniquely determines the system's throughput. We apply our approach to a case study, analyzing a simplified model of an emergency department from Assistance Publique - Hôpitaux de Paris.

Stability and Optimization of Perturbed Chemostat Systems

C. Alvarez-Latuz[1] T. Bayen[2] J. Coville[2]

[1,2]Avignon Université, Laboratoire de Mathématiques d'Avignon, France
claudia.alvarez-latuz@alumni.univ-avignon.fr, terence.bayen@univ-avignon.fr
[2]INRAE, Biostatistique et Processus Spatiaux, France jerome.coville@inrae.fr

Keywords: Chemostat system, Perturbed dynamics, Stability, Control, Numerical optimization.

The chemostat system is a classical model for microbial evolution in controlled environments, famously predicting the competitive exclusion principle where only one species survives when competing for a single nutrient. Yet, real-world observations often defy this principle, showing coexistence of multiple species (though low-concentration). First, we demonstrate that this discrepancy can be explained by introducing a perturbation term into the system dynamics. Under biologically reasonable assumptions on the perturbation, we prove that the system remains globally asymptotically stable around a coexistence steady-state (see [1]). This result is based on the application of the Malkin-Gorshin Theorem [4] and Smith and Waltman results on perturbed steady-states [5]. In this setting, we also address the problem of simultaneously optimizing two objective functions: the productivity of species, which has potential applications in industrial contexts, and biodiversity, quantified through the Simpson index. Our numerical results highlight how optimizing species productivity while maintaining a predefined threshold for biodiversity alters optimal strategies compared to those designed for productivity only as in [2, 3].

References

[1] C. Alvarez-Latuz, T. Bayen, J. Coville: *Global stability of perturbed chemostat systems*, in press, 2025, arXiv:2501.08011.

[2] T. Bayen, H. Cazenave-Lacroutz, J. Coville: *Stability of the chemostat system including a linear coupling*, Discrete Contin. Dyn. Syst. Ser. B, vol. 28, 3, pp. 2104–2129, 2023.

[3] T. Bayen, J. Coville, F. Mairet : *Stabilization of the chemostat system with mutations and application to microbial production*, Optimal Control Appl. Methods, vol.4, 6, pp. 3342–3360, 2023.

[4] T. Sari, B.S. Kalitin: *B-stability and its applications to the Tikhonov and Malkin-Gorshin theorems*, Differential Equations, Kluwer Academic Publishers-Plenum Publishers, Edinburgh, vol. 37, pp. 11–16, 2001.

[5] H.L. Smith, P. Waltman: *Perturbation of a globally stable steady-state*, Proc. Amer. Math. Soc., vol. 127, 2, pp. 447–453, 1999.

Optimal Electric Vehicle Charging with Dynamic Pricing, Customer Preferences and Power Peak Reduction

Miguel F. Anjos[1] Luce Brotcorne[2] Gaël Guillot[3]

[1]School of Mathematics and Maxwell Institute for Mathematical Sciences,
University of Edinburgh, Edinburgh, United Kingdom anjos@stanfordalumni.org
[2]INRIA Lille Nord-Europe, Lille, France luce.brotcorne@inria.fr
[3] Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France
gael.guillot@centralelille.fr

Keywords: electric vehicle charging, dynamic pricing, bilevel optimization, preference list, reserve price

We consider a provider of electric vehicle charging stations that operates a network of charging stations and use time varying pricing to maximize profit and reduce the impact on the electric grid. We propose a bilevel model with a single leader and multiple disjoint followers. The customers (followers) make decisions independently from each other. The provider (leader) sets the price of charging for each station at each time slot, and ensures there is enough energy to charge. The charging choice of each customer is represented by a combination of a preference list of (station, time) pairs and a reserve price. The proposed model thus accounts for the heterogeneity of customers with respect to price sensitivity and charging preferences. We define a single-level reformulation based on a reformulation approach from the literature on product line optimization, and we report computational results that highlight the efficiency of the new reformulation and the potential impact of our approach for reducing peaks on the electricity grid.

References

[1] M.F. Anjos, L. Brotcorne, and G. Guillot, Optimal Electric Vehicle Charging with Dynamic Pricing, Customer Preferences and Power Peak Reduction, *INFOR: Information Systems and Operational Research* 63 (2025), 439–458, https://doi.org/10.1080/03155986.2025.2463189.

Evolutionary Algorithms Are Robust to Noise out of the Box

Denis Antipov[1] Benjamin Doerr[2]

[1]LIP6, CNRS, Sorbonne Université, France, denis.antipov@lip6.fr
[2]Laboratoire d'Informatique (LIX), CNRS, École Polytechnique,
Institut Polytechnique de Paris doerr@lix.polytechnique.fr

Keywords: Evolutionary Algorithms, Theory, Runtime Analysis, Noisy Optimization

Randomized search heuristics (RHSs) are generally believed to be robust to noise. However, almost all mathematical analyses on how RSHs cope with a noisy access to the objective function assume that each solution is re-evaluated whenever it is compared to others. This is unfortunate, both because it wastes computational resources and because it requires the user to foresee that noise is present (as in a noise-free setting, one would never re-evaluate solutions).

In this work, we show the need for re-evaluations could be overestimated, and in fact, detrimental. For the classic benchmark problem of how the $(1 + 1)$ evolutionary algorithm optimizes the LeadingOnes benchmark, we show that without re-evaluations up to constant noise rates can be tolerated, much more than the $O(n^{-2} \log n)$ noise rates that can be tolerated when re-evaluating solutions [1].

This first runtime analysis of an evolutionary algorithm solving a single-objective noisy problem without re-evaluations could indicate that such algorithms cope with noise much better than previously thought, and without the need to foresee the presence of noise.

References

[1] D. Sudholt, Analysing the Robustness of Evolutionary Algorithms to Noise: Refined Runtime Bounds and an Example Where Noise is Beneficial, *Algorithmica*, 83:976–1011, 2021.

Modeling Aggregate Electric Vehicle Flexibility in Unit Commitment Problems with Generalized Polymatroids

Hélène Arvis[1,2,3] Olivier Beaude[1] Nicolas Gast[2] Stéphane Gaubert[3]
Bruno Gaujal[2] Juliette Lesturgie[1]

[1]EDF Lab Paris-Saclay, OSIRIS department, France helene.arvis@edf.fr
[2]Inria Grenoble, France
[3]Inria Saclay, France

Keywords: Combinatorial Optimization, Electric Vehicle Aggregation, G-Polymatroids, Unit Commitment Problem

A common practice for representing flexible populations of assets in Unit Commitment (UC) problems, which aim at optimizing electricity production while ensuring energy supply and demand balancing [7], is the aggregation of similar units in order to decrease the computational difficulty of the resolution [5]. However, in many cases the naive aggregation, consisting in summing individual constraints, is not exact and can lead to infeasibilities on the individual level. Interestingly, in the specific case of Electric Vehicles (EVs), individual constraint sets can be seen as generalized polymatroids, or g-polymatroids, a special class of polytopes introduced by Frank and Tardos [3] and also studied by Danilov and Koshevoy in the general setting of discrete convexity [2]. In particular, g-polymatroids ensure exact aggregation when summed as their Minkowski sum remains a g-polymatroid with the correct characteristics, and they can allow efficient optimization of linear objectives by the greedy algorithm. This observation has been made independently by Mukhi, Loho, and Abate [6]. We show how to integrate EV constraint sets in the UC problem by leveraging the g-polymatroid structure through price decomposition methods [1] and submodular function minimization techniques [4]. We numerically illustrate this approach on a realistic UC model of the French electricity production.

References

[1] Pierre Carpentier and Guy Cohen. *Décomposition-coordination en optimisation déterministe et stochastique.* Vol. 81. Mathématiques et Applications. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017. DOI: 10.1007/978-3-662-55428-9.

[2] Vladimir I Danilov and Gleb A Koshevoy. "Discrete convexity and unimodularity—I". In: *Advances in Mathematics* 189.2 (2004), pp. 301–324. DOI: 10.1016/j.aim.2003.11.010.

[3] András Frank and Éva Tardos. "Generalized polymatroids and submodular flows". In: *Mathematical Programming* 42.1-3 (1988), pp. 489–563. DOI: 10.1007/BF01589418.

[4] Satoru Fujishige and Shigueo Isotani. "A Submodular Function Minimization Algorithm Based on the Minimum-Norm Base". In: *Pacific Journal of Optimization* 4 (2009).

[5] Jelle Meus, Kris Poncelet, and Erik Delarue. "Applicability of a Clustered Unit Commitment Model in Power System Modeling". In: *IEEE Transactions on Power Systems* 33.2 (2018), pp. 2195–2204. DOI: 10.1109/TPWRS.2017.2736441.

[6] Karan Mukhi, Georg Loho, and Alessandro Abate. *Exact Characterization of Aggregate Flexibility via Generalized Polymatroids.* 2025. DOI: 10.48550/arXiv.2503.23458.

[7] Milad Tahanan et al. "Large-scale Unit Commitment under uncertainty". In: *4OR* 13.2 (2015), pp. 115–171. DOI: 10.1007/s10288-014-0279-y.

Gradient Descent and Flows with General Costs

Pierre-Cyril Aubin-Frankowski[1]

[1] CERMICS, ENPC, IP Paris, France, pierre-cyril.aubin@enpc.fr

Keywords: gradient descent, convergence, alternating minimization

This talk is part of the invited session *Beyond Euclidean Convexity: Methods and Algorithms* (organizers: Pierre-Cyril Aubin and Michel De Lara)

How to go beyond the square distance d^2 in optimization algorithms and flows in metric spaces? Replacing it with a general cost function $c(x, y)$ and using a majorize-minimize framework I will detail a generic class of algorithms encompassing Newton/mirror/natural/Riemannian gradient descent/Sinkhorn/EM by reframing them as an alternating minimization, each for a different cost $c(x, y)$. Let $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}\cup\{+\infty\}$ where X is any set. Choose another set Y and a function $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$. Define the upperbound

$$f(x) + g(x) \leq \phi(x, y) := g(x) + c(x, y) + \sup_{x' \in X} [f(x') - c(x', y)]. \tag{1}$$

Setting $f^c(y) := \sup_{x' \in X}[f(x') - c(x', y)]$, we do alternating minimization of the surrogate

$$y_{n+1} \in \operatorname{argmin}_{y \in Y} c(x_n, y) + f^c(y) + g(x_n), \tag{2}$$

$$x_{n+1} \in \operatorname{argmin}_{x \in X} c(x, y_{n+1}) + f^c(y_{n+1}) + g(x). \tag{3}$$

Rooted in cross-differences, the convergence theory to the infimum and to the continuous flow is investigated is based on a (discrete) evolution variational inequality (EVI) which enjoys similar properties to the EVI with d^2 regularizer. In the limit, a curve $(x_t)_{t\in[0,+\infty)}$ solves a c-EVI if

$$\frac{d}{dt}c(x, x_t) + \mu \cdot c(x, x_t) \leq g(x) - g(x_t) \quad \text{a.e. } t \in (0, +\infty),\ x \in X.$$

This provides a theoretical framework for studying splitting schemes beyond the usual implicit Euler in gradient flows on metric spaces (like the JKO scheme). This talk is based on the works [1] with Flavien Léger (INRIA Paris), and [2] with Giacomo Sodini and Ulisse Stefanelli (Uni Vienna).

References

[1] Flavien Léger, Pierre-Cyril Aubin-Frankowski, Gradient descent with a general cost, 2023, arXiv:2305.04917.

[2] Pierre-Cyril Aubin-Frankowski, Giacomo Enrico Sodini, Ulisse Stefanelli, Evolution variational inequalities with general costs, 2025, arXiv:2505.00559.

Enhanced Benders' Decomposition for Stochastic Unit Commitment Using Interval Variables

Mathis Azéma[1] Vincent Leclère[1] Wim Van Ackooij[2]

[1]CERMICS, Ecole des Ponts, IP Paris, France, mathis.azema,vincent.leclere@enpc.fr
[2]EDF R&D, France, wim.van-ackooij@edf.fr

Keywords: Unit Commitment, Benders' Decomposition, Stochastic optimization

The Unit Commitment (UC) problem is a fundamental challenge for energy producers and network operators aiming to determine an optimal generation schedule for electricity production units, such as thermal and hydro units, while adhering to technical and economic constraints [1]. A major source of complexity in UC arises from uncertainty, particularly in forecasting energy demand and wind power generation. To address this, decomposition strategies such as Lagrangian relaxation and Benders' decomposition have been studied to efficiently handle many scenarios [2].

Recently, a new formulation of the deterministic UC problem based on interval variables has been proposed to better model unit production decisions [3]. Building on this approach, we introduce a novel Benders' decomposition that leverages interval variables to generate stronger cuts enabling convergence in a few number of iterations. Through computational experiments, we demonstrate that our method significantly improves solution efficiency, enabling the resolution of large-scale instances. In particular, we show that our approach scales effectively for two-stage stochastic UC with many scenarios and performs efficiently under both a risk-neutral framework and an average value-at-risk measure.

References

[1] Miguel F Anjos, Antonio J Conejo, et al. Unit commitment in electric energy systems. *Foundations and Trends in Electric Energy Systems*, 1(4):220–310, 2017.

[2] Wim van Ackooij, Irene Danti Lopez, Antonio Frangioni, Fabrizio Lacalandra, and Milad Tahanan. Large-scale unit commitment under uncertainty: an updated literature survey. *Annals of Operations Research*, 271(1):11–85, 2018.

[3] Ben Knueven, Jim Ostrowski, and Jianhui Wang. The ramping polytope and cut generation for the unit commitment problem. *INFORMS Journal on Computing*, 30(4):739–749, 2018.

A Learning-Based Approach for Traffic State Reconstruction from Limited Data

Nail Baloul[1] Amaury Hayat[1] Thibault Liard[2] Pierre Lissy[1]

[1] Ecole nationale des ponts et chaussées, France, nail.baloul@enpc.fr, amaury.hayat@enpc.fr, pierre.lissy@enpc.fr

[2] Université de Limoges, France, thibault.liard@unilim.fr

Keywords: Machine Learning, Neural Networks, Optimization, Differential Equations, Traffic Management

We propose an efficient method for reconstructing traffic density with low penetration rate of probe vehicles, relying solely on the initial and final positions of a small subset of cars. Due to the scarcity of collected data, we generate artificial trajectories using microscopic dynamical systems [2] and design a machine learning model to approximate the traffic density.

In [1], the authors have incorporated physics-informed neural networks (PINNs) to enforce conservation laws via PDE-derived Lagrangian terms, which necessitated real-time measurements. In contrast, our method requires less information, leveraging only the initial and final positions of probe vehicles, thus simplifying data collection.

Our method uses a residual network (ResNet) to analyze traffic dynamics. The learning process is set up as a constrained optimization problem. Starting from observed initial positions, the network predicts future states while incorporating physics-based principles of traffic flow.

Ultimately, inspired by [3], we prove that, when using only synthetic data from dynamical systems, our learned traffic density approximation converges to the LWR macroscopic model as the number of vehicles increases.

References

[1] Barreau, M., Aguiar, M., Liu, J., Johansson, K.H. "Physics-informed Learning for Identification and State Reconstruction of Traffic Density." September 2021.

[2] Colombo, R.M., Rossi, E. "On the micro-macro limit in traffic flow." June 2014.

[3] Di Francesco M., Rosini M.D. "Rigorous derivation of nonlinear scalar conservation laws from follow-the-leader type models via many particle limit." September 2015.

A Technician Routing and Scheduling Problem with a Lexicographic Objective

Elise Bangerter[1] David Schindl[2] Meritxell Pacheco[1] Nour Elhouda Tellache[1] Rodolphe Griset[3]

[1]University of Fribourg, Switzerland, elise.bangerter@unifr.ch, meritxell.pacheco@unifr.ch, nourelhouda.tellache@unifr.ch

[2]Geneva School of Business Administration, Switzerland, david.schindl@hesge.ch

[3]EDF R&D, OSIRIS Department, France, rodolphe.griset@edf.fr

Keywords: Vehicle Routing, Column Generation, Lexicographic Objective

The operation and maintenance of electricity distribution networks involve numerous technical interventions that must be carried out daily, while minimizing the operational costs associated with their execution. The problem we address originates from a technician routing and scheduling problem at Electricité de France (EDF). Specifically, we consider a set of teams of technicians and a set of interventions. Each intervention is associated with a specific time window and a set of required skills, while each team has an individual skill profile. The problem consists in assigning teams of technicians to interventions and dispatching them on routes that fulfill skill requirements, respect time windows, and satisfy additional operational constraints such as mandatory lunch breaks and workload limits. The problem involves two objectives to be optimized in lexicographic order: the primary objective is to maximize the total duration of interventions covered; the secondary objective is to minimize the total operational cost, comprising both personnel and travel expenses. We propose two mathematical formulations for the problem under study: a compact formulation and a set packing formulation. For the latter, we design column generation-based algorithms that incorporate different strategies to address the lexicographic objective function. These algorithms are applied to benchmark instances provided by EDF, and their performance is evaluated through a comparative analysis to assess both solution quality and computational efficiency.

Control-Based Design of Online Optimization Algorithms

Nicola Bastianello[1]

[1]KTH Royal Institute of Technology, Sweden, nicolba@kth.se

Keywords: Online optimization, time-varying optimization, control theory

In applications ranging from control, to signal processing, to machine learning, recent technological advances have made the class of *online optimization* problems of central importance. These problems are characterized by cost functions that vary over time, capturing the complexity of dynamic environments and changing optimization goals. Formally, an online problem is characterized by a discrete sequence of costs

$$\boldsymbol{x}_k^* = \underset{\boldsymbol{x} \in \mathbb{R}^n}{\arg\min} \, f_k(\boldsymbol{x}) \tag{1}$$

where a new cost is revealed every $T_\mathrm{s} > 0$ seconds, and where we assume that $\{f_k\}_{k\in\mathbb{N}}$ are strongly convex and smooth.

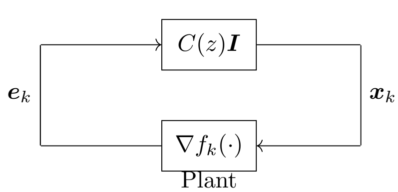
The goal then is to design online algorithms that track as closely as possible the optimal trajectory $\{\boldsymbol{x}_k^*\}_{k\in\mathbb{N}}$ *in real-time*. In particular, in [1, 2] we propose a novel approach to design such algorithms employing control theory.

The idea, depicted in Figure 1a, is to interpret the gradient of f_k as a plant to be controlled to zero. The resulting algorithm has been proved to converge exactly to the optimal trajectory [1], also for constrained problems [2]. Figure 1b reports the performance of the proposed algorithm (Control-based) against the state of the art online gradient, for different types of online problems.

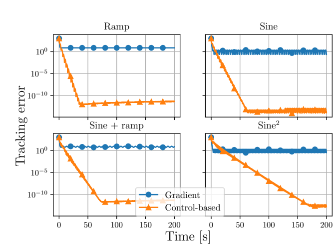
References

[1] Nicola Bastianello, Ruggero Carli, and Sandro Zampieri. Internal model-based online optimization. *IEEE Transactions on Automatic Control*, 69(1):689–696, January 2024.

[2] Umberto Casti, Nicola Bastianello, Ruggero Carli, and Sandro Zampieri. A control theoretical approach to online constrained optimization. *Automatica*, 176:112107, 2025.



(a) Online optimization as a control problem.



(b) Performance of the proposed algorithm.

A Knowledge Management Open Science Approach for Black-Box Optimization Benchmarking

Vitor Basto-Fernandes[1,2] Diederick Vermetten[2] Carola Doerr[2]

[1]Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Portugal, vitor.basto.fernandes@iscte-iul.pt
[2]Sorbonne Université, CNRS, LIP6, France, diederick.vermetten@lip6.fr, carola.doerr@lip6.fr

Keywords: Black-Box Optimization, Benchmarking, Open Science, Research Data Management

The black-box optimization research community has recently identified and made explicit in scientific *fora* (e.g. BeMCO 2024 [1], ROAR-NET COST Action [2]), the need for systematic, formal and open *data (knowledge) management* practices in this research field. Benchmarking data are generally not well structured, defined, or harmonized, for example, they suffer from the use of different taxonomies, nomenclatures, data representations, etc. This renders performance comparison and validation difficult or impossible, hindering the development and progress of this scientific area, as well as the corresponding potential for innovation and creation of economic value. Initiatives such as IOHprofiler [3], OPTION [4] and the Optimisation Problem Library [5], address some dimensions of this problem.

In a different scale and abstraction level, EU Data Strategy policies, initiatives, infrastructures, services and tools, such as the *Research and Innovation Common European Data Space* [6] and the *European Open Science Cloud - OpenAIRE*[7], are actively promoting research data management best practices.

We propose a roadmap for black-box optimization research data management to integrate with the EU Data Strategy agenda, benefiting from and contributing to the EU open science large-scale infrastructures, services, tools, standardization processes, and best practices. Our key goal is to make a step forward towards FAIR (findable, accessible, interoperable and reusable) research results in the black-box optimization research domain.

References

[1] Benchmarking in Multi-Criteria Optimisation, Lorentz Center - Leiden University, NL

[2] COST Action CA22137 Randomised Optimisation Algorithms Research Network

[3] IOHprofiler Iterative Optimization Heuristics Profiler

[4] OPTImization algorithm benchmarking ONtology (OPTION)

[5] OPL - Optimisation Problem Library

[6] Common European data spaces: Research and Innovation

[7] OpenAIRE Open Access Infrastructure for Research in Europe

A kite-based generator for airborne wind energy: modelling and optimisation

Antonin Bavoil[1]

[1]Université Côte d'Azur, CNRS, France, antonin.bavoil@univ-cotedazur.fr

Keywords: Kite, Modelling, ODE, Limit Cycle, Optimization)

Using kites to collect wind power and generate energy has been intensively studied in the last decade, see *e.g.* the survey by M. Diehl *et al* in [1]. In the framework of the KEEP (Kite Electrical Energy Power) funded by CNRS and gathering researchers from ENSTA Bretagne (well acquainted with the topic after previous studies on kites [2], most notably for boats [3]) and Université Côte d'Azur, we are interested in the analysis of a simple device composed of a kite attached to an arm; having the kite running along a well chosen curve will move the arm and generate electric power. We first build a simple point-mass mechanical model where the kite motion is prescribed to a conical surface modelled on an eight curve. The resulting differential equation can be expressed either as (i) a 5-dimensional second order DAE, or (ii) a dimension 2 second order ODE. For well chosen initial conditions, numerical integration of these two equivalent descriptions exhibit a limit cycle. We report on the optimization of the parameters of the device to maximize power on the limit cycle, as well as a validation of our approach using a third party trusted library [4].

References

[1] U. Ahrens, M. Diehl, R. Schmehl, Airborne Wind Energy, Springer, 2013.

[2] K. Desenclos, A. Nême, J.-B. Leroux, C. Jochum, A novel composite modelling approach for woven fabric structures applied to leading edge inflatable kites, Mechanics of Composite Materials, 2023.

[3] V. Podeur, D. Merdrignac, M. Behrel, K. Roncin, C. Fonti, C. Jochum, Y. Parlier, P. Renaud, Fuel economy assessment tool for auxiliary kite propulsion of merchant ship, Houille Blanche, 2018.

[4] J. De Schutter, R. Leuthold, T. Bronnenmeyer, E. Malz, S. Gros, M. Diehl, AWEbox: An Optimal Control Framework for Single- and Multi-Aircraft Airborne Wind Energy Systems, Energies, 2023.

Regularization of Optimal Control Problems on Stratified Domains using Additional Controls

Anas Bouali[1] Alain Rapaport[2] Térence Bayen[3]

[1]INRAE Montpellier, France, anas.bouali@inrae.fr
[2]INRAE Montpellier, France, alain.rapaport@inrae.fr,
[3]Avignon University, France, terence.bayen@univ-avignon.fr

Keywords: Optimal control, Fillipov solutions, Discontinuous dynamics, Regularization, Penalty method

This presentation focuses on hybrid optimal control. The goal is to implement a new regularization technique. This technique has the advantage of relying on a weaker transversality assumption than the commonly used transversality conditions. For this purpose, we consider a Mayer optimal control problem governed by a dynamics defined regionally, *i.e.*, the state space is stratified into a family of disjoint regions with nonsmooth interfaces, and, in each region, the dynamics is given by a smooth expression:

$$\dot{x} = f_j(x,u), \quad \text{if } \ \varphi_j(x) < 0.$$

It is shown that this problem is equivalent to a new optimal control problem, with additional controls v_j taking values in $[0,1]$ and a (smooth) dynamics as a convex combination of the smooth dynamics $\sum_{j=1}^{N} v_j f_j(x,u_j)$, along with the following mixed control-state constraint:

$$(1-2v_j)\varphi_j(x) = |\varphi_j(x)|.$$

Next, we introduce a family of auxiliary optimal control problems. In these problems, we first regularize the nonsmooth interfaces. In addition, we consider the convex combination of smooth dynamics (only) within a boundary layer. Furthermore, we add a penalization term to the cost function to account for the mixed control-state constraint. Our main result is that solutions to these (smooth) problems converge (up to a subsequence) to a solution of the original one. It is obtained thanks to a new hypothesis related to solutions to the auxiliary problems, which is weaker than the transverse crossing condition of the literature.

References

[1] A. Bouali, A. Rapaport and T. Bayen, Regularization of optimal control problems on stratified domains using additional controls, preprint, hal-04928858, 2025.

Speeding Up Hyper-Heuristics With Markov-Chain Operator Selection and the Only-Worsening Acceptance Operator*

Abderrahim Bendahi[1] Benjamin Doerr[2] Adrien Fradin[1] Johannes F. Lutzeyer[2]

[1]École Polytechnique, Institut Polytechnique de Paris, France

[2]Laboratoire d'Informatique (LIX), CNRS, École Polytechnique, Institut Polytechnique de Paris, France

{firstname.lastname}@polytechnique.edu

Keywords: Discrete optimization, evolutionary algorithms, runtime analysis

In discrete optimization, we seek to find "high-quality" solutions within a vast search space. Evolutionary algorithms form a family of randomized methods specifically designed to generate and refine promising candidates. Inspired by natural processes such as mutation, they explore the search space and employ heuristics to decide whether new candidates should be accepted or not. A notable approach [1], the move-acceptance hyper-heuristic (MAHH), alternates – via random coin flips – between an *elitist selection* (the only-improving operator, OI) and accepting any new candidate (the all-moves operator, AM), and was recently shown to escape local optima with remarkable efficiency.

In this talk [2], we propose two modifications to this algorithm that demonstrate impressive performances on a large class of benchmarks including the classic CLIFF_d and JUMP_m function classes and we provide an unified analysis of our method to this whole new class (contrary to [1]). **(i)** Instead of randomly choosing between the only-improving (OI) and any-move (AM) acceptance operator by flipping a coin, we take this choice via a simple two-state Markov chain. This modification alone reduces the runtime on JUMP_m functions with gap parameter m from $\Omega(n^{2m-1})$ to $O(n^{m+1})$. **(ii)** We then replace the all-moves (AM) acceptance operator with the operator that only accepts worsenings (OnlyWorsening, OW). Such a, counter-intuitive, operator has not been used before in the literature. However, our proofs show that our only-worsening operator can greatly help in leaving local optima, reducing, e.g., the runtime on Jump functions to $O(n^3 \log n)$ independent of the gap size.

In general, we prove a remarkably good runtime of $O(n^{k+1} \log n)$ for our Markov move-acceptance hyper-heuristic on all members of a new benchmark class SEQOPT_k, which contains a large number of functions having k successive local optima, and which contains the commonly studied JUMP_m and CLIFF_d functions for $k = 2$.

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

References

[1] Andrei Lissovoi, Pietro S. Oliveto, John Alasdair Warwicker. When move acceptance selection hyper-heuristics outperform Metropolis and elitist evolutionary algorithms and when not. *Artificial Intelligence*, 314:103804, 2023.

[2] Abderrahim Bendahi, Benjamin Doerr, Adrien Fradin, Johannes F. Lutzeyer. Speeding up hyper-heuristics with Markov-chain operator selection and the only-worsening acceptance operator. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2025* (to appear). arXiv:2506.01107.

* This work won a *Distinguished Paper Award* at IJCAI 2025.

Rational SOS certificates of any polynomial with integer coefficients

Matías Bender[1] Khazhgali Kozhasov[2] Elias Tsigaridas[3] Chaoping Zhu[4]

[1]Inria Saclay and École polytechnique, France, matias.bender@inria.fr
[2] Université Côte d'Azur, France, khazhgali.kozhasov@univ-cotedazur.fr
[3]Inria Paris and Sorbonne University, France, elias.tsigaridas@inria.fr
[4]Sorbonne University, France, zhu@imj-prg.fr

Keywords: Nonnegative polynomials, stereographic transformation, coercive functions, gradient perturbation, sum-of-squares decompositions, zero-dimensional ideal, rational univariate representation

Given a multivariate polynomial with integer coefficients, we present rational certificates of nonnegativity based on sum-of-squares decompositions, without making any the assumptions on the input polynomial. We introduce the stereographic transformation, which allows us to transform any polynomial to a coercive polynomial. We also employ perturbation methods to ensure that a coercive polynomial has a zero-dimensional gradient ideal. All these transformations preserve the nonnegativity of the input polynomial and reduce the problem to our previous work [1]. The complexity of our algorithm is single exponential in the number of variables. This is a joint work with Matías Bender, Khazhgali Kozhasov and Elias Tsigaridas.

References

[1] Bender, Matías R and Tsigaridas, Elias and Zhu, Chaoping, Rational SOS certificates of any polynomial over its zero-dimensional gradient ideal, hal-04909464

Average case analysis of the performance of Branch and Bound algorithms: Is there potential for quantum advantage?

Alexandre Bergerault[1] Hugo Izadi[1,2] Vassilis Apostolou[1]

[1]Quandela SAS, 7 Rue Léonard de Vinci, 91300 Massy, France, vassilis.apostolou@quandela.com

[2]Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

Keywords: Branch and Bound, QUBO, Semidefinite programming, Quantum Optimisation

Title of the session: Quantum computing and combinatorial optimization

Combinatorial optimization problems are ubiquitous across machine learning, finance, and operational research. Although worst-case analysis classifies these NP-hard problems as exponentially complex, Branch-and-Bound (BnB) algorithms, particularly those leveraging relaxation heuristics, often achieve practical efficiency. This work addresses the critical gap in understanding the average-case complexity of BnB in this setting. We develop an analytical framework based on the statistical properties of problem costs and their relaxations. Focusing on Quadratic Unconstrained Binary Optimization (QUBO) as a prototypical NP-hard problem for our analysis, we formally prove that its subproblem's costs are asymptotically normally distributed under specific spectral conditions on the objective function matrix.

Leveraging this, our analysis demonstrates a critical transition in the expected size of the BnB tree: it transitions from exponential growth to a more constrained, near-polynomial growth regime. This transition is accurately predicted by a critical pruning depth, $k^\star = n(1 - \alpha_{wc}\alpha_{ac})$, where α_{wc} is the worst case integrality gap (for global bounds) and α_{ac} is the average integrality gap of the SDP relaxation (for subproblems). Numerical experiments validate the normality premise across various QUBO instance types, including those derived from real-world NP-complete problems like Max-Cut, TSP, and Portfolio Optimization. Our experimental analysis on an SDP-based BnB solver demonstrates that empirical performance significantly outperforms theoretical worst-case predictions, with the predicted peak complexity depth aligning with our model's predictions. Furthermore, by comparatively analyzing Knapsack, Vertex Cover, and Binary TSP, we show that our model accurately predicts the vast differences in their practical tractability, which are not captured by worst-case NP-completeness equivalence alone.

These findings offer a novel, quantitative lens for understanding the performance of Branch-and-Bound algorithm, providing concrete guidance for developing improved classical, hybrid, and quantum optimization methods. This includes formalizing the strategic advantage of root node integrality gap strengthening and average integrality gap, and establishing new dimensions for robust benchmarking of optimization algorithms. Our work thus advances the understanding of what makes hard problems tractable.

Implementation of a Multi-year Generation and Transmission Expansion Planning Problem in an Industrial Context

Nicolas Bessin[1] Thomas Bittar[2] Antoine Oustry[2] Manuel Ruiz[2]

[1]École Nationale des Ponts et Chaussées, France, nicolas.bessin@eleves.enpc.fr
[2]RTE, France, thomas.bittar@rte-france.com, antoine.oustry@rte-france.com, manuel.ruiz@rte-france.com

Keywords: Generation and Transmission Expansion Planning, Stochastic Optimization, Scenario Tree, Benders Decomposition

As part of its public service mission, RTE (Réseau de Transport d'Électricité), the French transmission system operator, carries out prospective studies to assess the security of supply of electricity and to enlighten national strategic orientations for the evolution of the power system.

Because prospective studies aim at defining an economically efficient and low-carbon generation mix together with appropriate network development, they involve solving generation and transmission expansion planning problems. For this task, RTE relies on an internally developed open-source software, Antares-Xpansion. Currently, Antares-Xpansion addresses a two-stage optimization problem over a one-year horizon, where an investment decision (first stage) is made at the outset, followed by system operation (second stage) on an hourly time step, across several scenarios that capture operational uncertainties (demand, hydro inflows, generator availability).

In practice, investment decisions are path-dependent: building the power system of 2050 requires decisions to be made and implemented already in 2030, 2040, and so forth, each of which impacts system operations at those horizons. Such trajectory constraints stem primarily from financial limitations and from industrial realities - Rome was not built in a day! Moreover, these decisions are made amid structural uncertainties concerning key parameters (electricity consumption, location of new renewable power plants), for which only ranges can be anticipated.

Currently, Antares-Xpansion can only address single-year problems. Thus, investment trajectories must be constructed manually by solving the model sequentially for different individual years. This procedure leads to a myopic representation, unable to anticipate investment needs.

To improve the process, we implement a multi-year investment model in Antares-Xpansion, where investment decisions are taken on a scenario tree. The focus is to show how an efficient implementation has been carried on, taking into account constraints from the industrial context:

- How to efficiently reuse the Antares-Xpansion code that already handles the investment problem with a single first stage ?
- How was user interaction and data gathering extended from a single-horizon study to a multi-horizon study?

We also demonstrate how this implementation was validated using a realistic case study involving the resolution of an optimization problem involving hundreds of millions of variables.

Decentralised energy management of large-scale EV fleets: some results from the ANR EDEN4SG project

Anne Blavette[1] Eloann Le Guern–Dall'o[1] Eymeric Giabicani[1]
Xavier Ratte[2] Raphaël Féraud[2] Guy Camilleri[3] Patrick Maillé[4]
Hamid Ben Ahmed[1] Fatma Ezzahra Salem[2] Riadh Zorgati[5] Andréa Laugère [5]
Arthur Mepuis[6] Jesse-James Prince Abgodjan[6]

[1]ENS Rennes, Univ. Rennes, CNRS, IETR lab, Rennes, France, anne.blavette@ens-rennes.fr
[2]Orange Lannion & Paris, France, raphael.feraud@orange.com
[3]IRIT, University Paul Sabatier, Toulouse, France, guy.camilleri@irit.fr
[4]IRISA, IMT Atlantique, Rennes, France patrick.maille@imt-atlantique.fr
[5]EDF Lab, Palaiseau, France, riadh.zorgati@edf.fr
[6]SRD, Poitiers, France, arthur.mepuis@srd-energies.fr

Keywords: large-scale smart grids, reinforcement learning, bandits, grid constraints

The optimal strategy for integrating large-scale EV fleets with regard to the DSO constraints in a manner that maximizes the local use of renewable energy sources is an NP-hard problem. Moreover, due to stochastic weather and human behaviors, this optimization should be done under uncertainty and potentially online. Reinforcement learning is a tool of choice under these conditions. As part of this work, bandits algorithms have been investigated, in particular stochastic bandits based on Thompson sampling [1]. A comparative analysis was also performed with adversarial bandits which were expected to provide theoretical bounds, at the cost of performance degradation that we aimed to quantify [2]. In particular, the "Follow-The-Perturbed-Leader" (FTPL) combined with heavy-tailed distributions (Fréchet, Pareto) was considered. A last part of the work focussed on a comparative analysis of bandit algorithms with Markov-decision process-based reinforcement learning in the form of the proximal policy optimization (PPO) algorithm. This work compared the performance of both algorithm types used independently and in combination [3].

References

[1] E. Le Guern-Dall'o, R. Féraud, G. Camilleri, P. Maillé, F. Petit, R. Zorgati, H. Ben Ahmed, A. Blavette, "Muti-agent multi-armed bandits: application to EVs smart charging with grid constraints", *IEEE PowerTech conference, Kiel, Germany*, 2025.

[2] E. Giabicani, E. Le Guern-Dall'o, R. Féraud, G. Camilleri, P. Maillé, H. Ben Ahmed, A. Mepuis, J.-J. Prince Agbodjian, R. Zorgati, A. Blavette, "Decentralized and Robust EV Fleet Charge Scheduling Using Follow-The-Perturbed-Leader Under Multiple Uncertainties", *submitted to the PSCC 2026 conference, Limassol, Cyprus.*

[3] X. Ratte, Master thesis, *in preparation*, CentraleSupélec, 2025.

On the problem of minimizing the epidemic final size for the SIR model by social distancing

Anas Bouali

MISTEA, Univ. Montpellier, INRAE, Institut Agro, Montpellier, France
Contact: anas.bouali@inrae.fr

In this presentation, we revisit the problem of minimizing the epidemic final size in the SIR model through social distancing of bounded intensity:

$$\sup_{u(\cdot)\in\mathcal{U}} S(+\infty) \quad \text{subject to} \quad \|u(\cdot)\|_1 \leq K.$$

In the existing literature, this problem has been considered imposing a priori interval structure on the time period when interventions are enforced. We show that when considering the more general class of controls with an L^1 constraint on the confinement effort that reduces the infection rate, the support of the optimal control is still a single time interval. This shows that, for this problem, there is no benefit in splitting interventions on several disjoint time periods.

However, if the infection rate β is known in advance to vary over time as follows:

$$\beta(t) = \begin{cases} \beta_1, & \text{if } t \in [0, \tau), \\ \beta_2, & \text{if } t \in [\tau, \infty), \end{cases}$$

then we show that the optimal solution could consist of splitting the interventions in at most two disjoint time periods. Finally, we provide a couple of numerical examples[1] illustrating the structure of the optimal control derived in both cases: constant and piecewise constant transmission rate β.

References

1. P.-A. Bliman, A. Bouali, P. Loisel, A. Rapaport, A. Virelizier (2025). *Regularization of optimal control problems on stratified domains using additional controls*, submitted https://hal.science/hal-05194927v1

2. R. Balderrama, J. Peressutti, J.P. Pinasco, F. Vazquez, C.S. de la Vega (2022). *Optimal control for a SIR epidemic model with limited quarantine. Scientific Reports*, 12, 12583.

3. P.-A. Bliman, A. Rapaport (2023). *On the problem of minimizing the epidemic final size for SIR model by social distancing. IFAC-PapersOnLine*, 56, 4049–4054. https://doi.org/10.1016/j.ifacol.2023.10.1726

[1] The corresponding code and simulations are available at https://anasxbouali.github.io/SIRcontrol.jl.

A Characterization of Law-Invariant Coherent Risk Measures Through Optimal Transport

Riccardo Bonalli[1] Benoît Bonnet-Weill[1] Laurent Pfeiffer[1,2]

[1] Laboratoire des Signaux et des Systèmes, CentraleSupélec, CNRS, Université Paris-Saclay, France,
[2] Centre Inria de Saclay, France

riccardo.bonalli@centralesupelec.fr benoit.bonnet-weill@centralesupelec.fr
laurent.pfeiffer@inria.fr

Keywords: Coherent risk measures, Optimal transportation, duality

This talk deals with risk measures, a popular tool for modeling the aversion that one may have for the variability of a random cost. An example is the Conditional Value at Risk (CVaR), defined as the conditional expectation of the cost of interest in the worst $q\%$ cases, for a given level q. The popularity of CVaR lies in the existence of a convenient dual representation formula, taking the form of a simple minimization problem [1].

We will present a novel characterization of the risk measures satisfying two fundamental properties: law-invariance and coherence [2]. We will show that such risk measures can be represented as the value function of a generalized optimal transport problem. By generalized, we mean that the second marginal of the transport plan to be optimized is not fixed, but only restricted to lie in a given set of measures R. Our characterization is intimately related to Kusuoka's theorem [3], which states that any law-invariant and coherent risk measure can be represented as a supremum of a family of convex combination of CVaRs with different levels q.

Next we will present a duality theorem for the case where the set R is convex, extending the classical Kantorovich duality theorem in optimal transportation. This will allow us to derive a duality formula for risk measures, generalizing the one available for CVaR. The practical interest of such a formula will be discussed, in particular for the application of numerical methods based on epi-regularization [4].

References

[1] R.T. Rockafellar and S. Uryasev, Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.

[2] A. Shapiro, D. Dentcheva, and A. Ruszczyski, Lectures on stochastic programming, *SIAM Series on Optimization*, 2009.

[3] S. Kusuoka. On law invariant coherent risk measures, In *Advances in mathematical economics*, pages 83–95, Springer, 2001.

[4] D.P. Kouri and T.M. Surowiec. Epi-regularization of risk measures, *Mathematics of Operations Research*,45:774–795, 2020.

Mirror and Preconditioned Gradient Descent in Wasserstein Space

Clément Bonet[1] Théo Uscidda[2] Adam David[3]
Pierre-Cyril Aubin-Frankowski[4] Anna Korba[2]

[1]Ecole Polytechnique, CMAP, France, clement.bonet.mapp@polytechnique.edu
[2]ENSAE, CREST, France, theo.uscidda@ensae.fr, anna.korba@ensae.fr
[3] Institute of Mathematics, Technische Universität Berlin, Allemagne, david@math.tu-berlin.de
[4]ENPC, Cermics, France, pierre-cyril.aubin@enpc.fr

Keywords: mirror descent, preconditioned gradient descent, Wasserstein gradient

This talk is part of the invited session *Beyond Euclidean Convexity: Methods and Algorithms* (organizers: Pierre-Cyril Aubin and Michel De Lara)

As the problem of minimizing functionals on the Wasserstein space encompasses many applications in machine learning, different optimization algorithms on $\mathbb{R}^d$ have received their counterpart analog on the Wasserstein space. We focus here on lifting two explicit algorithms: mirror descent and preconditioned gradient descent. These algorithms have been introduced to better capture the geometry of the function to minimize and are provably convergent under appropriate (namely relative) smoothness and convexity conditions. Adapting these notions to the Wasserstein space, we prove guarantees of convergence of some Wasserstein-gradient-based discrete-time schemes for new pairings of objective functionals and regularizers. The difficulty here is to carefully select along which curves the functionals should be smooth and convex. We illustrate the advantages of adapting the geometry induced by the regularizer on ill-conditioned optimization tasks, and showcase the improvement of choosing different discrepancies and geometries in a computational biology task of aligning single-cells.

References

[1] C. Bonet, T. Uscidda, A. David, P.C. Aubin-Frankowski, A. Korba, Mirror and Preconditioned Gradient Descent in Wasserstein Space. *Advances in Neural Information Processing Systems*, 2024.

Flowing Datasets with Wasserstein over Wasserstein Gradient Flows

Clément Bonet[1] Christophe Vauthier[2] Anna Korba[2]

[1]ENSAE, CREST, IP Paris clement.bonet@ensae.fr, anna.korba@ensae.fr
[2]Université Paris-Saclay, Laboratoire de Mathématiques d'Orsay, christophe.vauthier@universite-paris-saclay.fr

Keywords: Wasserstein gradient flows, Optimal transport, Datasets

Many applications in machine learning involve data represented as probability distributions. The emergence of such data requires radically novel techniques to design tractable gradient flows on probability distributions over this type of (infinite dimensional) objects. For instance, being able to flow labeled datasets is a core task for applications ranging from domain adaptation to transfer learning or dataset distillation. In this setting, we propose to represent each class by the associated conditional distribution of features, and to model the dataset as a mixture distribution supported on these classes (which are themselves probability distributions), meaning that labeled datasets can be seen as probability distributions over probability distributions. We endow this space with a metric structure from optimal transport, namely the Wasserstein over Wasserstein (WoW) distance, derive a differential structure on this space, and define WoW gradient flows. The latter enables to design dynamics over this space that decrease a given objective functional. We apply our framework to transfer learning and dataset distillation tasks, leveraging our gradient flow construction as well as novel tractable functionals that take the form of Maximum Mean Discrepancies with Sliced-Wasserstein based kernels between probability distributions. This work was originally published at the International Conference on Machine Learning in 2025, where it was accepted for an oral presentation [1].

References

[1] C. Bonet, C. Vauthier, and A. Korba. Flowing Datasets with Wasserstein over Wasserstein Gradient Flows. In International Conference on Machine Learning. PMLR, 2025.

Optimisation des traitements microbiens : géométrie des trajectoires dans les systèmes Lotka–Volterra

Bernard Bonnard[1] Jérémy Rouot[2]

[1]Institut de Mathématiques de Bourgogne, & INRIA, McTAO Team, Sophia Antipolis.
bernard.bonnard@u-bourgogne.fr
[2]Laboratoire de Mathématiques de Bretagne Atlantique, Brest,
rouot@math.cnrs.fr

Keywords: Geometric optimal control, Biomathematics, Population dynamics, Gause–Lotka–Volterra–Kolmogorov equation

Dans cette présentation, nous étudions le contrôle optimal d'un système dynamique à trois espèces, modélisé par les équations de Lotka–Volterra, dans le but de réduire l'infection d'un microbiote complexe. Ce type de modèle s'inspire de travaux récents sur la dynamique intestinale, notamment Stein et al. (2013), qui ont utilisé des modèles écologiques pour analyser la stabilité et les interactions du microbiote intestinal à partir de séries temporelles. Le contrôle étudié correspond à un régime de dosage antibiotique ou probiotique, et notre approche s'appuie sur une analyse géométrique des trajectoires optimales. En appliquant le principe du maximum de Pontryagin et les méthodes de classification des contrôles développées dans Bonnard & Rouot (2024), nous paramétrons les trajectoires sous forme de géodésiques et identifions les ensembles limites α et ω, constitués de points d'équilibre isolés correspondant à des rayons anormaux. L'étude montre que ces géodésiques anormales forment un feuilletage d'une surface quadratique et permet de classifier la dynamique complète des trajectoires à l'aide d'un système d'invariants calculé de manière explicite et symbolique. Cette méthodologie offre un outil puissant pour analyser et optimiser les régimes de traitement microbiens afin de réduire les infections, tout en donnant une vision géométrique claire de la structure des trajectoires optimales dans ces systèmes dynamiques.

References

[1] B. Bonnard, J. Rouot, *Feedback Classification and Optimal Control with Applications to the Controlled Lotka-Volterra Model,* Optimization **74** no. 5 (2024), pp. 1173–1196.

[2] R.R. Stein, V. Bucci, N.C. Toussaint, C.G. Buffie, G. Rätsch, E.G. Pamer, et al., *Ecological modelling from time-series inference: insight into dynamics and stability of intestinal microbiota,* PLos Comp. Biology, **9** no. 12 (2013).

A Universal Uniform Approximation Theorem for Neural Networks

Olivier Bournez[1] Johanne Cohen[2] Adrian Wurm[3]

[1]Institut Polytechnique de Paris, Ecole Polytechnique, LIX, 91128 Palaiseau Cedex, France
bournez@lix.polytechnique.fr
[2]Equipe GALaC, LISN, CNRS, Université Paris-Saclay, 91190 Gif-sur-Yvette, France
johanne.cohen@lisn.fr
[3]Computer Science Institute, BTU Cottbus-Senftenberg D-03046, Cottbus, Germany
wurm@b-tu.de

Keywords: Models of computation, Complexity theory, Formal neural networks, Uniform approximation, Optimisation

One of the best-known fundamental results in neural network theory is that they are universal approximators: for any continuous function, any compact domain, and any precision ϵ, there exists a neural network within distance at most ϵ from the function over this compact domain. Mathematically, uniform approximation states that the class of functions realized by neural networks is dense in the set of continuous functions over compact domains. This holds under minimal assumptions (a single hidden layer and a non-polynomial activation function suffice.

We show the existence of a **fixed** recurrent network capable of approximating any computable function with arbitrary precision, provided that an encoding of the function is given in the initial input. While uniform approximation over a compact domain is a well-known property of neural networks, we go further by proving that our network ensures effective uniform approximation–simultaneously ensuring:

- **Uniform approximation in the sup-norm sense**, guaranteeing precision across the compact domain $[0,1]^d$;
- **Uniformity in the sense of computability theory** (also referred to as *effectivity* or *universality*), meaning the same network works for all computable functions.

Our result is obtained constructively, using original arguments. Moreover, our construction bridges computation theory with neural network approximation, providing new insights into the fundamental connections between circuit complexity and function representation.

We also show that the associated decision problems are complete– even under strict architectural constraints such as fixed, small network width and depth.

Furthermore, this connection extends beyond computability to complexity theory. The obtained network is efficient: if a function is computable or approximable in polynomial time in the Turing machine model, then the network requires only a polynomial number of recurrences or iterations to achieve the same level of approximation, and conversely. Moreover, the recurrent network can be assumed to be very narrow, strengthening the link our results and existing models of very deep learning, where uniform approximation properties have already been established.

A Theoretical Framework for Grokking: Interpolation Followed by Riemannian Norm Minimisation

Etienne Boursier[1] Scott Pesme[2] Radu-Alexandru Dragomir[3]

[1]INRIA, LMO, Université Paris-Saclay, Orsay, France, etienne.boursier@inria.fr
[2]INRIA, Grenoble, France, scott.pesme@inria.fr
[3]Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France, dragomir@telecom-paris.fr

Keywords: grokking, optimisation, theory of neural networks

Coined by Power et al. [2022], *grokking* refers to a two-phase training pattern: the training loss quickly drops to zero while the test loss plateaus, followed by a phase where the training loss stays at zero but the test loss steadily improves. We propose a new theoretical perspective on grokking via gradient flow dynamics with **weight decay**. In the limit of vanishing weight decay, we fully characterise the parameter trajectory and show it decomposes into two phases. First, the flow follows the unregularised path and converges to a manifold of critical points of the training loss. Then, in a slow drift phase, weight decay moves the parameters along this manifold, gradually shrinking their ℓ_2 norm. We argue this reduction explains grokking, since smaller weight norms often correlate with improved generalisation.

Informal statement of the result. For a generic training loss $F : \mathbb{R}^d \to \mathbb{R}$, consider the gradient flow with weight decay: $\dot{w}^\lambda(t) = -\nabla F(w^\lambda(t)) - \lambda w^\lambda(t)$. Assuming the unregularised flow is bounded, we prove:

Theorem 1 (Main result, informal)**.** *As $\lambda \to 0$, $w^\lambda(t)$ combines two dynamics:*

1. ***(Fast dynamics driven by F)*** *In a first phase, the weights follow the unregularised gradient flow and converge to a manifold of critical points of F.*
2. ***(Slow dynamics driven by the weight decay)*** *At time $t \approx 1/\lambda$, the iterates start slowly drifting along this manifold, following a Riemannian gradient flow that decreases the ℓ_2-norm of the weights.*

Connection to grokking. This optimisation result naturally explains grokking: in practice, deep models typically reach global minimisers that generalise poorly when initialised with large weights. The subsequent slow drift, driven by weight decay, reduces the ℓ_2-norm and simplifies the solution. Since smaller norms often align with better generalisation, this explains the delayed test improvement observed in practice.

References

Alethea Power, Yuri Burda, Harri Edwards, Igor Babuschkin, and Vedant Misra. Grokking: Generalization beyond overfitting on small algorithmic datasets. *arXiv preprint*, 2022.

Guaranteed performance in short-time quantum annealing

Arthur Braida[1] Simon Martiel[2] Ioan Todinca[3]

[1] Universite Paris Cite, IRIF, France, arthur.braida@irif.fr
[2] IBM Quantum
[3] Universite d'Orleans

Keywords: Approximation ratio, Maximum-Cut, Lieb-Robinson bound, locality

Quantum annealing (QA) is an analogue paradigm for quantum optimization that remains a leading candidate for demonstrating quantum advantage on combinatorial problems. Unlike gate-based methods such as the Quantum Approximate Optimization Algorithm (QAOA), QA evolves a system continuously under a time-dependent Hamiltonian, a feature that makes rigorous analysis challenging and has limited provable performance guarantees to date. Numerical comparisons between QA and QAOA are informative but constrained by the classical cost of simulating Schrodinger dynamics and by current annealer sizes, so analytical tools are needed to assess worst-case behavior.

Here we introduce a parametrized (tunable) variant of QA and use it to perform a precise 1-local analysis on regular graphs. The analysis rests on a tight Lieb-Robinson bound derivation tailored to QA dynamics by exploiting the commutativity graph construction, together with a simple prefactor on the initial (driver) Hamiltonian that improves analytical tightness without altering the algorithmic procedure. Using MaxCut on cubic graphs as a benchmark, we show that a linear-schedule, constant-time QA, when analyzed as a 1-local algorithm, achieves a worst-case approximation ratio exceeding 0.7020. This guarantee surpasses one-layer QAOA (=0.6925) and the best known classical 1-local algorithm (=0.6980).

Our results demonstrate that refined locality bounds and modest, analytically motivated Hamiltonian adjustments can materially tighten provable performance guarantees for QA, narrowing the gap between numerical evidence and rigorous theory.

Optimizing a Worldwide-Scale Shipper Transportation Planning in a Carmaker Inbound Supply Chain

Mathis Brichet[1,2] Maximilian Schiffer[3,4] Axel Parmentier[1]

[1] CERMICS, Ecole des Ponts, France, mathis.brichet@enpc.fr, axel.parmentier@enpc.fr
[2] Intelligence Artificielle Supply Chain, Renault, France, mathis.brichet@renault.com
[3] School of Management, Technical University of Munich, Germany, schiffer@tume.de
[3] Munich Data Science Institute, Technical University of Munich, Germany

Keywords: Shipper Network Design, Iterated Local Search, Flow Consolidation in Supply Chain

We study the shipper-side design of large-scale inbound transportation networks, motivated by Renault's global supply chain. We introduce the *Shipper Transportation Design Problem*, which integrates consolidation, routing, and regularity constraints. The resulting problem is a rich version of a multicommodity network design problem [1]. We propose a tailored Iterated Local Search (ILS) metaheuristic to solve it. The algorithm combines large-neighborhood search with MILP-based perturbations and exploits bundle-specific decompositions and giant container bounds to obtain scalable lower bounds and effective benchmarks. Computational experiments on real industrial data show that the ILS achieves an average gap of 7.9% to the best available lower bound on world-scale instances with more than 700,000 commodities and 1,200,000 arcs, delivering solutions showing a potential of 23.2% cost reduction compared to the Renault-based benchmark. To our knowledge, this is the first approach to solve shipper-side transportation design problems at such scale, being orders of magnitude bigger than what usual resolution methods from the literature can solve [2, 3, 4]. Our analysis further yields managerial insights: accurate bin-packing models are essential for realistic consolidation, highly regular plans offer the best balance between cost and operational stability, and outsourcing is only attractive in low-volume contexts, while large-scale networks benefit from in-house planning. The final section explores the network design perspectives emerging from our findings.

References

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

[2] I. Bakir, A. Erera, M. Savelsbergh, Motor carrier service network design, *Springer*, 2021.

[3] T.G. Crainic, M. Gendreau, Heuristics and Metaheuristics for Fixed-Charge Network Design, *Network Design with Applications to Transportation and Logistics*, 91–138, 2021.

[4] K. Lindsey, A. Erera, M. Savelsbergh, Improved integer programming-based neighborhood search for less-than-truckload load plan design, *Transportation science*, 50(4):1360–1379, 2016.

Commodity Maritime Flow Determination

Busch Nicolas[1] Balbiyad Saad[1] Naim Samir[1] Renart Paul[1]

[1]KPLER, France, nbusch@kpler.com, sbalbiyad@kpler.com, snaim@kpler.com, prenart@kpler.com

Keywords: Mixed Integer Linear Programming [1], Maritime Flow, Combinatorial Optimization

At KPLER, we develop advanced methodologies to estimate vessel-level trade flows — including product types and quantities — at global ports. By integrating fragmented data sources (such as draft surveys, public and private reports, analyst inputs, and port infrastructure constraints), we aim to reconstruct granular trade activity with precision, enabling deeper insights into maritime logistics and commodity markets. To achieve this, KPLER employs a vessel-by-vessel approach that provides a real-time, macro-level view of commodity volumes in transit, offering unparalleled insights into maritime logistics and global commodity markets.

To estimate quantities, we rely on draft change recordings, which are not always reliable and may contain data gaps. We complement this with all available product- and port-level information to enhance flow consistency. However, data sources are not always coherent, which necessitates sophisticated conflict resolution between constraints. Furthermore, vessels can exchange commodities at sea through ship-to-ship operations, leading to complex interactions between vessels and trade formations.

Currently, this is handled by separating tasks: first estimating quantities, then products, and finally building trades. However, this approach poorly handles product-quantity constraints that reflect real vessel behavior. For instance, some ports have specific export/import capabilities dependent on product type, which leads to inconsistencies between estimated quantities and the chosen product at the port level.

In this work, we designed a Mixed-Integer Linear Programming (MILP) model that incorporates all three tasks (quantity, product, and trade) to address the interdependence between quantity, product, and trade constraints. Additionally, the model must be computationally efficient, as it will be executed hundreds of thousands of times daily. This is particularly challenging due to the combinatorial nature of the problem.

References

[1] George Nemhauser, Laurence Wolsey, Integer and Combinatorial Optimization, 1988.

Operator Splitting for Large Scale Nonlinear Stochastic Optimal Control

François Caffier[1,2] Paul Malisani [1] Laurent Pfeiffer [2]

[1]IFP Energies nouvelles, francois.caffier@ifpen.fr, paul.malisani@ifpen.fr
[2]Inria Saclay, Laboratoire des Signaux et Sytèmes, Centrale Supélec, laurent.pfeiffer@inria.fr

Keywords: Optimal Control, Douglas-Rachford Splitting, Stochastic Optimization, Large Scale Optimization, Nonconvex Optimization, Aggregative Optimization.

This presentation addresses the numerical resolution of an aggregative stochastic optimal control problem in which agents are subject to a common source of uncertainty. The agents' dynamics may be nonlinear, rendering the problem typically nonconvex. Our problem is motivated by the practical framework for optimal and coordinated management of a large number of nonlinear energy dynamical systems under uncertainty.

A direct application of Rockafellar's expansion lemma [1, Lemma 1] allows a reformulation of our problem with separability across the agents and scenarios. This structure is particularly well-suited for a numerical resolution using Douglas-Rachford Splitting (DRS) algorithm, a well-known method for minimizing the sum of two convex functions, with proven weak convergence to a stationary point. In the nonconvex setting, when one function is smooth with a Lipschitz-continuous gradient and the other is proper and lower semicontinuous, convergence in finite dimensions to a point satisfying the first-order necessary optimality condition has been established in [2]. We adapt the proof of this result to our infinite-dimensional setting, where the second function is assumed to be convex. Under suitable conditions on the dynamics and running cost, we establish the existence of a cluster point for the sequence generated by the DRS scheme when applied to our problem.

References

[1] R. T. Rockafellar, Progressive decoupling of linkages in optimization and variational inequalities with elicitable convexity or monotonicity, *Set-Valued and Variational Analysis*, 27(4):863–893, 2019.

[2] A. Themelis and P. Patrinos, Douglas–Rachford splitting and ADMM for nonconvex optimization: Tight convergence results, *SIAM Journal on Optimization*, 30(1):149–181, 2020.

Particle Method for a Nonlinear Multimarginal Optimal Transport Problem

Adrien Cances[1] Quentin Mérigot[1] Luca Nenna[1]

[1] Université Paris-Saclay, Orsay, France
adrien.cances@universite-paris-saclay.fr,
quentin.merigot@universite-paris-saclay.fr,
luca.nenna@universite-paris-saclay.fr

Keywords: risk estimation, spectral risk measure, multimarginal transport, Lagrangian discretization

We investigate a Lagrangian particle discretization for a risk estimation problem within the framework of optimal transport. Given finitely many univariate probability densities corresponding to risk factors, as well as a certain danger function, one must find the riskiest joint law, in the sense of a certain spectral risk measure. The problem in question has been shown by Ennaji, Mérigot, Nenna, and Pass in [1] to be equivalent to standard multimarginal optimal transport problem with one additional marginal. To discretize the latter, we look for an approximate solution as a uniform point cloud of size N, and optimize the positions of the Dirac masses. The constraints on the marginals are dealt with via quadratic Wasserstein penalization terms, each of which is a semi-discrete optimal transport cost in dimension one. We prove that the discretized problem converges to the original one as the number of points goes to infinity, with a rate determined by the asymptotic uniform quantization errors of the marginals. Numerical experiments illustrate and support our theoretical findings.

References

[1] H. Ennaji, Q. Mérigot, L. Nenna, B. Pass, Robust risk management via multi-marginal optimal transport, *Journal of Optimization Theory and Applications*, 202(2), 554-581, 2024.

Moment Constrained Optimal Transport for Energy Demand Management

Julien Cardinal[1,2] Thomas Le Corre[1,2] Ana Bušić[1,2]

[1]Inria, Paris, France
[2]DI ENS, ENS, PSL University, Paris, France

Keywords: Mean Field Control, Optimal Transport, Smart Grids.

Demand-side management has emerged as a key strategy to enhance the flexibility and reliability of modern power systems, particularly in the context of the increasing penetration of renewable energy sources. This is achieved by leveraging the flexibility of a wide range of distributed energy resources. The aggregated control of such devices can provide valuable services to the grid (e.g. load shaping, frequency regulation, or congestion management), thus contributing to system stability and economic efficiency. In demand-side management applications, the number of controllable devices is typically very large, which naturally motivates the use of mean-field control (MFC). Optimal transport (OT) provides an intuitive approach for modeling the evolution of the aggregate distributions considered in the MFC.

We develop a tractable optimization framework for the real-time coordination of heterogeneous loads (e.g. electric vehicles and water heaters), combining ideas from OT and MFC. We first extend the framework of Moment Constrained Optimal Transport for Control (MCOT-C) [1] to a heterogeneous setting. We then propose a model predictive control approach where the agents' data is progressively discovered during the day. The proposed approach is validated through numerical experiments on real datasets [3, 4] for electric vehicles and water heaters, demonstrating the effectiveness of this method in imposing global restrictions while preserving agent-level dynamics[1]. The results show an emerging phenomenon with each system adapting to the (lack of) flexibility of the other. This proves to be very beneficial compared to a separate optimization of the different populations.

References

[1] T. Le Corre, A. Bušić, S. Meyn, Moment Constrained Optimal Transport for Control Applications. arXiv:2208.01958

[2] T. Le Corre, J. Cardinal, A. Bušić, Moment Constrained Optimal Transport for Thermostatically Controlled Loads, 2025, arXiv:2508.20059

[3] Elaad OpenDataset. Available online: `https://platform.elaad.io/analyses/ElaadNL_opendata.php`

[4] J. Albouys, N. Sabouret, Y. Haradji, M. Schumann, and C. Inard. *SMACH: Multi-agent Simulation of Human Activity in the Household*. Springer International Publishing. 2019.

[1]The code used to generate the results is available at `https://github.com/Kreyparion/Heterogeneous-MCOT`.

A two-level healthcare districting problem : A mathematical programming approach

Paulette Castillo[1,3] Sourour Elloumi [1,3] Franco Quezada [2]

[1] CEDRIC-Cnam, 292 rue saint Martin, 75141 Paris, Cedex 03, France
[2] Industrial Engineering Department, University of Santiago of Chile (USACH), Chile
[3] UMA-ENSTA, 828 Boulevard des Maréchaux, 91120 Palaiseau, France

Keywords: Healthcare System, Districting Problem, Mixed-Integer Linear Programming Formulation

The districting problem involves partitioning a geographical region into smaller subregions called districts. This optimization problem has been applied across various domains, including police districting, political redistricting, and healthcare districting, among others [1], with the primary objective of optimizing resource allocation and service delivery. Depending on the specific application, districting solutions must satisfy different criteria and constraints. There are fundamental properties that are particularly critical, such as contiguity [2], population homogeneity [3], which seeks to achieve balanced population distribution across districts, and spatial compactness, which aims to minimize intra-district distances by creating geometrically compact districts.The challenge lies in simultaneously optimizing these often conflicting objectives while satisfying additional domain-specific constraints.

This study investigates the healthcare districting problem, specifically focusing on primary and secondary care levels within healthcare systems. The objective is to minimize the maximum population-resource difference across all districts, with the aim of maximizing the population served or covered by the healthcare system and helping to improve accessibility to healthcare services [4]. The proposed approach considers essential districting properties including population homogeneity, spatial compactness, and territorial contiguity. A key aspect of this work is the simultaneous analysis of both care levels, recognizing that if they are managed separately, then imbalances could occur in the level that is not the focus.

One of the most important challenges in this type of problem is how to impose the contiguity property, so we analyze four different mathematical formulations to ensure this constraint. Additionally, we examine structural properties of these formulations. From this analysis, we derive different strategies such as symmetry-breaking constraints and variable fixing techniques that help us reduce computational time and improve solution quality for instances that are not solved to optimality.

References

[1] R. Z. Ríos-Mercado, *Optimal districting and territory design*, vol. 284, Springer, 2020.

[2] F. Pezzella, R. Bonanno, B. Nicoletti, A system approach to the optimal health-care districting, *European Journal of Operational Research*, vol. 8, no. 2, pp. 139–146, 1981.

[3] H. Farughi, S. Mostafayi, J. Arkat, Healthcare districting optimization using gray wolf optimizer and ant lion optimizer algorithms (case study: South Khorasan healthcare system in Iran), *Journal of Optimization in Industrial Engineering*, vol. 12, no. 1, pp. 119–131, 2019.

[4] S. M. Darmian, M. Fattahi, E. Keyvanshokooh, An optimization-based approach for the healthcare districting under uncertainty, *Computers & Operations Research*, vol. 135, p. 105425, 2022.

On Multidimensional Disjunctive Inequalities for Chance-Constrained Stochastic Problems with Finite Support

Diego Cattaruzza[1] Martine Labbé[2] Matteo Petris[3] Marius Roland[4] Martin Schmidt[5]

[1]University of Udine, Italy, diego.cattaruzza@uniud.it
[2]Université Libre de Bruxelles, Belgium, martine.labbe@ulb.be
[3]École des ponts ParisTech, France, matteo.petris@enpc.fr
[4]Inria Lille, France, mmmroland@gmail.com
[5]Trier University, Germany, martin.schmidt@uni-trier.de

Keywords: Chance constraints, Stochastic optimization, Mixed-integer linear optimization, Valid inequalities

We consider linear Chance-Constrained Stochastic Problems (CCSPs) with finite support, a class of optimization problems known for its challenging non-convex structure. To improve the performance of branch-and-cut solution methods, we introduce a new class of valid inequalities, termed multi-disjunctive inequalities. These inequalities are constructed based on a disjunctive property inherent to the mathematical formulation of CCSPs. We present a heuristic separation procedure for these inequalities and provide extensive numerical results. The computational study demonstrates that our proposed method significantly strengthens the formulation and improves solver performance compared to existing families of valid inequalities from the literature, such as quantile [1] and mixing inequalities [2].

References

[1] W. Xie and S. Ahmed, On quantile cuts and their closure for chance constrained optimization problems, *Mathematical Programming*, 172 (2018), pp. 621-646.

[2] J. Luedtke, S. Ahmed, and G. L. Nemhauser, An integer programming approach for linear programs with probabilistic constraints, *Mathematical Programming* 122 (2010), pp. 247-272.

Lipschitz Continuity of Diffusion Transport Maps from a Control Perspective

Louis-Pierre Chaintron[1] Giovanni Conforti[2] Katharina Eichinher[3]

[1]École Polytechnique Fédérale de Lausanne, Switzerland,
[2]University of Padova, Italy,
[3]Inria ParMA and LMO, Université Paris-Saclay, France.

Keywords: Heat flow transport map, Hamilton Jacobi Bellman equation, Generative modelling

Finding regular transport maps between measures is an important task in generative modelling and a useful tool to transfer functional inequalities. The most well-known result in this field is Caffarelli's contraction theorem, which shows that the optimal transport map from a Gaussian to a uniformly log-concave measure is globally Lipschitz. Note that for our purposes optimality of the transport map does not play a role. This is why several works investigate other transport maps, such as those derived from diffusion processes, as introduced by Kim and Milman. Here, we use the control interpretation of the driving transport vector field inducing the transport map and a coupling strategy to obtain Lipschitz bounds for this map for a big class of what we call asymptotically log-concave measures. This talk is based on a joint work with Louis-Pierre Chaintron and Giovanni Conforti.

One-Sided Linear Couplings with Application to Sparse Optimization

Jean-Philippe Chancelier[1] Michel De Lara[2] Adrien Le Franc[3] Seta Rakotomandimby[4]

[1]CERMICS, École nationale des ponts et chaussées, IP Paris, France, jean-philippe.chancelier@enpc.fr
[2]CERMICS, École nationale des ponts et chaussées, IP Paris, France, michel.delara@enpc.fr
[3]CERMICS, École nationale des ponts et chaussées, IP Paris, France, adrien.lefranc@enpc.fr
[4]CERMICS, École nationale des ponts et chaussées, IP Paris, France, seta.rakotomandimby@enpc.fr

Keywords: optimization, one-sided linear coupling, convex composite, cutting plane

This talk is part of the invited session *Beyond Euclidean Convexity: Methods and Algorithms* (organizers: Pierre-Cyril Aubin and Michel De Lara)

When a one sided linear (OSL) coupling is substituted for the usual duality product in the Fenchel-Moreau conjugacy, OSL-convex functions are substituted for the (closed) convex functions. OSL-convexity is accompanied by the generalization of the usual subdifferential and the associated minimization algorithms from usual convex optimization (Fermat's rule, proximal methods, cutting plane methods, ...). In fact, it has been proven in [3] that OSL-convex functions are convex composite. Algorithms and optimality conditions for the minimization of convex composite functions have been studied in [6, 2, 1, 5].

Here, we present OSL convexity and we study the difficulties for the OSL-minimization of the notorious sparsity inducing ℓ_0-pseudonorm, which happens to be OSL-convex for the so-called Capra coupling [4].

References

[1] James V. Burke and Michael C. Ferris. A gauss-newton method for convex composite optimization. *Mathematical Programming*, 71(2):179–194, 1995.

[2] James V. Burke, and R.A. Poliquin. Optimality conditions for non-finite valued convex composite functions. *Mathematical Programming*, 57(1):103–120, 1992.

[3] J.-P. Chancelier and M. De Lara. Hidden convexity in the ℓ_0 pseudonorm. *Journal of Convex Analysis*, 28(1):203–236, 2021.

[4] J.-P. Chancelier and M. De Lara. Constant along primal rays conjugacies and the ℓ_0 pseudonorm. *Optimization*, 71(2):355–386, 2022.

[5] N. Hallak and M. Teboulle. An adaptive Lagrangian-based scheme for nonconvex composite optimization. *Mathematics of Operations Research*, 48(4):2337–2352, 2023.

[6] R.T. Rockafellar. Directionally lipschitzian functions and subdifferential calculus. *Proceedings of the London Mathematical Society*, 3(2):331–355, 1979.

On Degree of a Branching Point in Branched Optimal Transport (aka Gilbert–Steiner Problem)

Danila Cherkashin[1]

[1]Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Bulgaria,
jiocb@math.bas.bg

Keywords: Gilbert–Steiner problem, branched optimal transport, spherical code

The Gilbert–Steiner problem is a generalization of both the Steiner tree problem and optimal mass transportation problems, allowing the use of additional (branching) points in a transport plan. This problem was first introduced for measures with finite support in a paper by Gilbert in 1967 [4]. The name "Steiner" refers to the branched structure of Steiner trees. A book by Bernot, Caselles, and Morel [1] provides a modern perspective on the problem.

The most widely studied model assumes that the cost of transporting mass m along a segment is proportional to m^p and linear in the segment length, where $0 \leq p < 1$. This cost function, $c(x) = x^p$, satisfies several natural conditions such as monotonicity, concavity, and $c(0) = 0$.

We determine all pairs (p, d) for which the Gilbert–Steiner problem in d-dimensional Euclidean space admits only branching points of degree 3. Namely, this occurs if and only if $d = 2$ or $p < 1/2$. This classification was obtained in [2], building on results from [5] and [3].

An *irrigation setup* is a subproblem in which one of the measures is restricted to be a Dirac measure. As an application of the main result, one can construct a universal solution to the irrigation Gilbert–Steiner problem for $d = 2$ or $p < 1/2$. Universality here means that this solution contains every possible combinatorial structure with the same p and d as a substructure.

References

[1] Marc Bernot, Vicent Caselles, and Jean-Michel Morel. *Optimal transportation networks: models and theory*. Springer, 2008.

[2] Danila Cherkashin. On branching points in the Gilbert–Steiner problem. *arXiv preprint arXiv:2507.13532*, 2025.

[3] Danila Cherkashin and Fedor Petrov. Branching points in the planar Gilbert–Steiner problem have degree 3. *Pure and Applied Functional Analysis*, 10:331–338, 2025.

[4] Edgar N. Gilbert. Minimum cost communication networks. *Bell System Technical Journal*, 46(9):2209–2227, 1967.

[5] Marcus G. Volz, Marcus Brazil, Charl J. Ras, Konrad J. Swanepoel, and Doreen A. Thomas. The Gilbert arborescence problem. *Networks*, 61(3):238–247, 2013.

Optimization landscape of ℓ_0-Bregman relaxations

Jonathan Chirinos-Rodríguez Cédric Févotte Emmanuel Soubies

IRIT, CNRS, Toulouse INP, Université de Toulouse, France.

{jonathan-eduardo.chirinos-rodriguez, cedric.fevotte, emmanuel.soubies}@irit.fr,

Keywords: sparse inverse problems, nonconvex optimization, ℓ_0-Bregman relaxations.

In this presentation, we study (noisy) linear systems, and their ℓ_0-regularized optimization problems, coupled with general convex, not necessarily quadratic, data fidelity terms. Specifically, we aim at finding solutions of $J_0(\mathbf{x}) := F_{\mathbf{y}}(\mathbf{A}\mathbf{x}) + \lambda_0\|\mathbf{x}\|_0$. Recent successful approaches for solving this class of problems have proposed to consider non-convex exact continuous relaxations that preserve global minimizers while reducing the number of local minimizers. Within this framework, we consider the class of ℓ_0-Bregman relaxations, proposed in [3], and whose associated relaxed problem becomes $J_\Psi(\mathbf{x}) := F_{\mathbf{y}}(\mathbf{A}\mathbf{x}) + B_\Psi(\mathbf{x}; \lambda_0)$. Our analysis is primarily built upon a novel property we introduce, termed the Bregman Restricted Strong Convexity (BRSC), which naturally generalizes well-known notions in the literature [1, 4]. Our first result is to establish sufficient conditions under which a critical point of J_Ψ is isolated in terms of sparsity, meaning that any other critical point must have a strictly larger cardinality. Furthermore, we analyze the exact recovery properties of such exact relaxations. Such result is presented in the following.

(Informal) Theorem. *Assume that $F_{\mathbf{y}}$ satisfies* BRSC *with $C_K > 0$, $K \geq 2\|\mathbf{x}^*\|_0$, has $L-$Lipschitz gradient, and*

$$\sqrt{\lambda_0} \in \left(\underline{\Lambda}(\mathbf{y}, C_K, L), \bar{\Lambda}(\mathbf{x}^*, \mathbf{y}, C_K, L)\right).$$

Then, the oracle solution $\mathbf{x}^{\text{or}}$ is the ***unique global minimizer*** *of J_Ψ (and so of J_0), and any other critical point $\mathbf{x}'$ of J_Ψ, $\mathbf{x}' \neq \mathbf{x}^{\text{or}}$ satisfies $\|\mathbf{x}' - \mathbf{x}^{\text{or}}\|_0 \geq K$.*

Finally, we particularize our results for several choices of the data fidelity term: we improve the state-of-the-art results obtained in [2] for the least-squares loss, and derive novel results for the (generalized) Kullback–Leibler divergence.

Acknowledgements. This work has been supported by the project EROSION (ANR-22-CE48-0004) and the Toulouse AI cluster ANITI (ANR-23-IACL-0002).

References

[1] Emmanuel J. Candés, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[2] M. Carlsson, D. Gerosa, and D. Olsson, An unbiased approach to compressed sensing, *Inverse Problems*. 36(11):115014, 2020.

[3] M. Essafri, L. Calatroni, and E. Soubies, Exact continuous relaxations of ℓ_0-regularized criteria with non-quadratic data terms, *Journal of Global Optimization*, 2025.

[4] Jie Shen and Ping Li. A tight bound of hard thresholding. *Journal of Machine Learning Research*, 18(208):1–42, 2018.

A COUNTEREXAMPLE TO SMOOTHNESS OF ENERGY-MINIMIZERS IN SUB-RIEMANNIAN GEOMETRY

YACINE CHITOUR, FRÉDÉRIC JEAN, ROBERTO MONTI, LUDOVIC RIFFORD, LUDOVIC SACCHELLI, MARIO SIGALOTTI, AND ALESSANDRO SOCIONOVO

Keywords. Sub-Riemannian geometry, regularity of geodesics, counterexample.

Abstract

In this talk, we present some recent results concerning the regularity of the curves minimizing the sub-Riemannian energy, also called sub-Riemannian geodesics (they are also length-minimizing curves that are parameterized by the arc-length). In particular, we discuss the first example of a non-smooth (i.e., non-C^∞) geodesic, which disproves a longstanding open conjecture. The example we exhibith is a class C^2 but not C^3 length-minimizer of a real-analytic (even polynomial) sub-Riemannian structure. Moreover, we can prove that sub-Riemannian structure of the counterexample belongs to a class of polynomial sub-Riemannian structures where geodesics are of class C^2. These results are based on the works [1, 2].

References

[1] Yacine Chitour, Frédéric Jean, Roberto Monti, Ludovic Rifford, Ludovic Sacchelli, Mario Sigalotti, and Alessandro Socionovo. Not all sub-Riemannian minimizing geodesics are smooth. *Preprint arXiv*, 2025.

[2] Alessandro Socionovo. Sharp regularity of sub-Riemannian length-minimizing curves. *Preprint arXiv*, 2025.

Optimization under Weakly-Concave Composite Structure: A Proximal-type Approach

Yi Chu[1] Welington de Oliveira[1] Wim van Ackooij[2]

[1]CMA, Mines Paris - PSL, France, yi.chu@minesparis.psl.eu

[2]EDF, France, wim.van-ackooij@edf.fr

Keywords: Nonsmooth optimization, Nonconvex optimization, Stochastic programming

We introduce a proximal-type optimization algorithm designed for minimizing nonsmooth and nonconvex composite functions. The composite structure comprises a weakly-concave outer function and a Lipschitz continuous inner mapping. By linearizing the outer function, the problem is approximated by a nonlinear, potentially nonconvex master program. The algorithm's iterates are defined as stationary points of this master program, guided by an improvement function that integrates both the objective and constraint into a single mathematical object. We establish convergence to critical points and present promising numerical results on chance-constrained optimization problems, demonstrating the practical effectiveness of the proposed approach.

Design of an electric vehicle sharing system under demand uncertainty: a bi-objective risk-averse stochastic programming approach

Christian Clavijo-Lopez[1] Mouna Kchaou-Boujelben [1,2] Céline Gicquel[3]

[1] College of Business and Economics, United Arab Emirates University, United Arab Emirates
[2] Emirates Center for Mobility Research, Al Ain, United Arab Emirates
[3] LISN, Université Paris Saclay, Paris, France
celine.gicquel@lisn.fr

Keywords: Car sharing, electric vehicles, network design, stochastic programming, mixed-integer linear programming, Benders decomposition

Car sharing is a concept that allows individuals to borrow cars on a short-term basis for a rental rate charged by the time or the distance driven. One-way station-based car sharing systems require the user to pick-up the car at any station of the network and drop it off at another station, not necessarily the same as the pick-up one. The recent development of this concept may be explained by its environmental and social benefits and by its potential to promote the use of electric vehicles. However, the success and prosperity of an electric vehicle sharing system (EVSS) heavily depend on the quality of planning of the system design and operations.

This work focuses on optimizing the strategic design of a one-way station-based EVSS. More precisely, we assume that the locations of the charging stations have already been fixed and seek to determine the number of chargers to be installed at each station, together with the fleet size. We seek to explicitly consider uncertainties on the demand for shared car trips both in terms of flow volume and of spatio-temporal distribution of this flow. We propose a bi-objective two-stage stochastic programming model to simultaneously optimize the economic performance of the system and its quality of service under uncertainties of customer trip demand. Our model makes use of a risk measure, the conditional value-at-risk, to control the risk of obtaining a poor customer service level in the worst-case scenarios. Moreover, to enhance the practical relevance of the model, we integrate a detailed representation of tactical decisions on fleet deployment, as well as operational decisions related to vehicle relocation and battery recharging into the modelling of the EVSS design problem.

This integrated approach results in the formulation of a large-size mixed-integer linear program. In order to solve it efficiently, we develop an original approximate Benders decomposition-based method in which explicit information from the scenario sub-problems is included into the master problem formulation under the form of aggregate second-stage variables and constraints. Numerical experiments on randomly generated instances highlight the performance of the developed procedure as compared to a state-of-the-art MILP solver. Finally, using a case study, we investigate the trade-off between economic profitability and quality of service.

Acknowledgment: This work was partially funded by a grant of the Labex Mathématique Hadamard (RD-505).

Bias-Optimal Bounds for SGD: A Computer-Aided Lyapunov Analysis

Daniel Cortild[1,2] Lucas Ketels[1,3] Juan Peypouquet[1] Guillaume Garrigos[3]

[1]University of Groningen, Groningen, Netherlands,
[2]University of Oxford, Oxford, United Kingdom,
[3]Université Paris Cité and Sorbonne Université, CNRS,
Laboratoire de Probabilités, Statistique et Modélisation, F-75013 Paris, France,
daniel.cortild@maths.ox.ac.uk, {l.ketels, j.g.peypouquet}@rug.nl, garrigos@lpsm.paris

Keywords: Stochastic Gradient Descent, Convex Optimization, Performance Estimation Problem, Lyapunov Analysis

The analysis of Stochastic Gradient Descent (SGD) often relies on making some assumption on the variance of the stochastic gradients, which is usually not satisfied or difficult to verify in practice. This paper contributes to a recent line of works which attempt to provide guarantees without making any variance assumption, leveraging only the (strong) convexity and smoothness of the loss functions. In this context, we prove new theoretical bounds derived from the monotonicity of a simple Lyapunov energy, improving the current state-of-the-art and extending their validity to larger step-sizes. Our theoretical analysis is backed by a Performance Estimation Problem analysis, which allows us to claim that, empirically, the bias term in our bounds is tight within our framework.

This talk is based on the paper [1].

References

[1] D. Cortild, L. Ketels, J. Peypouquet, and G. Garrigos. New Tight Bounds for SGD without Variance Assumption: A Computer-Aided Lyapunov Analysis, May 2025. arXiv preprint arXiv:2505.17965.

Integrating Revenue Management into the Supply Chain for New Vehicles' Distribution at Renault Group

Nicolas Corvol[1] Louis Bouvier[1] Axel Parmentier[1]

[1]CERMICS, ENPC, Champs sur Marne, France, nicolas.corvol@eleves.enpc.fr
louis.bouvier@enpc.fr, axel.parmentier@enpc.fr

Keywords: Combinatorial Optimization, Machine learning, Structured Learning, Markov Decision Process

Car dealers must daily decide which vehicles to purchase from the manufacturer to replenish their stock. This is an integrated sales and operations planning problem. On the sales side, dealers must anticipate customer choices. Operational costs include inventory and lot-sizing costs. We introduce a Markov Decision Process (MDP) that models this car dealer inventory replenishment problem.

Four aspects make this problem particularly challenging. First, the problem has a combinatorial nature. Any vehicle that can be produced can be ordered, leading to a catalog of over 10^{30} possible configurations from different option combinations. Second, the uncertainty is endogenous. We do not know what customers would have purchased if different vehicles had been offered. Third, the historical data for demand estimation is censored, which further complicates the problem. Finally, the problem is an MDP with combinatorial state and action spaces, a class of problems known to be difficult to solve.

The core contribution is a novel policy based on Combinatorial-Optimization Augmented Machine Learning (COAML) [1] [2],. Numerical experiments show the performance of the approach.

References

[1] Baty, L., Jungel, K., Klein, P. S., Parmentier, A., and Schiffer, M. (2024). Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows. *Transportation Science*. Available at: arXiv:2304.00789.

[2] Dalle, G., Baty, L., Bouvier, L., and Parmentier, A. (2022). Learning with Combinatorial Optimization Layers: A Probabilistic Approach. Available at: arXiv:2207.13513.

Branched Optimal Transport and Fractal Measures in Type-I Superconductors

Alessandro Cosenza[1], Michael Goldman [2], Melanie Koser [3], Felix Otto [4]

[1] Université Paris Cité, France, acosenza@math.univ-paris-diderot.fr
[2] École Polytechnique, France, michael.goldman@cnrs.fr,
[3] Humboldt University Berlin, Germany, kosermel@hu-berlin.de,
[4] MPI Leipzig, Germany, felix.otto@mis.mpg.de

Keywords: Branched optimal transport, fractal measures, type-I superconductors.

In this talk I will introduce a branched transport problem with weakly imposed boundary conditions. This problem was first derived as a reduced model for pattern formation in type-I superconductors in [1]. For minima of the reduced model with weak boundary conditions, it is conjectured in [2] that the dimension of the boundary measure is non-integer. The conjecture was linked to local scaling laws in [5]. I will present some recent advances in solving this conjecture. This talked is based on some works with Michael Goldman, Melanie Koser and Felix Otto [3, 4].

References

[1] S. Conti, M. Goldman, F. Otto, and S. Serfaty. "A branched transport limit of the Ginzburg-Landau functional". In: *J. Éc. polytech. Math.* 5 (2018), pp. 317–375.

[2] S. Conti, F. Otto, and S. Serfaty. Personal communication.

[3] A. Cosenza, M. Goldman, and M. Koser. *New dimensional bounds for a branched transport problem*. 2024. arXiv: 2411.14547 [math.AP].

[4] A. Cosenza, M. Goldman, and F. Otto. *Concentration phenomena in a branched transport problem in the half space*. In preparation.

[5] G. De Philippis, M. Goldman, and B. Ruffini. *From energy bounds to dimensional estimates in a branched transport model for type-I superconductors*. 2023. arXiv: 2304.12715 [math.AP].

Control of Collision Orbits

Riccardo Daluiso[1] Jean-Baptiste Caillau[1] Alain Albouy[2]
[1]Université Côte d'Azur, CNRS, Inria, LJAD
[2]CNRS, Observatoire de Paris

Keywords: Regularization, Min time trajectories, Collision, Control

In the n-body problem, the forces acting between particles approach infinity when the mutual distances approach zero. Therefore, at collision the dynamics has singularities. Since Levi-Civita and Sundman, the double collision has been *regularized*, i.e. the singularity has been made to disappear by means of algebraic transformations. Levi-Civita obtained first a regularizing transformation of the Kepler problem based on the inverse of the map $z \to z^2$ of the complex plane. This conformal map sends the orbits of the harmonic oscillator in the ones of the Kepler problem. The coordinates transformation is coupled with a slowing down of the motion by means of a time change defined by the differential relation $d\tau = 1/|z|dt$.

Based on classical results, the purpose of the work is the application of the regularization theory to optimal control. We consider the control of a spacecraft under the attraction of a celestial body, where the control is the thrust, and with the goal of minimizing the physical time. This system exhibits singularities in zero, and thus the theory fails in the study of collision orbits. By applying the Levi-Civita regularization to the controlled system, we obtain an affine system in the new time τ where the vector fields are polynomials, thus C^ω. In the reparametrized system, we can broaden the concept of controllability at collision, and apply Pontryagin's principle by allowing orbits that collide or come arbitrarily close to the center of attraction.
The extension of the procedure in three dimensions is not straightforward, as it encounters algebraic and topological obstructions. To extend the study to the spatial case, we need to make use of the so-called KS regularization. However, the KS-map is based on the Hopf fibration so it is a transformation $\mathbb{R}^4 \to \mathbb{R}^3$, and to each point in the Keplerian orbit corresponds a $\mathbb{S}^1$-orbit in the reparametrized system. This entails the addition of one dimension to the phase space, which is now seven-dimensional.

References

[1] A. Albouy, *The Levi-Civita regularizations*, New Frontiers of Celestial Mechanics: theory and applications, 2023.

[2] J.-B. Caillau et al, *Non-integrability of the minimum-time Kepler problem*, Journal of Geometry and Physics 132, pp. 452–459, 2018.

[3] T. Levi-Civita, *Sur la régularisation du problèm des trois corps*, Rend. Mat. Acc. Lincei 24, 1920.

[4] B. Bonnard, J.B. Caillau, and E. Trélat, *Geometric optimal control of elliptic Keplerian orbits*, Discrete and Continuous Dynamical Systems - Series S, pp. 929–956, 2005.

Stochastic Revelation in Sequential Decision-Making with a Pharmaceutical Application

Morteza Davari[1] Bernardo Pagnoncelli[1] Merve Bodur[2]

[1]SKEMA Business School, Lille, France morteza.davari@skema.edu, bernardo.pagnoncelli@skema.edu
[2]The University of Edinburgh, Scotland, UK, merve.bodur@ed.ac.uk

Keywords: Stochastic programming, sequential decision-making, uncertainty revelation

In many real-life situations, such as medical product launches, energy investments, or the rollout of new policies, decision-makers must act before knowing exactly when critical information will become available. We develop new mathematical models that incorporate uncertainty about what will happen and when that uncertainty will resolve. Traditional decision-making tools assume fixed timelines for information revelation; instead, we address more realistic scenarios where information arrives at unpredictable moments, making planning more complex and costly if mishandled.

In this work, we focus on a pharmaceutical application, in particular, the planning and production of medical devices, where firms must be mindful of regulatory approvals. This context has long been recognized as a challenge due to uncertain approval dates, which complicate inventory and production planning decisions (Hill and Sawaya, 2004). The change in demand resulting from regulatory uncertainty is closely connected to that of optimal product rollover, where the timing of phasing out old products and introducing new ones significantly affects demand and, as such, the firm's performance (Lim and Tang, 2006).

Our first contribution is to formulate the problem as a multistage stochastic programming problem (Birge and Louveaux, 2011; Shapiro et al., 2014). We can solve moderate instances using the extensive form of the problem, which provides us with insights into both the solution and how to develop efficient algorithms. We then propose decomposition methods that enable us to solve larger instances, such as those involving more countries and extended time periods, and explore the managerial insights that can be derived from the solutions.

References

[1] A. V. Hill and W. J. Sawaya, "Production planning for medical devices with an uncertain regulatory approval date," *IIE Transactions*, 36(4): 307–317, 2004.

[2] W. S. Lim and C. S. Tang, "Optimal product rollover strategies," *European Journal of Operational Research*, 174(2): 905–922, 2006.

[3] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer, 2nd edition, 2011.

[4] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, 2nd edition, 2014.

Conditional Infimum, Hidden Convexity and One-Sided Linear Couplings

Jean-Philippe Chancelier[1] Michel De Lara[2]

[1]CERMICS, École nationale des ponts et chaussées, IP Paris, France, jean-philippe.chancelier@enpc.fr
[2]CERMICS, École nationale des ponts et chaussées, IP Paris, France, michel.delara@enpc.fr

Keywords: Optimization, conditional infimum, one-sided linear coupling, hidden convexity, S-procedure

This talk is part of the invited session *Beyond Euclidean Convexity: Methods and Algorithms* (organizers: Pierre-Cyril Aubin and Michel De Lara)

Detecting hidden convexity is one of the tools to address nonconvex minimization problems, and find global minimizers. We introduce the notion of conditional infimum, develop the theory, and establish a tower property, relevant for minimization problems. We connect the conditional infimum with one-sided linear and subadditive couplings. Then, we illustrate how the conditional infimum is instrumental in revealing hidden convexity. Thus equipped, we provide a new sufficient condition for hidden convexity in nonconvex quadratic minimization problems, that encompasses and goes beyond known results (with the notion of blocked-sign pair matrix-vector). We also show how the conditional infimum is especially adapted to tackle the so-called S-procedure.

References

[1] A. Ben-Tal and D. den Hertog. Hidden conic quadratic representation of some nonconvex quadratic optimization problems. *Math. Program.*, 143:1–29, 2014.

[2] A. Ben-Tal, D. den Hertog, and M. Laurent. Hidden convexity in partially separable optimization. Working paper, Operations research, 2011.

[3] J.-P. Chancelier and M. De Lara. Conditional infimum and hidden convexity in optimization, Apr. 2021. Preprint hal-02976039, arXiv:2010.13323.

[4] M. De Lara and J.-B. Hiriart-Urruty. A fresh geometrical look at the general S-procedure. *Optimization Letters*, 16(4):1129–1135, May 2022.

[5] M. Volle and M. Barro. General formulation of the S-procedure. *Journal of Convex Analysis*, 31(4):1289–1304, 2024.

Distributionally Robust Standard Quadratic Optimization with Wasserstein Ambiguity

Abdel Lisser[1] Immanuel M. Bomze[2] Daniel de Vicente[2] Heng Zhang[1]

[1]Paris-Saclay University, abdel.lisser@l2s.centralesupelec.fr, heng.zhang@l2s.centralesupelec.fr

[2]University of Vienna, Austria, immanuel.bomze@univie.ac.at, dani.vice@hotmail.com

Keywords: Stochastic optimization, quadratic optimization, distributionally robust optimization, portfolio optimization

The standard quadratic optimization problem (StQP) consists of minimizing a quadratic form over the standard simplex. If the quadratic form is neither convex nor concave, the StQP is NP-hard. This problem has many interesting applications ranging from portfolio optimization to machine learning.

Sometimes, the data matrix is uncertain but some information about its distribution can be inferred, e.g. the first two moments or else a reference distribution (typically, the empirical distribution after sampling). In distributionally robust optimization, the goal is to minimize over all possible distributions in an ambiguity set defined based upon above mentioned characteristics. We explore the distributionally robust StQP and an ambiguity set based upon maximal Wasserstein distance to the sampled empirical distribution.

References

[1] J. Blanchet and K. Murthy, Quantifying distributional model risk via optimal transport, *Mathematics of Operations Research*, 44(2):565–600, 2019.

[2] I. M. Bomze, On standard quadratic optimization problems, *Journal of Global Optimization*, 13:369–387, 1998.

[3] I. M. Bomze, M. Kahr and M. Leitner, Trust Your Data or Not – StQP Remains StQP: Community Detection via Robust Standard Quadratic Optimization, *Mathematics of Operations Research*, 46(1):301–316, 2021.

[4] J. Cheng, E. Delage and A. Lisser, Distributionally robust stochastic knapsack problem, *SIAM Journal on Optimization* 24(3):1485–1506, 2014.

Estimating Maintenance Cost of Offshore Electrical Substations

Solène Delannoy-Pavy[1] Axel Parmentier[2] Vincent Leclère[2]
Manuel Ruiz[1] Cyrille Vessaire[1]

[1]Réseau de Transport d'Électricité (RTE), Paris, France firstname.name@rte-france.com
[2]CERMICS, ENPC, Paris, France firstname.name@enpc.fr

Keywords: Maintenance Optimization, Stochastic Programming

France aims to deploy 45 GW of offshore wind capacity by 2050. Both ownership and maintenance of the offshore substations linking these farms to the grid are the responsibility of the French Transmission System Operator (TSO). In the event of an unscheduled substation shutdown, the TSO must pay significant penalties to producers. Failures when weather conditions prevent access to the substation can quickly snowball into huge losses.

Although the optimization of maintenance for onshore [1] and offshore [2] wind turbines has previously been studied, the specific challenges of offshore substations highlighted above have not been dealt with. The problem is modeled as a Markov Decision Process, where each state reflects the asset's degradation level and ongoing maintenance, and actions correspond to maintenance decisions. The objective is to optimize maintenance planning to minimize penalties, which are proportional to curtailed energy when capacity is limited. Weather scenarios are incorporated to account for their impact on both power production and maintenance feasibility. A multihorizon stochastic optimization model is introduced [3], featuring a bimonthly strategic horizon for advance planning and a daily operational horizon to capture the penalty dynamics under uncertain weather conditions. A significant challenge arises from the exponential growth of the state space with the number of components, making it impossible to solve the optimization problem exactly for real-life substation models. To address this, we propose an approximate solution approach that yields maintenance policies with strong performance.

References

[1] Wu, Yan-ru and Zhao, Hong-shan, *Optimization maintenance of wind turbines using Markov decision processes*, 2010 International Conference on Power System Technology, 10.1109/POWERCON.2010.5666092

[2] Besnard, François and Fischer, Katharina and Bertling Tjernberg, Lina, A Model for the Optimization of the Maintenance Support Organization for Offshore Wind Farms, *Sustainable Energy, IEEE Transactions on Sustainable Energy*, 10.1109/TSTE.2012.2225454.

[3] Kaut, Michal and Midthun, Kjetil and Werner, Adrian and Tomasgard, Asgeir and Hellemo, Lars and Fodstad, Marte, Multi-horizon stochastic programming, *Computational Management Science*, 10.1007/s10287-013-0182-6.

Decomposition for handling storage in power generation expansion planning

Sophie Demassey

[1]Centre de Mathématiques Appliquées, Mines Paris - PSL, France, sophie.demassey@minesparis.psl

Keywords: long-term planning, mixed integer programming, block coordination

System periodicity is exploited in long-term planning to reduce the size of the model by reducing the time horizon to a limited set of representative periods. System dynamic is then captured by a fine time discretization within each period, but this temporal representation prevents to model interdependencies between the periods, including storage conservation.

We consider modelling non-periodic storage in integrated models for long-term expansion planning of power systems, from micro to continental grids, using temporal decomposition and block coordination between investment and operation. We apply an alternating direction method [1] to the deterministic case. We also consider uncertain loads in a robust two-stage model, solved with Bender's cut and column generation [2]. The methods are experimented on the planning of a microgrid and on a MARKAL-TIMES linear programming model of the EU power system [3].

References

[1] S. Demassey, V. Sessa, A. Tavakoli, Alternating direction method and deep learning for discrete control with storage, *Combinatorial Optimization, ISCO 2024*, Lecture Notes in Computer Science 14594:85-96, (2024).

[2] A. Billionnet, M.C. Costa, P.L. Poirion, Robust optimal sizing of a hybrid energy standalone system, *European Journal of Operational Research*, 254(2):565-575, (2016).

[3] G. Siggini, S. Demassey, E. Assoumou, Investment choices and capacities at risk in decarbonizing the EU electric system. *International Conference on Applied Energy* (2019).

Approximately optimal distributed controls for high-dimensional stochastic systems with pairwise interaction through controls

Elise Devey

Inria Paris - Université Paris Dauphine-PSL, France.

Keywords: High-dimensional stochastic control, Heterogeneous Interactions, Mean Field Control, Hamilton-Jacobi equation;

We investigate high-dimensional stochastic optimal control problems involving many cooperating agents who aim to minimize a convex cost functional. We analyze two settings: a full-information model, where each agent observes the states of all others, and a distributed model, where agents only observe their own states.

J. Jackson and D. Lacker already established a theoretical connection between these two models when agents interact through their state, see [1]. In this context, they provide a sharp non-asymptotic bound on the gap between the value functions associated with these two problems. In other words, they quantify to what extent an optimal solution of the distributed control problem can be a good approximation for the optimal control of the full-information problem.

Our paper builds on their work and provides similar results in the context of interaction through the controls. To derive these results, we follow the approach of [1], employing a parallel theory of distributed stochastic control alongside the classical framework, characterizing optimizers through a Hamilton-Jacobi-type equation. The main difficulty of studying interaction through controls instead of interaction through states is that the optimal controls have a less tractable structure, as we lose the decoupling property of the N Hamiltonians. So far, we have to restrict our study to pairwise interaction through controls.

The primary motivation for this study is the application to the management of the flexibility of the electrical system, as, for example, in [2]. Our model involves equipping each node of the electrical grid with small-scale local electricity storage, thereby easing the constraint of supply-demand balance. The goal is to optimize how much electricity each storage unit draws from or injects into the grid. The power flow in each transmission line is bounded by a maximum capacity, and this constraint can be modeled as a pairwise heterogeneous function of all control variables. As a result, the optimal control problem naturally falls within the framework developed in this work.

References

[1] J. Jackson and D. Lacker, Approximately optimal distributed stochastic controls beyond the mean field setting, *The Annals of Applied Probability*, 2025.

[2] A. Séguret, C. Alasseur, J.F. Bonnans, A. De Paola and N. Oudjane, Decomposition of Convex High Dimensional Aggregative Stochastic Control Problems, *Applied Mathematics & Optimization*, 2023

Bang–Bang Light Control in Optogenetic Bioprocess Models: Optimal Protein Production in Yeast

Baptiste Boerkmann[1] Térence Bayen[2] Mustafa Khammash[3] Walid Djema[1*]

[1]Inria, Université Côte d'Azur (UniCA), GreenOwl team, 06560 Valbonne, France
baptiste.boerkmann@inria.fr, walid.djema@inria.fr

[2]Avignon Université, Laboratoire de Mathématiques d'Avignon (EA 2151), 84018 Avignon, France
terence.bayen@univ-avignon.fr

[3]ETH Zürich, Department of Biosystems Science and Engineering, Basel, Switzerland
mustafa.khammash@bsse.ethz.ch

Keywords: Optimal control, optogenetics, synthetic biology, bioprocesses

In this talk, we revisit the optogenetic control framework introduced in [1], where light was used to regulate the unfolded protein response in *Saccharomyces cerevisiae* and proportional–integral–derivative (PID) feedback was shown to improve the yield of folded amylase. Building on this foundation, our recent study [2] investigated the corresponding optimal control problem in a batch culture setting. By reformulating the model and applying the Pontryagin Maximum Principle, we established that the optimal light input is of bang–bang type, a structure confirmed through numerical simulations and of potential practical interest. We will first summarize the key insights of this analysis, and then discuss ongoing efforts to extend the framework to more general bioprocess scenarios, including continuous cultures. Finally, we will see how the structure of the optimal solutions naturally raises the question of whether feedback-inspired strategies may be derived, opening perspectives for experimentally feasible implementations.

References

[1] Benisch, M., Benzinger, D., Kumar, S., Hu, H. and Khammash, M., 2023. *Optogenetic closed-loop feedback control of the unfolded protein response optimizes protein production.* Metabolic Engineering, 77, pp.32-40.

[2] Djema, W., Bayen, T. and Khammash, M., 2025. *Bang–Bang optimal light control for maximum protein production in yeast.* In: *Proceedings of the 64th IEEE Conference on Decision and Control (CDC)*, IEEE. Rio de Janeiro, Brazil.

*This work was supported by the French government through the France 2030 investment plan managed by the National Research Agency (ANR), as part of the Initiative of Excellence Université Côte d'Azur under reference number **ANR-15-IDEX-01**. Corresponding author: walid.djema@inria.fr.

Speeding Up the NSGA-II via Dynamic Population Sizes

Benjamin Doerr[1] Martin S. Krejca[1] Simon Wietheger[2]

[1]LIX, CNRS, École Polytechnique, IP Paris, France,
benjamin.doerr@polytechnique.edu, martin.krejca@polytechnique.edu
[2]Algorithms and Complexity Group, TU Wien,
swietheger@ac.tuwien.ac.at

Keywords: multi-objective optimization, NSGA-II, runtime analysis, dynamic population size, speed-up

Multi-objective evolutionary algorithms (MOEAs) are among the most widely and successfully applied optimizers for multi-objective problems. However, to store many optimal trade-offs (the *Pareto optima*) at once, MOEAs are typically run with a large, static population of solution candidates, which can slow down the algorithm. We propose the *dynamic* NSGA-II (dNSGA-II), which is based on the popular NSGA-II and features a non-static population size. The dNSGA-II starts with a small initial population size of four and doubles it after a user-specified number τ of function evaluations, up to a maximum size of μ. Via a mathematical runtime analysis, we prove that the dNSGA-II with parameters $\mu \geq 4(n+1)$ and $\tau \geq \frac{256}{50}en$ computes the full Pareto front of the OneMinMax benchmark of size n in $O(\log(\mu)\tau + \mu \log(n))$ function evaluations, both in expectation and with high probability. For an optimal choice of μ and τ, the resulting $O(n \log(n))$ runtime improves the optimal expected runtime of the classic NSGA-II by a factor of $\Theta(n)$. In addition, we show that the parameter τ can be removed when utilizing concurrent runs of the dNSGA-II. This approach leads to a mild slow-down by a factor of $O(\log(n))$ compared to an optimal choice of τ for the dNSGA-II, which is still a speed-up of $\Theta(n/\log(n))$ over the classic NSGA-II.

This work is part of the PGMO-funded project *From Mathematical Runtime Analyses to Better Evolutionary Multi-Objective Optimizers* (PI: Benjamin Doerr).

References

[1] B. Doerr, M. S. Krejca, S. Wietheger, Speeding up the NSGA-II via dynamic population sizes. *CoRR abs/2509.01739* (2025), arXiv:2509.01739.

Improved Runtime Guarantees for the SPEA2 Multi-Objective Optimizer

Benjamin Doerr[1] Martin S. Krejca[1] Milan Stanković[2]

[1] LIX, CNRS, École Polytechnique, IP Paris, France,
benjamin.doerr@polytechnique.edu, martin.krejca@polytechnique.edu
[2] École Polytechnique, IP Paris, France,
milan.stankovic@polytechnique.edu

Keywords: multi-objective optimization, evolutionary computation, runtime analysis, SPEA2

Together with the NSGA-II, the SPEA2 is one of the most widely used domination-based multi-objective evolutionary algorithms. For both algorithms, the known runtime guarantees are linear in the population size; for the NSGA-II, matching lower bounds exist. With a careful study of the more complex selection mechanism of the SPEA2, we show that it has very different population dynamics. From these, we prove runtime guarantees for the OneMinMax, LeadingOnesTrailingZeros, and OneJumpZeroJump benchmarks that depend less on the population size. For example, we show that the SPEA2 with parent population size $\mu \geq n - 2k + 1$ and offspring population size λ computes the Pareto front of the OneJumpZeroJump benchmark with gap size k in an expected number of $O((\lambda + \mu)n + n^{k+1})$ function evaluations. This shows that the best runtime guarantee of $O(n^{k+1})$ is not only achieved for $\mu = \Theta(n)$ and $\lambda = O(n)$ but for arbitrary $\mu, \lambda = O(n^k)$. Thus, choosing suitable parameters – a key challenge in using heuristic algorithms – is much easier for the SPEA2 than the NSGA-II.

This work is part of the PGMO-funded project *From Mathematical Runtime Analyses to Better Evolutionary Multi-Objective Optimizers* (PI: Benjamin Doerr).

Nash equilibrium of a non-atomic game with heterogeneous players sensitive to queueing time

Alix Dupont[1], Quentin Petit[1,2], Adrien Seguret[1,2]

[1]OSIRIS, EDF R&D, Palaiseau, France,
[2]Finance for Energy Market Research Centre (FIME), Paris, France.

Keywords: non-atomic game, queueing theory, electric vehicles, optimal transport

With the increasing adoption of Electric Vehicles (EVs), the charging demand at public charging stations with a fast charging mode is expected to grow, leading to congestion, such as long waiting time. Such congestion can influence user's behavior, creating interaction between them. From the point of view of a charging station operator, it is crucial to anticipate such congestion in order to design appropriate strategies (e.g., sizing, pricing, etc.).

In this context, we study a mixed-strategies and non-atomic game, where the players are the EV users, their strategies are the choice of a Charging Station Pool (CSP), and their individual cost include the waiting time at the chosen CSP. As in [1] and [2], it is assumed that the players (i.e., the EV users) impact on congestion is heterogeneous and continuously distributed.

Inspired by [3], we show that the Nash equilibrium is connected to the solution of a social welfare maximization, meaning that the game admits a potential. As a consequence, we demonstrate the existence and, under monotonicity assumption (leading to the convexity of the potential), the uniqueness of the Nash equilibrium. Under additional assumptions, the solution of the social planner, and hence the Nash equilibrium, induces a Voronoï-type partition of the domain, meaning that the Nash equilibrium is in pure strategies. In addition, we propose a numerical scheme that approximates the Nash equilibrium, and show a toy example.

References

[1] Yves Achdou, Guillaume Carlier, Quentin Petit, and Daniela Tonon, "A simple city equilibrium model with an application to teleworking", *Applied Mathematics & Optimization*, vol. 88, no. 2, p. 60, 2023.

[2] Alix Dupont, "Stratégies des Opérateurs pour la Recharge des Véhicules électriques en Espaces Publics avec un Comportement Piloté par l'Utilisateur", PhD thesis, Université d'Avignon, 2025.

[3] Gianluca Crippa, Chloé Jimenez and Aldo Pratelli, "Optimum and equilibrium in a transport problem with queue penalization effect", In *Walter de Gruyter GmbH & Co. KG*, 2009.

Exploiting Subgradient Sparsity in Max-Plus Neural Networks

Ikhlas Enaieh Olivier Fercoq

LTCI, Télécom Paris, Institut Polytechnique de Paris
ikhlas.enaieh@telecom-paris.fr, olivier.fercoq@telecom-paris.fr

Keywords: Artificial neuron design, tropical algebra, sparse updates, subgradient descent

Deep Neural Networks are powerful tools for solving machine learning problems, but their training often involves dense and costly parameter updates. In this work, we use a novel Max-Plus neural architecture in which classical addition and multiplication are replaced with maximum and summation operations respectively. Our interest in these architectures stems from their ability to naturally induce sparsity in both forward and backward passes [3], making them appealing alternatives to dense networks. Indeed, only neurons that contribute to the maximum affect the loss. The $(\max, +)$ and $(\min, +)$ algebras have been applied in previous works, achieving competitive performance in various tasks. For instance, max-plus operators have been used for filter selection and model pruning [4], and the Linear-Min-Max-Plus architecture has been shown to be a universal approximator for continuous functions [1]. These results indicate that moving toward structured sparse architectures does not necessarily sacrifice expressivity.

However, standard backpropagation fails to exploit this sparsity, leading to unnecessary computations. To address this, we propose a sparse subgradient algorithm that explicitly exploits this algebraic sparsity. By tailoring the optimization procedure to the non-smooth nature of Max-Plus models, our method achieves more efficient updates while retaining theoretical guarantees. This highlights a principled path toward bridging algebraic structure and scalable learning. One of the fundamental tools we use to manage sparse updates is short computational trees [2], that allow an efficient use of computing resources in high dimensions. From the theoretical result [1], we also propose a novel initialization for min-max-plus networks, showing encouraging performance in preliminary numerical experiments.

References

[1] Y. Luo and S. Fan. Min-max-plus neural networks. *arXiv preprint arXiv:2205.09997*, 2022.

[2] Y. Nesterov. Subgradient methods for huge-scale optimization problems. *Mathematical Programming*, 146(1-2):275–297, 2014.

[3] A. Tsiamis and P. Maragos. Sparsity in max-plus algebra and systems. *Discrete Event Dynamic Systems*, 29(1):163–189, 2019.

[4] Y. Zhang, S. Blusseau, S. Velasco-Forero, I. Bloch, and J. Angulo. Max-plus operators applied to filter selection and model pruning in neural networks. 2019.

Toward ℓ_0 solution path via ℓ_0 Bregman relaxations

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

[1]IRIT, CNRS, Toulouse INP, Université de Toulouse, Toulouse, France.
(mhamed.essafri@irit.fr, emmanuel.soubies@cnrs.fr)
[2]MaLGa Centre, DIBRIS, Università di Genova & MMS, Istituto Italiano di Tecnologia, Genoa, Italy.
(luca.calatroni@unige.it)

Keywords: ℓ_0 regularization path, non-quadratic data terms, non-convex optimization

Given a possibly underdetermined matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$, with $M \leq N$, and an observation vector $\mathbf{y} \in \mathbb{R}^M$, we consider optimization problems of the form

$$\hat{\mathbf{x}} \in \underset{\mathbf{x}\in\mathbb{R}^N}{\operatorname{argmin}}\, J_0(\mathbf{x}), \quad \text{with} \quad J_0(\mathbf{x}) := F_{\mathbf{y}}(\mathbf{A}\mathbf{x}) + \lambda_0\|\mathbf{x}\|_0 + \frac{\lambda_2}{2}\|\mathbf{x}\|_2^2, \tag{1}$$

where $F_{\mathbf{y}} : \mathbb{R}^M \to \mathbb{R}_{\geq 0}$ denotes a data-fidelity functional, $\|\cdot\|_0$ is the ℓ_0 pseudo-norm counting the number of nonzero components of its input, and $\lambda_0 > 0$ controls the trade-off between data fidelity and sparsity. The ℓ_2 regularization, weighted by $\lambda_2 \geq 0$, acts as a stability term: it penalizes large coefficients, mitigates overfitting in noisy settings, and improves the well-posedness of the optimization problem (existence of a solution) for some data terms. Since the ℓ_0 function is non-continuous and non-convex, Problem (1) is known to be NP-hard [1]. Recently, the family of ℓ_0 Bregman relaxations (B-rex) [2] has been proposed to build a class of *exact continuous relaxations* for ℓ_0-regularized criteria with general fidelity terms of the form (1). These relaxations are exact in the sense that, they preserve the set of global minimizers of the original ℓ_0 problem while eliminating certain local minimizers, thereby simplifying the optimization landscape.

Building on these properties, we introduce a new algorithm, named `L0PathBrex`, to compute the path of solutions across a range of sparsity levels [3]. Our approach exploits the structure of B-rex to build efficient warm-start strategies so as to refine progressively the estimated path of solutions. We detail the methodological aspects of this approach and provide illustrative examples as well as numerical experiments. Extensive comparisons with alternative strategies such as the `L0Learn` package [4], highlight the effectiveness of the proposed algorithm.

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, in *Proceedings of SAMPTA*, Bordeaux, 2019.

[2] M. Essafri, L. Calatroni, and E. Soubies, Exact continuous relaxations of ℓ_0-regularized criteria with non-quadratic data terms, *Journal of Global Optimization*, To appear, 2025.

[3] M. Essafri, L. Calatroni, and E. Soubies, Regularization path via ℓ_0 Bregman relaxations, in *Proceedings of Colloque GRETSI*, Strasbourg, 2025.

[4] H. Hazimeh and R. Mazumder. Fast best subset selection: Coordinate descent and local combinatorial optimization algorithms, in *Operations Research*, 68(5), pp. 1517-1537, 2020.

Finite-Difference Least Square Method for Solving Hamilton-Jacobi Equations Using Neural Networks

Carlos Esteve-Yagüe[1] Richard Tsai[2] Alex Massucco[3]

[1]University of Alicante, Departamento de Matemáticas, Spain, c.esteve@ua.es
[2]University of Texas at Austin, Oden Institute, USA, ytsai@math.utexas.edu
[3]University of Cambridge, DAMTP, UK, am3270@cam.ac.uk

Keywords: Hamilton-Jacobi equations, deep learning, finite-difference methods, least squares principle, optimal control

In recent years, advancements in deep learning and new optimisation algorithms have motivated the use of artificial neural networks to solve non-linear problems in high-dimensional setups. One of the crucial steps during the implementation of any deep learning method is the choice of the loss functional, which is used to train the neural network parameters, typically through a gradient-based method. In this talk, I will consider the approximation of the viscosity solution for Hamilton-Jacobi equations by means of an artificial neural network. I will discuss the choice of the loss functional, which should be such that any critical point approximates the viscosity solution. I will present some recent results concerning loss functionals involving a consistent and monotone numerical Hamiltonian of Lax-Friedrichs type. Using the numerical diffusion built in the numerical Hamiltonian, we are able to prove that any critical point solves the associated finite-difference problem and, therefore, approximates the viscosity solution. I will also present a method in which the numerical diffusion of the numerical scheme is decreased during the training, allowing for approximations with less numerical diffusion. This talk is based on the recent paper[1].

References

[1] C. Esteve-Yagüe, R. Tsai and A. Massucco. Finite-difference least square methods for solving Hamilton-Jacobi equations using Neural Networks, *Journal of Computational Physics* 524, 113721, 2025.

Probabilistic Groundwater Control

M.H. Farshbaf-Shaker[1] H. Heitsch[2] R. Henrion[2]

[1]HTW Berlin - University of Applied Sciences, Germany, Mohammad.Farshbaf@htw-berlin.de
[2]Weierstrass Institute for Applied Analysis and Stochastics, Germany, holger.heitsch@wias-berlin.de, rene.henrion@wias-berlin.de

Keywords: PDE constrained optimization, dynamic probabilistic constraints.

In this talk we consider a probabilistic control of ground water level in the soil of some area between two parallel channels (Dirichlet boundary conditions). The ground water level is influenced by a meteorological source term involving the net difference between time dependent precipitation and evaporation over ground, [5]. The aim is to keep the ground water level in the middle of the area in a given tolerance from a nominal level by applying a suitable control. Since the source term is uncertain, we consider a worst case approach by requiring that the state constraint (evaluated at one specific point) had to be kept in a certain region of tolerance for all times in the considered horizon and for no matter which realization of the source term within a specified set. The big challenge in dealing with probabilistic constraints consists in the fact that no analytical formula for the value and gradient of the probability function is available in general. A successful tool to achieve this goal in the context of random vectors with elliptically symmetric (in particular Gaussian) and related distributions (e.g., log-normal, truncated Gaussian, mixture of Gaussian) is the so-called spheric-radial decomposition (SRD), which allows the simultaneous computation of values and gradients by means of spherical integrals, [1, 2, 3, 4]. Moreover, we will discuss the derivation of an analytic formula for the control-times-uncertainty-to state operator and possibly nonsmooth necessary optimality conditions.

References

[1] W. Van Ackooij and R. Henrion, (Sub-) Gradient formulae for probability functions of random inequality systems under Gaussian distribution. *SIAM/ASA J. Uncertainty Quantification*, 5 (2017) 63-87.

[2] W. Van Ackooij, R. Henrion and P. Pérez-Aros, Generalized gradients for probabilistic/robust (probust) constraints, Optimization, 69 (2020) 1451-1479.

[3] M.H. Farshbaf-Shaker, M. Gugat, H. Heitsch and R. Henrion, Optimal Neumann boundary control of a vibrating string with uncertain initial data and probabilistic terminal constraints, *SIAM J. Control Optim.*, 58 (2020) 2288-2311.

[4] A. Hantoute, R. Henrion and P. Pérez-Aros, Subdifferential characterization of probability functions under Gaussian distribution, *Mathematical Programming*, 174 (2019) 167–194.

[5] B.S. Mordukhovich, Optimal control of the groundwater regime on engineering reclamation systems,*Water Resources*, 12 (1986), 244-253.

Strong Total Cover and Mixed Dominating Set Problems: Polyhedra and Algorithms

Manuel Aprile[1] Luca Ferrarini[2] Emiliano Lancini[3]

[1] University of Padua, Italy, manuel.aprile@unipd.it
[2] Université Sorbonne Paris Nord, LIPN, France, ferrarini@lipn.univ-paris13.fr
[3] LAMSADE, Université Paris Dauphine - PSL, France, emiliano.lancini@lamsade.dauphine.fr

Keywords: Polyhedral Combinatorics, Integer Programming, Mixed Dominating Set, Strong Total Cover

The Strong Total Cover Problem (STCP) is a novel covering problem defined on an undirected graph $G = (V, E)$, where the goal is to find a minimum subset $D \subseteq V \cup E$ such that every element of $V \cup E$ is covered: a vertex covers itself and all incident edges, while an edge covers itself and its two endpoints. This model generalizes the classical Vertex Cover and Edge Cover problems, but the stricter covering rules make it fundamentally different from the well-studied Mixed Dominating Set Problem (MDSP), where a vertex dominates its neighbors and incident edges, while an edge dominates its endpoints and adjacent edges. In this talk, we provide an overview of some results on STCP from polyhedral and algorithmic perspectives, while highlighting its connections to MDSP. We introduce the associated polytope, namely the Strong Total Cover Polytope, and derive several families of valid inequalities, some of which are shown to be facet-defining. In particular, we obtain complete linear descriptions for specific graph classes such as triangulated cactus graphs and cycles. From a computational perspective, we establish complexity results on general graphs and identify classes of graphs for which the problem can be solved in polynomial time. Finally, we investigate approximation aspects of both STCP and MDSP through the study of the maximum subdeterminant of the corresponding constraint matrices.

References

[1] Alavi, Y., Liu, J., Wang, J., and Zhang, Z. (1992). On total covers of graphs. *Discrete Mathematics*, 100(1):229–233.

[2] Bouchakour, M., Contenza, T., Lee, C., and Mahjoub, A. (2008). On the dominating set polytope. *European Journal of Combinatorics*, 29(3):652–661.

[3] Dublois, L., Lampis, M., and Vangelis, T. P. (2021). New algorithms for mixed dominating set. Discrete Mathematics and Theoretical Computer Science, 23(1).

[4] Ferrarini, L., Fiorini, S., Kober, S., and Yuditsky, Y. (2024). Total Matching and Subdeterminants. In Basu, A., Mahjoub, A.R., Salazar González, J.J. (eds) *Combinatorial Optimization. ISCO 2024. Lecture Notes in Computer Science*, 14594.

An iterative deal approach to increase self-consumption for a high-voltage energy community

Baptiste Ferrato[1,3] Sandra Ulrich Ngueveu[1] Frédéric Messine[2]
Christophe Aubigny[3] Florian Allard[3]

[1]LAAS-CNRS, Université de Toulouse, INP, Toulouse, France, {bferrato, ngueveu}@laas.fr
[2]LAPLACE, ENSEEIHT, Toulouse, France, frederic.messine@laplace.univ-tlse.fr
[3]Selection-EnR, CRT Primes, Tarbes, France {christophe.aubigny, florian.allard}@selection-enr.com

Keywords: Mathematical Programming, Game Theory, Energy Community

Within an energy community, the electricity consumed is generally not correlated with the energy produced. Thus, two types of periods can occur: (1) a period when consumption exceeds production, meaning consumers must purchase the shortfall from a supplier, and (2) a period when consumption is lower than production, meaning producers must sell their surplus back to the grid. Limiting these periods of deficit and surplus would maximize self-consumption within the community, thereby reducing energy costs.

The objective of this work is to establish a strategy to encourage companies to modify their consumption in order to make the best use of the energy produced within the community.

Demand side management encompasses various approaches aimed at aligning consumption with production. Among these, load shifting is the most widespread method, but in our context, flexible load shaping is more appropriate [1]. This approach relies on the willingness and the ability of consumers to adjust their demand in exchange for incentives.

The issue addressed in this work has been mainly studied in the context of low-voltage networks [1, 2]. The case of high-voltage networks, typically associated with companies, has been explored much less and has several specifities. In particular, for legal and regulatory reasons, the communities are composed of few consumers but of high impact on the electricity network, compared to low-voltage networks were there are a large number of consumers with low individual impact on the electricity network. In addition, the consumption profiles of such consumers are less predictable than those of individuals, making their management more complex.

We define consumer flexibility as his ability to shift consumption throughout the day. Assuming the consumer flexibilities are unknown, we propose an iterative deal mechanism aimed at increasing the energy community's self-consumption rate. The problems of calculating the deals and the consumer responses can be formulated as MILPs.

References

[1] Hala Alsalloum. Gestion décentralisée des interactions complexes entre producteurs et consommateurs d'énergie électrique. *PhD thesis*, Université de technologie de Troyes, 2021.

[2] M.N.Q. Macedo, J.J.M. Galo, L.A.L. de Almeida, and A.C. de C. Lima. Demand side management using artificial neural networks in a smart grid environment. *Renewable and Sustainable Energy Reviews*, 41 :128-133, 2015.

Integrating crew scheduling and cyclic crew rostering for rail freight with train delays

Héloïse Gachet[1] Frédéric Meunier[2] Juliette Pouzet[2]

[1]École nationale des ponts et chaussées, SNCF DTIPG, heloise.gachet@enpc.fr
[2]École nationale des ponts et chaussées, frederic.meunier@enpc.fr
[3]SNCF DTIPG, juliette.pouzet@sncf.fr

Keywords: Rail freight, crew scheduling, crew rostering, column generation, stochastic optimization

In the railway industry, as well as in other branches of transportation, the construction of yearly schedules for drivers is a key step of the planning process [1]. Due to the high complexity of the problem, the SNCF applies a classical two-step approach. In a first step—the Crew Scheduling—, driving tasks (assumed to be the same every week) are combined together to form daily duties. In a second step—the Cyclic Crew Rostering—, these duties are organized so as to form cyclic rosters, one per team of drivers.

The sequential nature of this approach results in sub-optimal solutions [2]. Additionally, this approach does not account for uncertainties, which is known to lead to increased costs due to train delays.

In this work, we propose to modify this approach by handling the two steps simultaneously and by including stochastic constraints, so as to tackle both suboptimality and train delays. This "integrated" approach relies on column generation and models exactly the impact of train delays on crew rosters when the delay distributions are known. A contribution is the modeling of the pricing sub-problem as a Resource Constrained Shortest Path Problem, within the setting of Parmentier [3]. In particular, distributions are kept as a resource and the stochastic order is used to discard paths in a label correcting path enumeration algorithm.

In the case where there is no train delays, the integrated approach has already been able to provide a solution to a regional freight instance of 280 trains, with a cost 8% lower that the one achievable with the current sequential approach. In the case with train delays, the integrated approach showed significant reduction in the impact of these delays on the previous solution, while having a cost lower than with the current sequential approach.

References

[1] T. Breugem, T. Dollevoet, D. Huisman, Analyzing a Family of Formulations for Cyclic Crew Rostering, *ATMOS 2020*, 85:7:1–7:16, 2020.

[2] D.-Y. Lin, M.-R. Tsai, Integrated Crew Scheduling and Roster Problem for Trainmasters of Passenger Railway Transportation, *IEEE Access*, 7:27362–27375, 2019.

[3] A. Parmentier, Algorithms for non-linear and stochastic resource constrained shortest path, *Math Meth Oper Res*, 89:281-317, 2019.

Metric Extrapolation and Generalized Wasserstein Barycenter

Thomas Gallouët [1] Andrea Natale [1] Gabriele Todeschi[2]

[1]Université Paris Saclay-LMO et Inria ParMA, France, thomas.gallouet@inria.fr,andrea.natale@inria.fr
[2]École des ponts, France, gabriele.todeschi@enpc.fr,

Keywords: Wasserstein extrapolation, Wasserstein generalized Barycenter, convex optimization

In this talk we will propose a notion of geodesic extrapolation in the Wasserstein space and more generally in a metric space. This notion appears to be a convex optimization problem. We then discuss an extension of this problem: the Generalized Wasserstein Barycenter. It is a Wasserstein barycenter with multiple positive and negative weights. Finally we will discuss a convex relaxation of this problem and some cases when the relaxation appears to be tight.

References

[1] Gallouët, T., Natale, A., Todeschi, G. (2025). Metric extrapolation in the Wasserstein space. *Calculus of Variations and Partial Differential Equations*, 64(5), 147. arXiv:2407.10516.

Last-Iterate Complexity of SGD for Convex and Smooth Stochastic Problems

Guillaume Garrigos[1] Daniel Cortild[2] Lucas Ketels[1,2] Juan Peypouquet[2]

[1] Université Paris Cité and Sorbonne Université, CNRS, France, garrigos@lpsm.paris, ketels@lpsm.paris
[2]University of Groningen, Netherlands, j.g.peypouquet@rug.nl, d.cortild@rug.nl

Keywords: Stochastic gradient descent, Convex smooth minimization, Last iterate, Algorithmic complexity.

Stochastic Gradient Descent (SGD) is an algorithm widely used to tackle high dimensional convex (possibly smooth) optimization problems that are very common in machine learning. This algorithm has first been proposed in the pioneering work of [1] but was studied under very strong variance assumptions that are virtually never verified in practice. Since then, an extensive body of work has emerged to develop frameworks that are both realistic and provide rigorous performance guarantees for SGD. In this talk, we propose to provide a brief yet thorough account of the assumptions underlying the analysis of SGD, highlighting the contexts where each has been relevant. However, despite a very rich literature, some open questions still surround SGD. Historically, most results on SGD in the convex and smooth setting were presented in the form of bounds on the ergodic function value gap and until very recently it remained unknown if the last iterate could achieve similar performance. We will present a result that answers this question positively, a result that was independently established in [2].

References

[1] Robbins, H and Monro, S (1951). A Stochastic Approximation Method, *Then Annals of Mathematical Statistics*, 22(3):400-407.

[2] Attia, A., Schliserman, M., Sherman, U., Koren, T.: Fast Last-Iterate Convergence of SGD in the Smooth Interpolation Regime. arXiv preprint arXiv:2507.11274 (2025)

SAA Consistency for Optimal Control Problems with Almost Sure State Constraints

Caroline Geiersbach[1] Johannes Milz[2]

[1]University of Klagenfurt, Austria, caroline.geiersbach@aau.at
[2]Georgia Institute of Technology, USA, johannes.milz@isye.gatech.edu

Keywords: conic constraints, sample average approximation, regularization, PDE-constrained optimization

This talk is concerned with a class of risk-neutral stochastic optimization problems defined on a Banach space with almost sure conic-type constraints. This kind of problem appears in the context of optimal control with random differential equation constraints where the state of the system is further constrained almost surely. For this class of problems, we investigate the consistency of optimal values and solutions corresponding to sample average approximation (SAA) as the sample size is taken to infinity. Consistency is also shown in the case where a Moreau–Yosida-type regularization of the constraint is used. The existence of Lagrange multipliers can be guaranteed under Robinson's constraint qualification with an appropriate choice of function space for the constraint. Our assumptions allow us to also show consistency of SAA Karush–Kuhn–Tucker conditions. This work provides theoretical justification for the numerical computation of solutions frequently used in the literature and in experimentation.

References

[1] C. Geiersbach, J, Milz, Sample-based consistency in infinite-dimensional conic-constrained stochastic optimization, *Submitted*, 2025, arXiv:2507.06982

Classification-Based Linear Surrogate Modeling of Constraints for AL-CMA-ES

Oskar Girardin[1] Nikolaus Hansen[1] Dimo Brockhoff[1] Anne Auger[1]

[1]Inria, CMAP, CNRS, École polytechnique
Institut Polytechnique de Paris, France, firstname.lastname@inria.fr

Keywords: Derivative-free optimization, Blackbox optimization, Continuous optimization, Constrained optimization, Surrogate modeling

We introduce linear surrogate functions for modeling inequality constraints to solve constrained blackbox optimization problems with the Augmented Lagrangian CMA-ES [1, 2]. Each surrogate is constructed from a binary classifier that predicts the sign of the constraint value. The classifier, and consequently the resulting algorithm, is invariant under sign-preserving transformations of the constraint values and can handle binary, flat, and deceptive constraints. Somewhat surprisingly, we find that adopting a sign-based classification model of the constraints allows to solve classes of constrained problems which cannot be solved with the original Augmented Lagrangian method using the true constraint value.

References

[1] Asma Atamna, Anne Auger, and Nikolaus Hansen. On invariance and linear convergence of evolution strategies with augmented Lagrangian constraint handling. *Theoretical Computer Science*, 832:68–97, September 2020. ISSN: 03043975. doi: 10.1016/j.tcs.2018.10.006.

[2] Nikolaus Hansen and Andreas Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001. MIT Press.

Towards Trustworthy AI for Power System Operation

Bastien Giraud[1] Spyros Chatzivasileiadis[2]

[1]DTU Denmark, bagir@dtu.dk
[2]DTU Denmark, spchatz@dtu.dk

Keywords: Trustworthy AI, Dynamic Security Assessment, Physics-Informed Neural Networks, Neural Network Verification

As power systems integrate increasing shares of renewables, distributed energy resources, and complex market mechanisms, ensuring secure and efficient operation becomes ever more challenging. Conventional approaches such as dynamic security assessment and AC Optimal Power Flow (AC-OPF) are computationally demanding, which has motivated the use of machine learning (ML) to reduce computational cost and enhance situational awareness. However, several barriers remain: the lack of high-quality training data can limit model performance; power system physics must be respected to ensure feasibility; and the black-box nature of ML methods hinders their adoption in critical infrastructure.

In this talk, I will present recent advances addressing these challenges. First, I will highlight the role of high-quality datasets and show how to generate samples that meaningfully capture system security boundaries, reflecting both static and dynamic criteria [1]. Second, I will present physics-informed graph neural networks as promising tools for efficient N-k contingency screening [2]. Finally, I will discuss how neural network verification can be used to rigorously enforce constraints [3], and introduce verification-informed ML techniques for AC-OPF that explicitly minimize worst-case violations during training.

Together, these approaches provide certified guarantees of constraint satisfaction, ensure physically realizable operating points, and enable substantial computational savings, offering a pathway toward trustworthy ML for secure power system operation.

References

[1] B. Giraud, L. Charles, A. M. Nakiganda, J. Vorwerk, and S. Chatzivasileiadis, "A Dataset Generation Toolbox for Dynamic Security Assessment: On the Role of the Security Boundary," *arXiv preprint arXiv:2501.09513*, 2025.

[2] A. Nakiganda, S. Chatzivasileiadis, "Graph Neural Networks for Fast Contingency Analysis of Power Systems", 2025. Online https://arxiv.org/abs/2310.04213

[3] B. Giraud, R. Nellikath, M. M. Alowaifeer, J. Vorwerk, and S. Chatzivasileiadis, "Neural Networks for AC Optimal Power Flow: Improving Worst-Case Guarantees during Training", preprint, 2025. *To be available on ArXiV.*

Bollobás–Meir Conjecture for the TSP in the Unit Cube Holds Asymptotically

Alexey Gordeev[1]

[1]Umeå University, Sweden, gordalserg@gmail.com

Keywords: travelling salesman problem, sphere packing, spherical code, perfect matching

In 1992, Bollobás and Meir [1] showed that for any n points in the k-dimensional unit cube $[0,1]^k$, $k \geq 2$, one can always find a tour $x_1, \ldots, x_n$ through these n points with

$$\sum_{i=1}^{n} |x_i - x_{i+1}|^k \leq c_k \cdot k^{k/2},$$

where $|x-y|$ is the Euclidean distance between x and y, $x_{n+1} = x_1$, and c_k is an absolute constant depending only on k. Remarkably, this bound does not depend on n, the number of points. They further conjectured that the best possible constant for every $k \geq 2$ is $c_k = 2$ and showed that it cannot be taken lower than that. The conjecture is only known to be true for $k = 2$ due to Newman [2]. Recently, Balogh, Clemen and Dumitrescu [3] revised the conjecture for $k = 3$ by showing that $c_3 \geq 4 \cdot (\frac{2}{3})^{\frac{3}{2}} > 2.17$. They also gave the best currently known bounds $c_k \leq \frac{2}{3} \cdot 6.709^k$ and, for large k, $c_k \leq 2.91^k \cdot (1 + o_k(1))$.

We reduce the gap between the lower and upper bounds on c_k from exponential to linear by showing that $c_k \leq \min(6(k+1), 2e(k+2))$. Moreover, we prove that for large k the conjecture holds asymptotically, i.e. show that $c_k = 2 + o_k(1)$. We obtain similar results for related problems on Hamiltonian paths and perfect matchings in the unit cube. To achieve this, we substantially improve the ball packing argument used in previous works and exploit connections with optimal (with respect to the minimum angle) spherical codes.

References

[1] B. Bollobás, A. Meir, A Travelling Salesman Problem in the k-Dimensional Unit Cube, *Operations Research Letters*, 11(1):19–21, 1992.

[2] D. J. Newman, A Problem Seminar, Springer, New York, 1982.

[3] J. Balogh, F. C. Clemen, A. Dumitrescu, On a Traveling Salesman Problem for Points in the Unit Cube, *Algorithmica*, 86(9):3054–3078, 2024.

Counter-examples in first-order optimization: a constructive approach and application to Heavy-Ball algorithm

Baptiste Goujaud[1] Aymeric Dieuleveut[2] Adrien Taylor[3]

[1]Telecom SudParis, France, baptiste.goujaud@telecom-sudparis.eu
[2]Ecole Polytechnique, France, aymeric.dieuleveut@polytechnique.edu
[3]INRIA Paris, France, adrien.taylor@inria.fr

Keywords: Optimization, Heavy-Ball, Performance Estimation, Cycle, Non-acceleration

While many approaches were developed for obtaining worst-case complexity bounds for first-order optimization methods in the last years, there remain theoretical gaps in cases where no such bound can be found. In such cases, it is often unclear whether no such bound exists (e.g., because the algorithm might fail to systematically converge) or simply if the current techniques do not allow finding them. In this work, we propose an approach to automate the search for cyclic trajectories generated by first-order methods. We then 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 $1 - O(\kappa)$), 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 techniques. 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 a class of functions that also satisfy higher-order regularity conditions.

Strongly Quasiconvex Functions: What We Know (so far)

Sorin-Mihai Grad[1] Felipe Ignacio Lara Obreque[2] Raúl Tintaya Marcavillaca[2]

[1]University of Wonderland, France, sorin-mihai.grad@ensta.fr

[2]Instituto de Alta investigación, University of Tarapacá, Arica, Chile, flarao@academicos.uta.cl

[3]Centro de Modelamiento Matemático, Universidad de Chile, Santiago de Chile, Chile rtintaya@dim.uchile.cl

Keywords: Strongly quasiconvex functions, generalized convexity, proximal point algorithms

Introduced by Polyak in 1966, the class of strongly quasiconvex functions includes some interesting nonconvex members, like the square root of the Euclidean norm or ratios with a nonnegative strongly convex numerator and a concave and positive denominator. In this talk we survey the most relevant examples of strongly quasiconvex functions and results involving them available in the literature at the moment. In particular, we recall some recent algorithms for minimizing such functions, and hint toward some directions where additional investigations would be welcome.

References

[1] S.-M. Grad, F. Lara, R.T. Marcavillaca, Strongly quasiconvex functions: what we know (so far), *Journal of Optimization Theory and Applications*, 205:38, 2025.

[2] S.-M. Grad, F. Lara, R.T. Marcavillaca, Relaxed-inertial proximal point type algorithms for quasiconvex minimization, *Joural of Global Optimization*, 85:615–635, 2023.

[3] F. Lara, On strongly quasiconvex functions: existence results and proximal point algorithms, *Journal of Optimization Theory and Applications*, 192:891–911, 2022.

Airports Energy Optimization with Optimal Control and AI

Gabriel Gros
Telecom Paris & VINCI Concessions, gabriel.gros@telecom-paris.fr

Keywords: Optimal Control, Optimization of large-scale systems, Statistical Estimation

We address the problem of energy-efficient control of *Heating, Ventilation, and Air Conditioning* (HVAC) systems in airports which account for more than 40% of their overall consumption [1]. For this task, classical PID (*Proportional Integral Derivative*) controllers are generally used but do not explicitly account for energy consumption and lack predictive capability regarding future disturbances. As airports represent an especially demanding case due to large-scale nonlinear dynamics and exogenous factors such as weather and passenger flows, there is a need for new control algorithms dedicated to the HVAC.

To do so, we formulate the task as a *Constrained Rolling Horizon Optimal Control* problem where the objective is to reduce energy use while ensuring comfort and respecting operational constraints [2]. We focus on *direct multiple-shooting transcription* to obtain a *Non-Linear Problem*, which can then be solved via *Sequential Quadratic Programming* or *Interior-Point* methods as they naturally handle constraints and scale to large systems. It provides a solid baseline compared with PID controllers and our preliminary results confirm measurable gains in efficiency while maintaining comfort.

The main challenges lie in obtaining accurate information about the system dynamics and handling uncertainty. First, we investigate parametric estimation of equipment consumption in order to construct a cost function directly from operational data. Secondly, we consider grey-box approaches such as *Autoregressive Models* for interpretability and black-box parameterizations such as recurrent and graph-based *Neural Networks* to capture nonlinear spatio-temporal couplings between airport zones. In parallel, we study both statistical and learning-based approaches for uncertainty estimation. Among these, *Variational Autoencoders* (VAEs) appear as a promising direction to capture latent variability and generate probabilistic forecasts to be integrated into predictive control. We also plan to explore robust and stochastic model predictive control to compare worst-case guarantees with probabilistic estimation performance.

This aims to contribute both methodologically and practically by establishing scalable strategies for embedding data-driven models into Optimal Control. One hopes that the developed methods can generalize across airports with diverse topologies and operating conditions.

References

[1] S. Taheri et al. Model predictive control of heating, ventilation, and air conditioning (hvac) systems: A state-of-the-art review. *Journal of Building Engineering*, 60, 2022.

[2] J. Drgna et al. All you need to know about model predictive control for buildings. *Annual Reviews in Control*, 50:190–232, 2020.

Preconditioned Primal Dual Algorithms in Convex Optimization

Huiyuan Guo[1] Juan José Maulén[2] Juan Peypouquet[1]

[1]University of Groningen, the Netherlands, hazel.guo@rug.nl, j.g.peypouquet@rug.nl,
[2]University of O'Higgins, Chile, jmaulen@dim.uchile.cl

Keywords: Convex optimization, Saddle point problem, Primal dual algorithm

We study a family of preconditioned primal dual algorithms for convex-concave saddle point problems by the dynamics introduced in [1]. For algorithms of proximal type, which have a simple form in the case of linear constrained problems, we establish a non ergodic linear convergence rate. These algorithms are different from the primal dual algorithms in [2] and [3]. For algorithms of gradient type or with a gradient type subiteration, we establish a non ergodic convergence rate of $\mathcal{O}(1/k)$ and show that the sequence of iterates weakly converges to a primal dual optimal solution. Finally, we present numerical experiments to indicate our well performed preconditioned primal dual algorithms.

References

[1] V. Apidopoulos, C. Molinari, J. Peypouquet, S. Villa. Preconditioned primal-dual dynamics in convex optimization: non-ergodic convergence rates, (2025), arXiv:2506.00501.

[2] A. Chambolle, T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging, *Journal of mathematical imaging and vision*, 40(1):120–145, 2011.

[3] H. Luo. A primal-dual flow for affine constrained convex optimization, *ESAIM: Control, Optimisation and Calculus of Variations*, 28:33, 2022.

Distributed games with jumps: An α-potential game approach

Xin Guo[1] Xinyu Li[2] Yufei Zhang[3]

[1]University of California, Berkeley, USA xinguo@berkeley.edu
[2]University of Oxford, UK xinyu.li@maths.ox.ac.uk
[3]Imperial College London, UK yufei.zhang@imperial.ac.uk

Keywords: Potential game, Nash equilibrium, distributed game, policy gradient algorithm, game-theoretic motion planning

Motivated by game-theoretic models of crowd motion dynamics, this paper analyzes a broad class of distributed games with jump diffusions within the recently developed α-potential game framework. We demonstrate that analyzing the α-Nash equilibria reduces to solving a finite-dimensional control problem. Beyond the viscosity and verification characterizations for the general games, we explicitly and in detail examine how spatial population distributions and interaction rules influence the structure of α-Nash equilibria in these distributed settings, and in particular for crowd motion games.

Our theoretical results are supported by numerical implementations using policy gradient-based algorithms, further demonstrating the computational advantages of the α-potential game framework in computing Nash equilibria for general dynamic games.

Flow of Gas Mixtures in Networks: Modeling and Multi-Objective Optimization

Simone Göttlich[1] Michael Schuster[2] Alena Ulke[1]

[1]University of Mannheim, Germany, ulke@uni-mannheim.de,
[2]Friedrich-Alexander University Erlangen-Nürnberg, Germany

Keywords: pipeline networks, gas mixtures, hydrogen blending, multi-objective optimization

As climate change necessitates a transition away from fossil fuels, hydrogen emerges as a promising alternative. Additionally, hydrogen-natural gas mixtures provide a viable interim strategy for reducing greenhouse gas emissions. As a result, there is a growing interest in modeling and optimizing the flow of these mixtures in networks. Operating these networks requires satisfying physical and safety constraints while balancing multiple objectives, such as maximizing profit, minimizing costs, increasing the hydrogen share, or reducing emissions, which often conflict with each other.

To address these challenges, we present a mathematical model, the so-called *mixture model*, that describes the flow of gas mixtures in networks. The model is based on the physical principles of the isothermal Euler equations and the mixing of incoming flow at nodes. Building on this model, we formulate a framework for multi-objective optimization problems that reflect the competing objectives occurring in gas network operation. This framework allows us to compute optimal compromises between the objectives and we show that for a certain class of objectives such optimal compromises exist. We also provide a numerical example to showcase the effectiveness of the proposed idea.

Decision Focused Scenario Generation for Contextual Two-Stage Stochastic Linear Programming

Jonathan Hornewall[1] Solène Delannoy-Pavy[2] Tito Homem-de-Mello[3] Vincent Leclere[1]

[1] ENPC, Institut Polytechnique Paris, jonathan.hornewall@enpc.fr, vincent.leclere@enpc.fr
[2] RTE, ENPC, Institut Polytechnique Paris, solene.delannoy-pavy@enpc.fr
[3] Universidad Adolfo Ibáñez, tito.hmello@uai.cl

Keywords: contextual stochastic programming, stochastic programming, decision-focused learning, log-barrier methods, scenario generation

We introduce a decision-focused scenario generation framework for contextual two-stage stochastic linear programs that bypasses explicit conditional distribution modeling. A neural generator maps a context x to a fixed-size set of scenarios $\{\xi_s(x)\}_{s=1}^{S}$. For each generated collection we compute a first-stage decision by solving a single log-barrier regularized deterministic equivalent whose KKT system yields closed-form, efficiently computable derivatives via implicit differentiation.

The network is trained end-to-end to minimize the true (unregularized) downstream cost evaluated on observed data, avoiding auxiliary value-function surrogates, bi-level heuristics, or differentiation through generic LP solvers. Unlike single-scenario methods, our approach natively learns multi-scenario representations; unlike distribution-learning pipelines, it scales without requiring density estimation in high dimension. We detail the barrier formulation, the analytic gradient structure with respect to second-stage data, and the resulting computational trade-offs.

Preliminary experiments on contextual synthetic instances illustrate that the method can rival current state-of-the-art methods, even when trained on small amounts of training data.

A Saddle Point Algorithm for Stochastic Multi-Objective Optimization with Inequality Constraints

Zachary Jones[1] Pietro Congedo[1] Olivier Le Maître[2]

[1]Inria Saclay and CMAP École Polytechnique, France, zachary.jones@inria.fr, pietro.congedo@inria.fr
[2]CNRS, Inria, and CMAP École Polytechnique,France, olivier.le-maitre@inria.fr

Keywords: Stochastic Optimization, Constrained Optimization, Multi-Objective Optimization

Consider the problem of constrained stochastic multi-objective optimization. We are interested in minimizing a set of objectives subject to inequality constraints, all of which are twice differentiable and have the functional form of an expectation of a stochastic quantity of interest.

Since we rarely have access to analytic expressions for the expectations of either the objectives or the constraints, we try to find the minima through stochastic approximation. However, complications arise due to the uncertain nature of the stochastic quantities of interest. Not only is it complicated to project onto the constraint set, but the evolution of the primal and dual problems are now coupled, leading to oscillations that can inhibit convergence to a saddle point without setting an arbitrary upper limit [1]. Moreover, even after arriving at a stationary solution which is asymptotically viable, it is still relatively unexplored how to assess its admissibility.

We propose a saddle point algorithm to solve constrained stochastic multi-objective optimization problems. The first ingredient to our approach is the stochastic multi-gradient, which we use as a common direction of descent for all objectives [2]. The second is a dual step which dynamically adjusts the value of the Lagrange multipliers for the problem.

We go on to prove the convergence of our approach almost surely to the Pareto set.

References

[1] Harold Joseph Kushner and Dean S Clark. *Stochastic approximation methods for constrained and unconstrained systems*. Vol. 26. Springer Science Business Media, 2012.

[2] F. Poirion Q. Mercier and J.-A. Désidéri. "SMGDA: an uncertainty based multi- objective optimization approach. Illustration to an airplane composite material". In: *Procedia engineering* 199 (2017), pp. 1199–1203

Dominance Properties for Electricity Distribution Planning with Battery Cycling Considerations

Natalia Jorquera-Bravo [1,2,3] Sourour Elloumi[1,2,4] 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
[3]PDSPS, Department of Industrial Engineering, University of Santiago of Chile.
[4]ENSIIE, 91025 Evry, France.

Keywords: Production Planning, Battery Degradation, Self Consumption

In this work, we study collective self-consumption energy communities formed within residential buildings, where houses jointly invest in photovoltaic (PV) panels and a shared electricity storage system. Each house can satisfy its electricity demand using PV generation, the shared battery, or the main grid. Any surplus PV energy can either be stored in the battery or sold back to the grid; however, direct energy exchanges between houses are not permitted. Each house is equipped with a smart meter that monitors electricity consumption over time.

A central community manager is responsible for allocating the PV-generated electricity at each time step. The operational plan must satisfy key technical constraints: (i) a house cannot charge and discharge the battery simultaneously, and (ii) a house cannot buy and sell electricity to the grid within the same period.

We incorporate battery health considerations by explicitly tracking the number of charging cycles. A charging cycle is defined as the period between the transition from discharging to charging and the subsequent transition from charging to discharging. This feature allows us to better capture the long-term operational impact of storage usage.

We formulate this problem as a mixed-integer linear programming (MILP) model to determine an optimal electricity distribution plan for the community, accounting for battery cycling dynamics. Moreover, we exploit structural properties of the problem to show that it can be solved efficiently using dynamic programming techniques inspired by classical lot-sizing algorithms [1, 2]. Finally, we test our algorithms on real-life instances to validate their performance and practical applicability.

References

[1] Harvey M. Wagner and Thomson M. Whitin, Dynamic version of the economic lot size model, Management science, INFORMS (1958)

[2] Stephen F. Love, Bounded production and inventory models with piecewise concave costs, Management science, INFORMS (1973)

Learning Value Functions with High-order Information

Dante Kalise[1]

[1]Imperial College London, United Kingdom, dkaliseb@ic.ac.uk

Keywords: Optimal control, value function, high-dimensional approximation, supervised learning, high-order information

This talk is devoted to recent results on the approximation of high-dimensional HJB PDEs with supervised learning. We will discuss supervised learning of high-dimensional value functions using high-order data, a.k.a. Sobolev training [1]. We expand our previous work on gradient-augmented learning for value functions [2] to incorporate second-order (Hessian) information. Unlike gradient information that is available directly from PMP solvers, the Hessian values is recovered by solving a non-autonomous Riccati equation depending on the optimal trajectory. Hence, it is crucial to understand the dimension/nonlinearity regime where the cost of obtaining second-order information is justified in term of data efficiency.

References

[1] W. Czarnecki, S. Osindero, M. Jaderberg, G. Swirszcz, and R. Pascanu, Sobolev training for neural networks *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, 4281–4290, 2017.

[2] B. Azmi, D. Kalise, and K. Kunisch, Optimal Feedback Law Recovery by Gradient-Augmented Sparse Polynomial Regression, *Journal of Machine Learning Research*, 22(48):1–32, 2021.

Optimal Control and Stabilisation of Fokker-Planck and McKean-Vlasov Equations

Dante Kalise[1] Lucas M. Moschen[1] Grigoris A. Pavliotis[1]

[1]Department of Mathematics, Imperial College London, UK, d.kalise@imperial.ac.uk, lucas.moschen22@imperial.ac.uk, g.pavliotis@imperial.ac.uk

Keywords: PDE-constrained optimisation, Fokker-Planck equations, McKean-Vlasov equations, numerical methods, stabilisation

We investigate optimal control strategies for probability distributions governed by linear and nonlinear Fokker-Planck (FP) equations, with a particular focus on McKean-Vlasov (MV) dynamics. These equations describe the time evolution of interacting particle systems under a stochastic external force and play a central role in statistical physics and collective behaviour. By controlling these dynamics, we aim to accelerate convergence to equilibrium and stabilise metastable or unstable steady states, i.e., to control the long-time behaviour of FP equations.

We develop two approaches inspired by [1]. Both modify the confining potential via space- and time-dependent controls. First, for the linear FP equations [2], we consider the case in which there is a unique equilibrium and introduce a spectral optimal control framework based on the ground-state transformation to a Schrödinger operator. A gradient-based iterative scheme that combines Pontryagin's maximum principle with Barzilai-Borwein updates computes controls that target slow modes, thus accelerating the convergence to the unique steady state. Second, for MV (nonlinear FP) [3], we treat equations that may admit multiple equilibria, for which our goal is to accelerate convergence to a chosen steady state. We design feedback controls derived from a linear-quadratic problem built on the linearisation around this steady state and involving a Riccati operator. We rigorously prove that the proposed approach stabilises the desired equilibrium at a chosen rate $\delta > 0$.

Numerical experiments on Kuramoto and $O(2)$ models (with magnetic field), ill-conditioned Gaussian, and double-well potentials in one and two dimensions demonstrate that our approaches accelerate convergence to equilibrium and stabilise otherwise unstable states. These results fall within PDE-constrained optimisation and stabilisation, connecting continuous optimisation, spectral methods, and controlled sampling.

References

[1] T, Breiten, K. Kunisch, L. Pfeiffer, Control strategies for the Fokker-Planck equation, *ESAIM: COCV*, 24(2): 741-763, 2025.

[2] D. Kalise, L. M. Moschen, G. A. Pavliotis and U. Vaes, A Spectral Approach to Optimal Control of the Fokker-Planck Equation, *IEEE Control Systems Letters*, 9: 504-509, 2025.

[3] D. Kalise, L. M. Moschen, G. A. Pavliotis, Linearization-Based Feedback Stabilization of McKean-Vlasov PDEs, *arXiv*, 2025.

From Cables to Qubits: A Decomposed Variational Quantum Optimization Pipeline

Paul-Niklas Kandora[1], Adrian Asmund Fessler[1], Robert Fabian Lindermann[1], Phil Arnold[2], Andreas Hempel[2], Steffen Rebennack[1]

[1]Karlsruhe Institute of Technology (KIT), Germany,
{paul-niklas.kandora,steffen.rebennack}@student.kit.edu ,
{adrian.fessler,robert.lindermann}@kit.edu

[2]Vinci Energies, Switzerland, phil.arnold@vinci-energies.com,
andreas.hempel@actemium.de

Keywords: Quadratic unconstrained binary optimization, QUBO, Variational Quantum Eigensolver, VQE, cable routing optimization, quantum computing

Abstract

The Cable Routing Optimization Problem (CROP) is a multi-flow routing task central to industrial layouts and smart manufacturing installations. We formulate CROP as a cable-wise separable, block-diagonal Quadratic Unconstrained Binary Optimization Problem (QUBO) and derive conservative penalty bounds that preserve feasibility. Exploiting this structure, we introduce a decomposition pipeline that builds one QUBO per cable, transform each QUBO into a Hamiltonian and solves the subproblems with Variational Quantum Eigensolver (VQE). Finally the solutions per cable are merged into a global routing assignment. This procedure reduces the per-run qubits from the full problem size to those of a single cable subproblem. We test our performance on different cable routing optimization problems varying in size utilizing Qiskit's SamplingVQE. Our findings indicate that a decomposed VQE approach attains feasible, and optimal, layouts over a range of cable-routing problems.

Prospect Theoretic Models of Irrationality in Aggregative Games

Ashok Krishnan K.S.[1,2] Hélène Le Cadre[3] Ana Bušić[1,2]

[1]Inria, Paris, France.
[2]DI ENS, École Normale Supérieure, PSL Research University Paris, France.
[3]Inria, Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000, Lille, France.
{ashok-krishnan.komalan-sindhu, helene.le-cadre, ana.busic} @inria.fr

Keywords: Game Theory , Prospect Theory, Aggregative Games

Rational behaviour of agents is a classical assumption in game theory. However, as seen from various experiments, rationality is limited when human agents make decisions. Prospect theory models the non rational behaviour of human agents arising from a subjective perception of risk. While Prospect theory has been used in economics, it has not been widely applied in game theory. We apply Prospect theory in games with aggregative structure. Such games arise, for example, in electricity markets where strategic end users are exposed to risk. We show that the irrationality of the agents is seen to modify the structure of the set of Nash equilibria [1]. Further, we obtain conditions under which equilibria are preserved or vanished. Taking a system-wide perspective, we show how the impact of irrational behaviour on the equilibria can be controlled by learning optimal values of a system price parameter [2]. Thus, even in the presence of irrationality, a coordinator can drive system behaviour in a desired direction.

References

[1] Ashok Krishnan K.S., Hélène Le Cadre, and Ana Bušić. "How Irrationality Shapes Nash Equilibria: A Prospect-Theoretic Perspective." accepted for presentation at CDC 2025.

[2] Ashok Krishnan K.S., Hélène Le Cadre, and Ana Bušić. "Achieving a Collective Target Through Incentives." accepted for presentation at NETGCOOP 2025.

Code Evolution Graphs: Understanding Large Language Model Driven Design of Algorithms

Niki van Stein[1] Anna V. Kononova[1] Lars Kotthoff[2,3] Thomas Bäck[1]

[1]LIACS, Leiden University, The Netherlands,
n.van.stein@liacs.leidenuniv.nl, a.kononova@liacs.leidenuniv.nl,
t.h.w.baeck@liacs.leidenuniv.nl
[2]University of St Andrews, UK, lk223@st-andrews.ac.uk
[3]Sorbonne Université, France

Keywords: heuristic optimization, machine learning, code generation

Large Language Models (LLMs) have demonstrated great promise in generating code, especially when used inside an evolutionary computation framework to iteratively optimize the generated algorithms. However, in some cases they fail to generate competitive algorithms or the code optimization stalls, and we are left with no recourse because of a lack of understanding of the generation process and generated codes. We present a novel approach to mitigate this problem by enabling users to analyze the generated codes inside the evolutionary process and how they evolve over repeated prompting of the LLM. We show results for three benchmark problem classes and demonstrate novel insights. In particular, LLMs tend to generate more complex code with repeated prompting, but additional complexity can hurt algorithmic performance in some cases. Different LLMs have different coding "styles" and generated code tends to be dissimilar to other LLMs. These two findings suggest that using different LLMs inside the code evolution frameworks might produce higher performing code than using only one LLM.

An Efficient and Tight Model for Nuclear Power Plant Constraints in Flexible Operation

Francisco Labora[1] Andres Ramos[2]

[1]Instituto de Investigación Tecnológica - Universidad Pontificia Comillas, Spain,
flabora@comillas.edu, aramos@comillas.edu

Keywords: Power system modeling, Renewable energy integration, Unit commitment,Nuclear power plants

Nuclear power plants have historically been inflexible generators, designed for baseload operation, generating a steady output of high power at low variable cost. This conflicts with the flexibility demands that high renewable penetration impose over power systems. However, it is possible, as some countries like France demonstrate daily [1, 2] , to operate nuclear power plants in a more flexible manner to perform load-following operations. Due to the unique characteristics of nuclear reactions as heat sources, reactors are subject to certain constraints, such as Xenon transients, which require specific modeling to represent their behavior [3]. We present a tight formulation for this constraint, which enhances existing literature formulations [4, 5]. We also present novel formulations for modeling the minimum stable time problem with more nuance. We finally provide a case study based on the RTS-GMLC to show the increase in computational efficiency of our improved formulation and the advantages presented by our novel formulations.

References

[1] C. Cany, C. Mansilla, P. da Costa, G. Mathonniere, T. Duquesnoy, A. Baschwitz, Nuclear and intermittent renewables: Two compatible supply options? The case of the French power mix, *Energy Policy*, vol. 95, pp. 135–146, 2016. doi: 10.1016/j.enpol.2016.04.037.

[2] C. Cany, C. Mansilla, G. Mathonniere, P. da Costa, Nuclear power supply: Going against the misconceptions. Evidence of nuclear flexibility from the French experience, *Energy*, vol. 151, pp. 289–296, 2018. doi: 10.1016/j.energy.2018.03.064.

[3] H. Ludwig, T. Salnikova, A. Stockman, U. Wass, Load cycling capabilities of German Nuclear Power Plants (NPP), *VGB PowerTech*, 2011.

[4] J. D. Jenkins, Z. Zhou, R. Ponciroli, R. B. Vilim, F. Ganda, F. de Sisternes, A. Botterud, The benefits of nuclear flexibility in power system operations with renewable energy, *Applied Energy*, vol. 222, pp. 872–884, 2018. doi: 10.1016/j.apenergy.2018.03.002.

[5] R. Ponciroli, Y. Wang, Z. Zhou, A. Botterud, J. Jenkins, R. B. Vilim, F. Ganda, Profitability evaluation of load-following nuclear units with physics-induced operational constraints, *Nuclear Technology*, vol. 200, no. 3, pp. 189–207, 2017. doi: 10.1080/00295450.2017.1388668.

Delayed Feedback in Online Non-Convex Optimization: A Non-Stationary Approach with Applications

Felipe Lara[1] Cristian Vega[1]

[1]Instituto de Alta investigación (IAI), Universidad de Tarapacá, Chile,

felipelaraobreque@gmail.com, cristianvegacereno6@gmail.com.

Keywords: Non-convex online optimization, Delayed algorithms, Quasar-convexity, Bandit, Quadratic fractional programming.

We study non-convex online optimization problems with delay and noise by evaluating dynamic regret in the non-stationary setting when the loss functions are quasar-convex. In particular, we consider scenarios involving quasar-convex functions either with Lipschitz gradients or with weak smoothness, and in each case, we establish bounded dynamic regret in terms of cumulative path variation, achieving sub-linear rates. Furthermore, we illustrate the flexibility of our framework by applying it to both theoretical settings, such as zeroth-order (bandit), and practical applications with quadratic fractional functions. Moreover, we provide new examples of non-convex functions that are quasar-convex by proving that the class of differentiable strongly quasiconvex functions is strongly quasar-convex on convex compact sets. Finally, several numerical experiments validate our theoretical findings, illustrating the effectiveness of our approach.

Robust Mean Field Control: Stochastic Maximum Principle and Risk-averse Mean Field Games

François Delarue[1] Pierre Lavigne[2]

[1]University of Nice Côte d'Azur, France, francois.delarue@unice.fr
[2]University of Nice Côte d'Azur, France, pierre.lavigne@unice.fr.

Keywords: Robust mean field control, Risk-averse mean field games, Stochastic maximum principle, Quadratic backward stochastic differential equation.

This paper introduces and examines a novel class of robust mean field control problems, formulated in a min-max framework where mean field interactions are subject to uncertainty:

$$\sup_q \inf_\psi \left\{ \mathcal{R}(\psi, q) - \mathcal{S}(q) \right\}. \tag{P}$$

where $\mathcal{R}(\psi, q) = \mathcal{G}(q_T, X_T^\psi) + \mathbb{E}[\int_0^T q_s \ell(s, \psi_s) \mathrm{d}s]$ and $\mathcal{S}(q) = \mathbb{E}[\int_0^T q_s f^\star(s, Y_s^\star, Z_s^\star) \mathrm{d}s]$. The first player, called central planner, controls its state process X via the control $\psi \in \mathcal{A}$ and the second player, called the nature, controls an equivalent probability measure q, parametrized by $(Y^\star, Z^\star)$,

$$X_t = \eta + \int_0^t b(s, X_s, \psi_s) \mathrm{d}s + \int_0^t \sigma(s, \psi_s) \mathrm{d}W_s, \quad q_t = 1 + \int_0^t q_s Y_s^\star \mathrm{d}s + \int_0^t q_s Z_s^\star \cdot \mathrm{d}W_s.$$

As a toy example, suppose that $\mathcal{G}(q_T, X_T^\psi) = \mathbb{E}[q_T g(X_T^\psi)]$. The problem solved by the central planner is a standard stochastic control problem and the problem solved by Nature coincides with the risk-aversion problem presented in [1, Chapter 6.4]. In the latter reference, the driver f (dual of $f^\star$) is assume to be of linear growth, while we only impose quadratic quadratic growth assumptions, which is one of the major source of difficulty in the analysis.

The first contribution of this article is the proof of the stochastic maximum principle for the problem (P). Under appropriate concavity-convexity conditions, we show that this problem has a unique solution, where the minimizer is fully characterized by the solution of a FBSDE. This result extends stochastic mean-field control theory to the context of stochastic robust mean-field problems.

The second contribution is to establish existence and uniqueness results for equilibria in risk-averse mean field games. We obtain this result in two different cases, depending on whether the game is variational or not. Up to our knowledge, this work is the first to propose and study risk averse mean field games under variational form. In the latter case, we interpret the mean field game FBSDE system as the first-order conditions of a robust mean field control problem.

Applications to finance, control of systemic risk or particle models are presented.

References

[1] Pham, Huyên, Continuous-time stochastic control and optimization with financial applications, *Springer Science & Business Media*, 2009.

High-Resolution Inertial Dynamics with Time-Rescaled Gradients for Nonsmooth Convex Optimization

Manh Hung LE* Andrea SIMONETTO†

*Unité de Mathématiques Appliquées, ENSTA Paris, manh-hung.le@ensta.fr
†Unité de Mathématiques Appliquées, ENSTA Paris, andrea.simonetto@ensta.fr

Keywords: nonsmooth convex optimization; high-resolution dynamics; acceleration methods; Hessian-driven damping; Moreau envelope.

We study nonsmooth convex minimization through a continuous-time dynamical system that can be seen as a high-resolution ODE of Nesterov Accelerated Gradient (NAG) adapted to the nonsmooth case. We apply a time-varying Moreau envelope smoothing to a proper convex lower semicontinuous objective function f and introduce a controlled time-rescaling of the gradient, coupled with a Hessian-driven damping term, leading to the inertial dynamic:

$$\boxed{\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\,\frac{d}{dt}\big[\delta(t)\,\nabla f_{\gamma(t)}\big(x(t)\big)\big] + \left(1 + \frac{\beta}{t}\right)\delta(t)\,\nabla f_{\gamma(t)}\big(x(t)\big) = 0.}$$

Here the parameters have the following roles:

- $\alpha > 0$: controls the *vanishing viscous damping* term α/t;
- $\beta > 0$: weights the *Hessian-driven damping* contribution;
- $\gamma(t) > 0$: specifies the *time-varying smoothing parameter* in the Moreau envelope $f_{\gamma(t)}$;
- $\gamma(t) > 0$: governs the *time-rescaling of the gradient*.

It should be noted that the factor $1 + \beta/t$ in the last term of the dynamic comes naturally from the high-resolution ODE framework (see, for example, [1]). We provide a well-posedness result for this dynamical system, and construct a Lyapunov energy function capturing the combined effects of inertia, damping, and smoothing. For an appropriate scaling, the energy dissipates and yields decay of the objective function and gradient, stabilization of velocities, and weak convergence of trajectories to minimizers under mild assumptions. Conceptually, the system is a nonsmooth high-resolution model of Nesterov's method that clarifies how time-varying smoothing and Hessian-driven damping jointly govern acceleration and stability. The framework links and extends lines of work on high-resolution ODEs for NAG, Hessian-driven damping, and Moreau-envelope/proximal dynamics.

References

[1] H. Attouch, J. Fadili, *From the Ravine Method to the Nesterov Method and Vice Versa: A Dynamical System Perspective*, SIAM J. Optim. 32(3) (2022), 10.1137/22M1474357

Generalization Meets Nonsmoothness in Wasserstein Distributionally Robust Optimization

Tam Le[1] Jérôme Malick[2]

[1]Université Paris Cité, LPSM, France, ngc.tam.le@gmail.com,
[2]Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, France, jerome.malick@cnrs.fr

Keywords: distributionally robust optimization, generalization, nonsmooth optimization, machine learning, optimal transport

Wasserstein distributionally robust optimization (WDRO) [1] offers an attractive framework for model fitting in machine learning as it systematically accounts for data uncertainty. While the Wasserstein distance is often associated with the curse of dimensionality, recent works have established generalization bounds with dimension-free sample rates for WDRO [2], similar to those in empirical risk minimization. However, exact generalization bounds—which fully capture WDRO's performance—remain unexplored in general settings. In this work [3], we derive an exact generalization bound for WDRO with dimension-free sample rates. Our analysis relies on a careful study of the nonsmooth structure of the robust objective, leading to a unified and broadly applicable framework.

References

[1] Mohajerin Esfahani, P., Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. Mathematical Programming, 171(1), 115-166.

[2] Gao, R. (2023). Finite-sample guarantees for Wasserstein distributionally robust optimization: Breaking the curse of dimensionality. Operations Research, 71(6), 2291-2306.

[3] T. Le and J. Malick, "Universal generalization guarantees for Wasserstein distributionally robust models," in *Proceedings of the Thirteenth International Conference on Learning Representations (ICLR)*, 2025.

Chattering in Homogeneous Linearly Controlled Systems

Enrico Le Donne[1] Sebastiano Nicolussi Golo[1] Nicola Paddeu[1]

[1]University of Fribourg, Switzerland, enrico.ledonne@unifr.ch
sebastiano.nicolussigolo@unifr.ch, nicola.paddeu@unifr.ch.

Keywords: Chattering, Linearly controlled systems, Homogeneous control systems.

The phenomenon of chattering appears in various areas of applied optimal control, as for example quantum control [1], medicine [2], and aerospace [3]. We say that an optimal trajectory *chatters* if it is the unique optimal trajectory joining two points and its control has infinitely many points of discontinuity. We investigate the problem of chattering in homogeneous linearly controlled systems, with control taking value on a polyhedron. For such systems, we give a quantitative bound on the times that an optimal control can switch face of the polyhedron. As a consequence, we show that the control of an optimal trajectory cannot switch infinitely many times between two different faces. Our argument is based on sub-differential calculus and Lie group theory.

References

[1] R. Robin, U. Boscain, M. Sigalotti, and D. Sugny, "Chattering phenomenon in quantum optimal control," *New Journal of Physics*, vol. 24, Paper No. 123037, 15 pp., 2022. doi: 10.1088/1367-2630/acab24.

[2] U. Ledzewicz and H. Schättler, "Singular controls and chattering arcs in optimal control problems arising in biomedicine," *Control and Cybernetics*, vol. 38, no. 4B, pp. 1501–1523, 2009.

[3] E. Trélat, "Optimal Control and Applications to Aerospace: Some Results and Challenges," *Journal of Optimization Theory and Applications*, vol. 154, no. 3, pp. 713–758, 2012. doi: 10.1007/s10957-012-0050-5.

A Characterization of Capra-Convex Sets

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

[1]CERMICS, École nationale des ponts et chaussées, IP Paris, France, `first-name.last-name@enpc.fr`

Keywords: Convexity, Fenchel-Moreau conjugate, one-sided linear coupling

This talk is part of the invited session *Beyond Euclidean Convexity: Methods and Algorithms* (organizers: Pierre-Cyril Aubin and Michel De Lara)

We focus on a specific form of abstract convexity known as Capra-convexity, where a constant along primal rays (Capra) coupling replaces the scalar product used in standard convex analysis to define generalized Fenchel-Moreau conjugacies. A key motivating result is that the ℓ_0 pseudonorm — which counts the number of nonzero components in a vector — is equal to its Capra-biconjugate. This implies that ℓ_0 is a Capra-convex function, highlighting potential applications in statistics and machine learning, par- ticularly for enforcing sparsity in models. Building on prior work characterizing the Capra-subdifferential of ℓ_0 and the role of source norms in defining the Capra-coupling, we present in this talk a characterization of Capra-convex sets.

References

[1] A. Le Franc, J.-P. Chancelier, M. De Lara, S. Rakotomandimby. What are Capra-Convex Sets? Working paper, 2025.

A probabilistic approach for a local day-ahead network congestion signal

Théo Lechat Alexandre Marié Emily Little Lucille Huriez

RTE R&D, Paris La Défense, France

{theo.lechat, alexandre.marie, emily.little, lucille.huriez}@rte-france.com

Keywords: Security analysis, probabilistic load flow, forecast error model, Schaake Shuffle.

In order to accelerate the connection of renewable energies and flexibility resources while limiting costs for the community, RTE–the French transmission system operator–has adopted a strategy of optimal grid sizing, which consists of operating the grid as close as possible to its limits. To this end, RTE aims to better anticipate congestion risks at the local level on D-1.

Short-term network security analysis studies generally involve considering the best forecasts for consumption and renewable energy production (wind and solar) on each bus and simulating the network, with load flow calculations and sensitivity analysis, in the states N ("pre-contingency") and $N-1$ or $N-k$ ("post-contingency"). The result is a deterministic value for the flows on each line, for each state, which can be compared with the Permanent Admissible Transmission Loading (PATL). In this paper, we propose to construct a D-1 congestion signal using a new security analysis approach consisting of calculating a probability distribution of flows for each line and each network state, taking into account historical forecast errors (consumption, wind and solar).

The proposed approach consists in identifying, for each monitored branches, the influential load or renewable generation assets on which we apply an uncertainty model based on historical forecast errors, centered on our best forecast. The resulting probability distributions are then combined, with Probabilistic DC load flow approach using Monte Carlo Simulations [1] and Schaake Shuffle technique [2], to estimate a probability distribution of the flow. The position of the branch PATL in relation to this distribution is then used to construct the congestion signal. This approach was evaluated using historical data from an area in Picardie (France) where 46 MW of wind energy is connected. By comparing our congestion signal calculated based on the D-1 grid forecast with actual flows, we were able to assess the relevance of our approach for this use case.

References

[1] Ali Reza Abbasi and Mohammad Mohammadi. Probabilistic load flow in distribution networks: An updated and comprehensive review with a new classification proposal. *Electric Power Systems Research*, 222:109497, 2023.

[2] Martyn Clark, Subhrendu Gangopadhyay, Lauren Hay, Balaji Rajagopalan, and Robert Wilby. The schaake shuffle: A method for reconstructing space–time variability in forecasted precipitation and temperature fields. *Journal of Hydrometeorology*, 5(1):243–262, 2004.

A Spatial Branch-and-Price-and-Cut Algorithm for the Continuous Network Design Problem

Michael W. Levin[1] David Rey[2]

[1]University of Minnesota, USA, mlevin@umn.edu
[2]SKEMA Business School, France, david.rey@skema.edu

Keywords: Bilevel optimization, transportation, user equilibrium, network design, spatial branch-and-bound

Transportation network design, or the problem of optimizing infrastructure for a societal goal, subject to individual travelers optimizing their behavior for their own preferences arises frequently in many contexts. However, it is also an NP-hard problem due to the leader-follower or bi-level structure involving a follower objective that is different from yet significantly affects the leader objective. Creating exact methods has been particularly difficult for the continuous network design problem (CNDP), in which leader variables are continuous, because of the many nonlinearities arising therein. We present an exact algorithm for the CNDP based on using the high-point relaxation, i.e., the system optimal CNDP, and value function cuts to find lower bounds and solving the traffic assignment follower problem to find upper bounds. We introduce a convex relaxation and a spatial branching scheme to handle the non-convexity of value function cuts. This leads to a spatial branch-and-cut algorithm that gradually cuts out bi-level infeasible points from the feasible region of the CNDP. To enhance computational performance, we outer-approximate nonlinear convex functions and use column generation to obtain a sequence of linear programs that can be solved relatively quickly. We show that, for a predefined ϵ, our spatial branch-and-price-and-cut algorithm converges to ϵ-optimality. Compared to prior work on exact methods for CNDP, we can find ϵ-optimal solutions for the same small test networks in much less time, or solve CNDP on problem instances based on networks that are two orders of magnitude larger than those used in the literature.

A multi-stage stochastic integer programming model and branch-and-SDDP algorithm for industrial production planning under intermittent renewable energy

Ruiwen Liao[1] Franco Quezada[2] Céline Gicquel[1] Safia Kedad-Sidhoum[3]

[1] Université Paris Saclay, LISN, France, celine.gicquel@lisn.fr

[2]University of Santiago of Chile, Faculty of Engineering, Industrial Engineering Dpt, Chile

[3] Conservatoire National des Arts et Métiers, Laboratoire CEDRIC, France

Keywords: Energy-efficient manufacturing, lot-sizing, renewable energy, multi-stage stochastic integer programming, Stochastic Dual Dynamic Programming algorithm

We consider a manufacturing facility producing multiple finished products on a single capacitated machine. The production activities in this facility require energy under the form of electricity. We assume that this electricity is provided by a grid-connected decentralized energy system comprising renewable energy conversion devices (e.g., photovoltaic panels) and an energy storage system (e.g., lithium-ion batteries). We seek to simultaneously plan the industrial production and the energy supply of this facility over a short-term horizon comprising a few days and divided into a discrete set of periods. Specifically, this plan should determine, for each period, the timing of the startup operations to be carried out on the machine and the size of the production lots for the items to be manufactured, as well as the amount of energy to be purchased from the national grid and the quantities to be charged into or discharged from the batteries.

The joint optimization of the short-term production plan and of the energy supply plan is made particularly challenging by the uncertainties on the demand for the finished products and on the renewable power generation. In this work, we consider a multi-stage decision process in which the actual value of the uncertain parameters is revealed progressively over time and some planning decisions can be adjusted after the realization of these parameters. This leads us to develop a multi-stage stochastic integer programming approach for the problem.

We introduce a dynamic-programming type formulation of our multi-stage stochastic integer problem taking the form of a set of nested single-stage sub-problems linked together by cost-to-go functions. The sub-problem corresponding to stage 0 is a mixed-binary linear program while the sub-problems corresponding to the later stages are linear programs.

We propose a novel solution approach based on the hybridization of a branch-and-cut algorithm with a stochastic dual dynamic programming algorithm. Through computational experiments on a diverse set of randomly generated instances, we show that our hybrid algorithm outperforms benchmark methods in both solution quality and speed, particularly for medium to large instances involving a large number of decision stages.

Optimality Conditions and Algorithms for Optimization on Nonlinear Metric Spaces

Russell Luke[1] Titus Pinta[2] Qinyu Yan[3]

[1]Georg-August-Universität Göttingen, Germany, r.luke@math.uni-goettingen.de
[2]École Nationale Supérieure de Techniques Avancées, France, titus.pinta@ensta.fr
[3]Georg-August-Universität Göttingen, Germany, qinyu.yan@math.uni-goettingen.de

Keywords: optimization on nonlinear space, directional derivative, optimality condition, Polyak-Łojasiewicz property, strong convexity

While gradient descent method on Euclidean spaces is deeply studied, similar results on nonlinear metric spaces are still underdeveloped. Due to the lack of linearity, the usual definition of subgradient can not be directly adapted. In this work, directional derivatives along geodesics of functions defined on a p-uniformly convex metric space (G, d) are introduced. The necessary condition for being a local minimizer of f turns out to have nonnegative directional subderivative along all geodesics. If f is strongly convex, then this condition is also sufficient. Based on optimality conditions, algorithms such as basic descent method and steepest descent method are designed and analyzed. Meanwhile, a linear convergence result of the steepest descent method under Polyak-Łojasiewicz property or strong convexity is proved. This work is intended to serve as a foundational result for further research on this topic.

Synthetic Nonnegative Cross-Curvature and Spaces of Measures

Flavien Léger[1] Gabriele Todeschi[2] François-Xavier Vialard[3]

[1]Inria, France, flavien.leger@inria.fr
[2]CERMICS, École nationale des ponts et chaussées, IP Paris, France, gabriele.todeschi@enpc.fr
[3]LIGM, Université Gustave Eiffel, France, francois-xavier.vialard@univ-eiffel.fr

Keywords: Spaces of measures, optimal transport, non-smooth geometry

This talk is part of the invited session *Beyond Euclidean Convexity: Methods and Algorithms* (organizers: Pierre-Cyril Aubin and Michel De Lara)

Nonnegative cross-curvature is a geometric condition related to a pseudo-Riemannian metric defined on a space $X \times Y$ endowed with a cost $c : X \times Y \to \mathbb{R}$. Originally introduced in relation to the regularity theory of optimal transport, it found recently a surprising link with optimization, providing a convenient framework to generalize classical optimization algorithms [1]. In order to work with infinite dimensional spaces and non-smooth costs, we introduced a synthetic formulation for nonnegative cross-curvature and generalized several known results [2]. This condition is similar in spirit (and actually tightly related) to nonnegative curvature bounds *à la Alexandrov*. In particular, our main interest are spaces of (probability) measures. We showed that several distances and divergences satisfies this condition, such as the Hellinger and the Fisher-Rao distances or the Kullback-Leibler divergence. Our main result concerns however optimal transport costs $\mathcal{T}_c$ defined on $\mathcal{P}(X) \times \mathcal{P}(Y)$:

$$\mathcal{T}_c(\mu, \nu) \coloneqq \inf_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} c(x, y) \mathrm{d}\gamma(x, y) \tag{1}$$

where $\Gamma(\mu, \nu)$ is the set of admissible transport plans, i.e. the subset of probability measures on $X \times Y$ with first marginal μ and second marginal ν. Then, the space $(\mathcal{P}(X) \times \mathcal{P}(Y), \mathcal{T}_c)$ has nonnegative cross-curvature if and only if $(X \times Y, c)$ does as well. Notably, this implies that the celebrated Wasserstein distance squared W_2^2 (for $X = Y = \mathbb{R}^d$ and $c(x, y) = |x - y|^2$ in (1)) has non-negative cross-curvature, which opens up interesting perspectives related to optimization in the Wasserstein space.

References

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

[2] F. Léger, G. Todeschi, and F.-X. Vialard, *Nonnegative cross-curvature in infinite dimensions: synthetic definition and spaces of measures*, arXiv preprint arXiv:2409.18112, (2024).

From Nash to Cournot-Nash via Γ-convergence

João Miguel Machado[1] Guilherme Mazanti[2] Laurent Pfeiffer[2]
[1]LMCRC, France, joao-miguel.machado@ceremade.dauphine.fr
[2]Centrale Supelec, France, guilherme.mazanti@inria.fr, laurent.pfeiffer@inria.fr

Keywords: Γ-convergence, mean field limit, potential games, Nash equilibria, Cournot-Nash equilibria, random measures.

In this talk we present results from [2], where we study the convergence of an N-player game to a limit model with a continuum of players as $N \to \infty$. We focus on the convergence of Nash equilibria to Cournot-Nash equilibria. Given a positive integer N and a tuple $(x_i)_{i=1}^N \subset \mathcal{X}$, where x_i is the type of player i, each player minimizes an individual cost depending on their type and a symmetric pairwise interaction of the form:

$$\text{minimize } g_N(y_i; y{-}i) \overset{\text{def.}}{=} c(x_i, y_i) + L(y_i) + \frac{2}{N}\sum_{j\neq i} H(y_j, y_i). \tag{1}$$

In the continuum case, Cournot-Nash equilibria are described by a measure $\gamma \in \mathscr{P}(\mathcal{X} \times \mathcal{Y})$ with $(\pi_\mathcal{X})\sharp\gamma = \mu$, the fixed distribution of players, and satisfying the following fixed point condition

$$\gamma\left(\left\{(x, y) : y \in \text{argmin } c(x, \cdot) + L(\cdot) + 2\int_\mathcal{Y} H(z, \cdot)\, \mathrm{d}\nu(z)\right\}\right) = 1 \tag{2}$$

where $\nu = (\pi_\mathcal{Y})_\sharp\gamma$. The N-player game, is obtained by describing players as an *i.i.d.* sample of μ. We show convergence in two cases: *open loop*, where players choose strategies before the sample is realized, and *closed loop*, where they choose with knowledge of it.

We show that both the N-player games and the continuum game have a potential structure and characterize equilibria by a stationarity condition on the potential function, improving on the minimality condition of Blanchet and Carlier [1]. We then show that the potential functions in both open and closed loop cases Γ-converge to the continuum potential, while the first is stated as a convergence of random measures, the second is a Γ-convergence with full probability. Combining this with the stationarity characterization, we prove that any convergent sequence of Nash equilibria has a Cournot-Nash equilibrium as its limit. Finally, we discuss applications to Lagrangian Mean-Field Games such as [3].

References

[1] A. Blanchet and G. Carlier. Optimal transport and Cournot-Nash equilibria. *Math. Oper. Res.*, 41(1):125â 145, 2016.

[2] J.M. Machado, G. Mazanti, L. Peiffer, From Nash to Cournot-Nash via Γ-convergence, arXiv:2509.02205, 2025

Deep Active Learning-driven Acceleration of Smart Grids Security Assessment

Alban Marie[1] Juan J. Cuenca[2] Emanuel Aldea[3]
Guy Camilleri[4] Anne Blavette[1]

[1]ENS Rennes, Univ. Rennes, CNRS, IETR lab, Rennes, France, firstname.lastname@ens-rennes.fr
[2]CentraleSupélec, CNRS, IETR lab, Rennes, France, juan.cuenca@centralesupelec.fr
[3]SATIE CNRS UMR 8029, Univ. Paris Saclay, Gif-sur-Yvette, France, emanuel.aldea@universite-paris-saclay.fr
[4]IRIT-SMAC, Univ. de Toulouse, CNRS, Toulouse INP, UT3 Toulouse, France, guy.camilleri@irit.fr

Keywords: smart grids, security assessment, power flow, deep active learning, machine learning

The reliable and efficient operation of smart grids is contingent on fast and accurate security assessment. Performing such an assessment is increasingly difficult because of the growing complexity and uncertainty of modern power grids. As traditional power flow simulations based on iterative methods are computationally intensive, machine learning (ML)-based approaches have attracted attention. However, such models typically require large amounts of labelled data, which is time-consuming to acquire via power flow results oracles. To mitigate this, we propose a deep active learning (DAL)-driven framework [1] that actively selects the most informative operational points for labelling, thereby reducing reliance on exhaustive iterative methods-based power flow computations. To the best of the authors' knowledge, deep active learning has never been applied to this type of problem, although recent and rare works emerge on the application of this method to power systems problems [2]. We evaluate multiple DAL query strategies - including Monte Carlo dropout and batch active learning by diverse gradient embeddings (BADGE) - on a binary classification task to detect congestion in a low voltage network (namely the IEEE European low voltage test network (ELVTN)). Results show that DAL methods significantly reduce the number of training labelled samples required over the random baseline on the considered security assessment dataset. Our findings suggest that deep active learning is a promising avenue for accelerating the training of ML-models for smart grid applications, by reducing dependence on costly labelling requiring power flow results from iterative methods [3].

References

[1] P. Ren et. al., "A Survey of Deep Active Learning", ACM Computing Surveys, 2021.

[2] Z. Shi et al., "Intelligent Power System Stability Assessment and Dominant Instability Mode Identification Using Integrated Active Deep Learning", IEEE Transactions on Neural Networks and Learning Systems, 2024.

[3] A. Marie et al., "Deep Active Learning-driven Acceleration of Smart Grids Security Assessment", IEEE ISGT Europe Int. Conf., Malta, Oct. 2024.

Tikhonov Regularized Exterior Penalty Methods For Hierarchical Variational Inequalities

Daniel Cortild[1] Meggie Marschner[2] Mathias Staudigl[2]

[1]University of Oxford, UK

[2]University of Mannheim, Germany, meggie.marschner@uni-mannheim.de

Keywords: Hierarchical Variational Inequalities, Bilevel Optimization, Tikhonov Regularization, Krasnoselskii-Mann Iteration

We consider hierarchical variational inequalities consisting in a (upper-level) variational inequality whose feasible set is given by the solution set of a (lower-level) variational inequality. Purely hierarchical convex bilevel optimization problems and certain multi-follower games are particular instances of nested variational inequalities. Working within a real Hilbert space setting we combine Tikhonov and proximal regularization terms to develop a double loop algorithm with convergence guarantees towards a solution of the nested VI problem. A Krasnoselskii-Mann (KM) iteration serves as the inner loop scheme, involving a set of user-provided parameters that are updated over time. The involved fixed-point encoding map is user defined, where the Forward-Backward splitting and the Three Operator Splitting scheme are possible instances for the KM map. Defining two Gap functions, one measuring optimality and one measuring feasibility, we are able to derive ergodic convergence rates for the inner- and the upper-level. We present various applications that fit into our framework and some preliminary numerical results.

Session: Recent Advances In Multi-Objective Optimization: Theory Meets Practice

References

[1] M. Marschner, M. Staudigl, Tikhonov Regularized Exterior Penalty Methods For Hierarchical Variational Inequalities, *arXiv preprint arXiv:2508.20872*, 2025.

Computing Usage Values for Prospective Studies in Energy Systems Using Spatial Decomposition

Camila Martinez Parra[1,2], Jean-Marc Janin[1],
Michel De Lara[2], Jean-Philippe Chancelier[2], Pierre Carpentier[3]
[1] Réseau de Transport d'Électricité, France,
[2] CERMICS, École nationale des ponts et chaussées, IP Paris, France
[3] UMA, ENSTA Paris, IP Paris, France
Keywords: Prospective studies, stochastic multistage optimization, spatial decomposition

The increasing penetration of renewable energy requires greater use of storage resources to manage system intermittency. As a result, there is growing interest in evaluating the opportunity cost of stored energy, or usage values, which can be derived by solving a multi-stage stochastic optimization problem. Stochasticity arises from net demand (the aggregation of demand and non-dispatchable generation), the availability of dispatchable generation, and inflows when the storage facilities considered are hydroelectric dams.

We aim to compute these usage values for each market zone of the interconnected European electricity system, in the context of prospective studies currently conducted by RTE using the Antares simulation tool. The energy system is mathematically modelled as a directed graph, where nodes represent market zones and arcs represent interconnection links between these zones. In large energy systems, spatial complexity (thirty nodes in the system, each with at most one aggregated storage unit) compounds temporal complexity (a one-year horizon modelled with two timescales: weekly subproblems with hourly time steps).

This work addresses three main sources of complexity: temporal, spatial, and stochastic. We begin by examining the interaction between temporal and stochastic structures through a detailed study of information structures in a two-timescale setting for a single-node system, supported by numerical experiments. Based on this, we extend the framework to the full multinode case by incorporating a spatio-temporal decomposition scheme. To efficiently compute usage values, we apply Dual Approximate Dynamic Programming (DADP), which enables tractable decomposition across both time and space. This approach yields nodal usage values that depend solely on the local state of each node, independently of the others. To assess the quality of this approximation, we conduct numerical studies on a three-node system and compare the usage values obtained via DADP with those computed using traditional methods such as Stochastic Dynamic Programming (SDP) and Stochastic Dual Dynamic Programming (SDDP). Finally, we conduct numerical studies on a more realistic system composed of thirty nodes (modelling part of Europe) and show that DADP scales well.

References

[1] François Pacaud, Michel De Lara, Jean-Philippe Chancelier, and Pierre Carpentier, Distributed multistage optimization of large-scale microgrids under stochasticity, *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 204-211, 2021.

No-Regret Gaussian Process Optimization of Time-Varying Functions

Eliabelle Mauduit[1] Eloïse Berthier[2] Andrea Simonetto[1]

[1]UMA, ENSTA, France, eliabelle.mauduit@ensta.fr, andrea.simonetto@ensta.fr
[2]U2IS, ENSTA, France, eloise.berthier@ensta.fr

Keywords: Gaussian Processes, Time Series, Time-Varying Optimization

We study sequential time-varying optimization, where the underlying reward function evolves over time and is assumed to lie in a Reproducing Kernel Hilbert Space (RKHS). While classical Gaussian Process Upper Confidence Bound (GP-UCB) algorithms achieve sublinear regret in static settings [1], such guarantees typically fail when the reward function drifts unpredictably [2], creating both theoretical and algorithmic challenges.

To address this, we propose SparQ-GP-UCB, a novel framework that handles time variations via uncertainty injection (UI), which increases the noise variance of past observations to account for drift. To mitigate the resulting information loss, we introduce a sparsification and querying mechanism that strategically updates a limited set of past observations, effectively restoring no-regret guarantees. This leads to a sparse inference problem under heteroscedastic noise, which we solve by extending sparse GP methods [3].

We prove that no-regret is achievable under this framework, provided a logarithmic number of side queries per time step. This offers a concrete solution to the open problem of achieving no-regret in time-varying settings. Experiments on synthetic and real-world datasets demonstrate that SparQ-GP-UCB outperforms existing baselines, achieving lower cumulative regret. Building on this, we introduce W-SparQ-GP-UCB, which employs a windowing strategy to reduce the number of queries per iteration from $\mathcal{O}(\log(T))$ to $o(1)$, and provide a lower bound on the queries required for no-regret.

These results demonstrate the theoretical rigor and practicality of our approach, making time-varying optimization broadly applicable to scenarios where continual adaptation is essential.

References

[1] Srinivas, N., Krause, A., Kakade, S.M., Seeger, M.W.: Information-theoretic regret bounds for Gaussian Process optimization in the bandit setting. IEEE Transactions on Information Theory **58**(5), 3250–3265 (2012)

[2] Bogunovic, I., Scarlett, J., Cevher, V.: Time-varying Gaussian Process bandit optimization. In: International Conference on Artificial Intelligence and Statistics. vol. 51, pp. 314–323 (2016)

[3] Burt, D.R., Rasmussen, C.E., van der Wilk, M.: Convergence of sparse variational inference in Gaussian Processes regression. Journal of Machine Learning Research **21**(131), 1–63 (2020)

An Augmented Lagrangian Algorithm with inexact proximal method and its application to a Mean-Field Control problem

Guilherme Manazanti[1] Thibault Moquet[1] Laurent Pfeiffer[1]

[1] CentraleSupélec, Université Paris-Saclay, France, guilherme.mazanti@inria.fr, thibault.moquet@universite-paris-saclay.fr, laurent.pfeiffer@inria.fr

Keywords: Convex optimization, nonsmooth optimization, mean-field games

In this presentation, we will describe our recent results on the convergence of the Augmented Lagrangian Algorithm with inexact proximal method, as well as its application to the numerical approximation of a solution to a mean-field control problem.

The Augmented Lagrangian is an algorithm for the minimization of a convex cost function which can be written as the sum of two convex functions f and g. Motivated by problems in which the direct minimization of $f + g$ is difficult but that of the sum of f with a regularization of g is simpler, the Augmented Lagrangian algorithm proceeds by iteratively minimizing the latter sum, modifying the regularization of g at each iteration. This regularization is obtained through a dual principle, so that our regularized problem is the dual problem associated to an augmented dual cost which penalizes the distance to the dual parameter $\bar{\mu}$. A more complete definition can be found for instance in [1].

Our contribution consists here in studying the behaviour of this algorithm when we cannot find an exact solution of the regularized problems. More precisely, we use the Frank–Wolfe Algorithm for the resolution of these problems. We show a sublinear convergence speed for this method, in terms of the number of calls to the oracle for the Frank–Wolfe Algorithm.

In a second part, we show that this algorithm can be used to find a numerical solution of a mean-field control problem. Our main idea consists in putting the contribution of the running cost and the coupling constraint in the function f. The oracle then consists in a numerical scheme for the coupled system of PDEs Hamilton–Jacobi–Bellman and Fokker–Plank.

References

[1] Jorge Nocedal and Stephen J. Wright, *Numerical optimization*, Springer, 1999

Best-Response Learning in Budgeted α-Fair Kelly Mechanisms

Cleque Marlain Mboulou-Moutoubi[1] Younes Ben Mazziane[1]
Francesco De Pellegrini[1] Eitan Altman[1,2]

[1]LIA, Avignon university, Avignon, France, cleque-marlain.mboulou-moutoubi@univ-avignon.fr, younes.ben-mazziane@univ-avignon.fr, francesco.de-pellegrini@univ-avignon.fr
[2]INRIA, Sophia Antipolis, France,eitan.altman@sophia.inria.fr *

Keywords: decentralized resource allocation, auctions, game theory, Kelly mechanism, α-fair allocation.

The Kelly mechanism is a proportional allocation auction widely adopted in decentralized resource allocation systems to share an infinitely divisible resource among competing agents. We analyze the sequential game it induces when agents have α-fair utilities and behave strategically. Our main result proves that synchronous best-response updates drive bids to the unique Nash equilibrium at a linear rate for $\alpha \in \{0, 1, 2\}$. Extensive simulations reveal that best-response dynamics reach equilibrium significantly faster than previously proposed no-regret learning algorithms [1].

References

[1] Cleque Marlain Mboulou-Moutoubi, Younes Ben Mazziane, Francesco De Pellegrini and Eitan Altman, Best-response learning in budget-constrained α-Fair Kelly mechanisms, *NET-GCOOP*, 2025, hal-05329798v1.

*This work was supported by the Horizon Europe CoGNETs project (Grant Agreement Number 101135930).

On Constructing Optimal Risk Scores via Mathematical Optimization

Cristina Molero-Río[1] Claudia D'Ambrosio[2]

[1]Universidad de Sevilla, Spain mmolero@us.es
[2]LIX CNRS, École Polytechnique, France, dambrosio@lix.polytechnique.fr

Keywords: Mixed-integer non-linear optimization, outer-approximation, interpretable machine learning

The widespread adoption of Machine Learning (ML) in sensitive domains such as healthcare, criminal justice, and policy-making has emphasized the importance of interpretability for fostering trust [4] and meeting regulatory requirements [1]. Mathematical Optimization has proven to be a powerful approach to designing interpretable ML models, as is the case for Risk Scores [5, 2]. Risk Scores—simple linear classifiers based on logistic regression— stand out for their transparency since users can easily compute predictions with integer additions.

Recent work [3] proposes a convex Mixed-Integer Non-Linear Optimization (MINLO) formulation to handle the construction of risk scores. However, the resulting MINLO problems are computationally demanding, due to their large number of integer variables and constraints, thus posing a major challenge for existing solvers even for small-sized instances. In this talk, we will revisit the work in [3] and propose alternative solution strategies. In particular, we will leverage the convexity of the original problem and explore an outer-approximation method. The practical relevance of our approach is tested in both synthetic and real-world datasets.

References

[1] B. Goodman, S. Flaxman, European union regulations on algorithmic decision-making and a “right to explanation". *AI Magazine*, 38(3):50–57, 2017.

[2] J. Liu, C. Zhong, B. Li, M. Seltzer, C. Rudin, Fasterrisk: fast and accurate interpretable risk scores, *Advances in Neural Information Processing Systems*, 35:17760–17773, 2022.

[3] C. Molero-Río, C. D'Ambrosio, Optimal Risk Scores for Continuous Predictors, *Lecture Notes in Computer Science* 15508 (2025), 148–162, https://link.springer.com/chapter/10.1007/978-3-031-82481-4_11.

[4] C. Rudin, C. Chen, Z. Chen, H. Huang, L. Semenova, C. Zhong, Interpretable machine learning: Fundamental principles and 10 grand challenges. *Statistics Surveys*, 16:1–85, 2022.

[5] B. Ustun, C. Rudin, Learning optimized risk scores, *Journal of Machine Learning Research*, 20(150):1–75, 2019.

Preconditioned Primal-Dual Dynamics in Convex Optimization: Non-Ergodic Convergence Rates

Vassilis Apidopoulos[1] Cesare Molinari[2] Juan Peypouquet[3] Silvia Villa[2]

[1]Archimedes, Athena Research Center (Greece), vassilis.apid@gmail.com
[2]Machine Learning Genoa Center (MaLGa) - Dipartimento di Matematica
Universitá degli Studi di Genova (Italy),
cecio.molinari@gmail.com, silvia.villa@unige.it
[3]University of Groningen (The Netherlands), j.g.peypouquet@rug.nl

Keywords: Convex optimization, Saddle-point problem, Primal-Dual methods.

We introduce and analyze a continuous primal-dual dynamical system in the context of the minimization problem $f(x) + g(Ax)$, where f and g are convex functions and A is a linear operator. In this setting, the trajectories of the Arrow-Hurwicz continuous flow may not converge, accumulating at points that are not solutions. Our proposal is inspired by the primal-dual algorithm [1], where convergence and splitting on the primal-dual variable are ensured by adequately preconditioning the proximal-point algorithm. We consider a family of preconditioners, which are allowed to depend on time and on the operator A, but not on the functions f and g, and analyze asymptotic properties of the corresponding preconditioned flow. Fast convergence rates for the primal-dual gap and optimality of its (weak) limit points are obtained, in the general case, for asymptotically antisymmetric preconditioners, and, in the case of linearly constrained optimization problems, under milder hypotheses. Numerical examples support our theoretical findings, especially in favor of the antisymmetric preconditioners.

References

[1] A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, *Journal of Mathematical Imaging and Vision*, 40 (2011) 120-145.

Optimal Linear Interpolation under Differential Information: application to prediction of perfect flows

Soumyodeep Mukhopadhyay[1,2,3] Didier Rullière[1,2] Rodolphe Le Riche[2,1] Xavier Bay[1,2]
Laurent Genest[3] David Gaudrie[3]

[1] École des Mines de Saint-Étienne, F-42023 Saint-Étienne, France s.mukhopadhyay@emse.fr
[2] UMR CNRS 6158 - LIMOS - Laboratoire informatique modélisation et optimisation des systèmes,
[3] Stellantis, France

Keywords: Physics-informed Machine Learning, Co-Kriging, Constrained Optimization

Approximation of functions satisfying partial differential equations is paramount for simulation of physical fluid flows and other problems in physics. Recent studies related to physics-informed machine learning [1] have proven useful as a data-driven complement to numerical models for Partial Differential Equations (PDEs). However, their efficiency and convergence depend on the availability of vast training datasets. For sparse observations, Gaussian process regression or Kriging [2, 3] has emerged as a powerful interpolation model, offering principled estimates and uncertainty quantification. Several attempts have been made to adapt Gaussian processes conditioned on linear PDEs via artificial or collocation observations [4] and kernel design. These methods suffer from scalability issues in higher dimensions and limited generalizability. The aim of our study is to explore the extension of the Kriging predictor - the best linear unbiased predictor - in the presence of linear PDE information at finite collocation points. We propose two approaches: 1) A collocated co-Kriging [3, §24.2] with primary observations of the physical field and auxiliary differential observations; 2) A constrained Kriging optimization problem strongly satisfying linear PDE constraints at the points of prediction through a Lagrangian formulation. This work highlights a trade-off between the computational efficiency of the Lagrange multipliers approach and the strict interpolation of observations.

References

[1] Karniadakis, George Em, et al. "Physics-Informed Machine Learning." *Nature Reviews Physics*, vol. 3, no. 6, May 2021, pp. 422–40, doi:10.1038/s42254-021-00314-5.

[2] Rasmussen, Carl Edward, and Christopher K. I. Williams. "Gaussian Processes for Machine Learning". *MIT Press*, 2008. Adaptive Computation and Machine Learning.

[3] Wackernagel, Hans. "Multivariate Geostatistics", *Springer Berlin Heidelberg*, 2003. doi:10.1007/978-3-662-05294-5_11.

[4] Thore Graepel. 2003. "Solving noisy linear operator equations by Gaussian processes: application to ordinary and partial differential equations". *In Proceedings of the 20th International Conference on Machine Learning* (ICML'03). AAAI Press, 234–241.

Solving Parametric Linear Matrix Inequalities and Convergence Analysis in Optimization

Simone Naldi[1] Mohab Safey El Din[2] Adrien Taylor[3] Weijia Wang[2]

[1]Université de Limoges, CNRS, XLIM, Limoges, France, simone.naldi@unilim.fr
[2]Sorbonne Université, CNRS, LIP6, Paris, France,
mohab.safey_el_din@sorbonne-universite.fr, weijia.wang@sorbonne-universite.fr
[3]Inria, École normale supérieure, PSL Research University, Paris, France, adrien.taylor@inria.fr

Keywords: Linear matrix inequality; Quantifier elimination; Semidefinite programming; Convergence analysis

We consider linear matrix inequalities (LMIs) $A = A_0 + x_1A_1 + \cdots + x_nA_n \succeq 0$ with the A_i's being $m \times m$ symmetric matrices, with entries in a ring $\mathcal{R}$. When $\mathcal{R} = \mathbb{R}$, the feasibility problem consists in deciding whether the x_i's can be instantiated to obtain a positive semidefinite matrix. When $\mathcal{R} = \mathbb{Q}[y_1, \ldots, y_t]$, the problem asks for a formula on the parameters $y_1, \ldots, y_t$, which describes the values of the parameters for which the specialized LMI is feasible. This problem can be solved using general quantifier elimination algorithms, with a complexity that is exponential in n. In [1], we leverage the LMI structure of the problem to design an algorithm that computes a formula Φ describing a dense subset of the feasible region of parameters, under genericity assumptions. The complexity of this algorithm is exponential in n, m and t but becomes polynomial in n when m and t are fixed. We apply the algorithm to a parametric sum-of-squares problem and to the convergence analyses of certain first-order optimization methods, which are both known to be equivalent to the feasibility of certain parametric LMIs, hence demonstrating its practical interest.

References

[1] S. Naldi, M. Safey El Din, A. Taylor, W. Wang, Solving generic parametric linear matrix inequalities, to appear in *Proceedings of the 2025 International Symposium on Symbolic and Algebraic Computation*, arXiv:2503.01487.

On convergence rates of regularized unbalanced optimal transport: the discrete case

Luca Nenna[1] Paul Pegon[2] Louis Tocquec[3]

[1]Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay & INRIA, ParMA, France
luca.nenna@universite-paris-saclay.fr

[2]Université Paris-Dauphine, CNRS, CEREMADE & INRIA, MoKaplan, France
pegon@ceremade.dauphine.fr

[3]Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay & INRIA, MoKaplan and ParMa, France
louis.tocquec@universite-paris-saclay.fr

Keywords. optimal transport, unbalanced optimal transport, entropic regularization, convex analysis, Csiszàr divergence, asymptotic analysis.

Abstract

Unbalanced optimal transport (UOT) is a natural extension of optimal transport (OT) allowing comparison between measures of different masses. It arises naturally in machine learning by offering a robustness against outliers. The aim of this work is to provide convergence rates of the regularized transport plans and potentials towards their original solution when both measures are weighted sums of Dirac masses.

References

[1] Lenaic Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Unbalanced optimal transport: Dynamic and kantorovich formulations. *Journal of Functional Analysis*, 274(11):3090–3123, 2018.

[2] Roberto Cominetti and J San Martín. Asymptotic analysis of the exponential penalty trajectory in linear programming. *Mathematical Programming*, 67:169–187, 1994.

[3] Alex Delalande. Nearly tight convergence bounds for semi-discrete entropic optimal transport. In *International Conference on Artificial Intelligence and Statistics*, pages 1619–1642. PMLR, 2022.

[4] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal entropy-transport problems and a new hellinger–kantorovich distance between positive measures. *Inventiones mathematicae*, 211(3):969–1117, 2018.

[5] Thibault Séjourné, Gabriel Peyré, and François-Xavier Vialard. Unbalanced optimal transport, from theory to numerics. *Handbook of Numerical Analysis*, 24:407–471, 2023.

[6] Jonathan Weed. An explicit analysis of the entropic penalty in linear programming. In *Conference On Learning Theory*, pages 1841–1855. PMLR, 2018.

A Sequential Bayesian Approach to Gaussian Process Quantile Regression for Optimization

Hugo Nicolas[1] Olivier Le Maître[2]

[1]Inria, École polytechnique, IP Paris, France, hugo.nicolas@polytechnique.edu
[2]CNRS, École polytechnique, IP Paris, France, olivier.le-maitre@polytechnique.edu

Keywords: Quantile regression, Gaussian processes, Bayesian optimization, Uncertainty Quantification

Quantile regression extends the classical least-squares regression to the estimation of conditional quantiles of a response variable. In the frequentist approach, the quantile regression problem is cast as the minimization of a loss function, possibly complemented with regularization terms. The Bayesian counterpart formulates the problem as posterior inference over a function space. Of particular interest is Gaussian process quantile regression, which formulates the regression problem as posterior inference over the latent conditional quantiles, where prior knowledge is encoded in the form of a Gaussian process.

Existing approaches to Gaussian process quantile regression either perform the regression directly on observed data [1] or resort to sparse approximations to mitigate computational costs [2]. However, in the latter case, the approximation is typically defined over a small set of latent auxiliary variables that act as compact representations of the quantile, but whose locations are fixed in advance, ultimately resulting in suboptimal predictive performance. In this work, we introduce an adaptive strategy that exploits the Gaussian process predictive variance to infill the set of auxiliary variable locations. Inference of the posterior distribution over these auxiliary variables is recast as its Laplace approximation. The impact of finite training data on the auxiliary variables is estimated through bootstrap resampling. Building on this, we introduce an active learning strategy that acquires new observations of the response variable via rejection sampling, with the sampling density guided by the uncertainties in the auxiliary estimates. Our algorithm combines adaptive auxiliary variable allocation and active learning, leading to a sequential approach that ensures a rich and well-balanced representation of the quantile function.

Finally, we extend our quantile regression method and its enrichment criteria to the quantile minimization problem within a Bayesian optimization framework.

References

[1] L. Torossian, V. Picheny, R. Faivre, A. Garivier, A review on quantile regression for stochastic computer experiments. *Reliability Engineering & System Safety*, **201**, 106858, 2020.

[2] V. Picheny, H. Moss, L. Torossian, N. Durrande, Bayesian quantile and expectile optimisation. *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, 1623–1633, 2022.

Sampled-Based Guided Quantum Walk: Non-variational quantum algorithm for combinatorial optimization

Quantum computing and combinatorial optimization

Ugo Nzongani[1,2] Dylan Laplace Mermoud[2,3] Giuseppe Di Molfetta[1] Andrea Simonetto[2,3]

[1]Aix-Marseille Université, CNRS, LIS, France ugo.nzongani@lis-lab.fr

[2]Unité de Mathématiques Appliquées, ENSTA, Institut Polytechnique de Paris, France

[3]CEDRIC, Conservatoire National des Arts et Métiers, 75003 Paris, France

Keywords: Quantum walks, Combinatorial optimization, Quantum algorithms

We introduce SamBa–GQW, a novel quantum algorithm for solving binary combinatorial optimization problems of arbitrary degree with no use of any classical optimizer. The algorithm is based on a continuous-time quantum walk on the solution space represented as a graph. The walker explores the solution space to find its way to vertices that minimize the cost function of the optimization problem. The key novelty of our algorithm is an offline classical sampling protocol that gives information about the spectrum of the problem Hamiltonian. Then, the extracted information is used to guide the walker to high quality solutions via a quantum walk with a time-dependent hopping rate. We investigate the performance of SamBa–GQW on several quadratic problems, namely MaxCut, maximum independent set, portfolio optimization, and higher-order polynomial problems such as LABS, MAX-k-SAT and a quartic reformulation of the travelling salesperson problem. We empirically demonstrate that SamBa–GQW finds high quality approximate solutions on problems up to a size of $n = 20$ qubits by only sampling n^2 states among 2^n possible decisions. SamBa–GQW compares very well also to other guided quantum walks [1] and QAOA [2].

References

[1] Sebastian Schulz, Dennis Willsch, and Kris- tel Michielsen. Guided quantum walk. *Phys. Rev. Res.*, 6:013312, Mar 2024.

[2] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate opti- mization algorithm, 2014.

Lot-sizing Problem Under Decision-Dependent Uncertainty: A Probing-Enhanced Stochastic Programming Approach

Bernardo K. Pagnoncelli[1] Franco Quezada[2] Céline Gicquel[3]
Safia Kedad-Sidhoum[4]

[1]SKEMA Business School, Lille, France, bernardo.pagnoncelli@skema.edu
[2]University of Santiago of Chile, Faculty of Engineering, Industrial Engineering Department, Chile
franco.quezada@usach.cl,
[3]Université Paris-Saclay, CNRS, LRI, France, gicquel@lri.fr
[4]Conservatoire National des Arts et Métiers, Paris, France, safia.kedad_sidhoum@cnam.fr

Keywords: Decision-dependent uncertainty, Probing, Stochastic programming, Lot-sizing, Branch-and-cut, Value of information

We study the Multi-Item Capacitated Lot-Sizing Problem (MCLSP) under decision-dependent uncertainty, where managers can proactively acquire information before committing to production decisions. Motivated by settings in which demand is correlated with observable covariates, such as competitor actions or macroeconomic indicators, we adopt a probing-enhanced stochastic programming framework. In this setting, decision-makers strategically choose which covariates to probe, updating conditional demand distributions and enabling more informed production planning.

Our contributions are threefold. First, we present the first formulation of the MCLSP with decision-dependent uncertainty, embedding probing actions into a three-stage stochastic program. Second, we propose a compact Big-M-free reformulation that eliminates non-anticipativity constraints, yielding stronger relaxations and improved stability. Third, we extend classical (k, U) inequalities to this setting and introduce a new family of value-function-based inequalities that link probing with expected recourse costs. These cuts strengthen the formulation and accelerate convergence.

To solve the model, we design a branch-and-cut algorithm that incorporates the proposed inequalities and a tailored primal heuristic. This method exploits the decomposable structure of the problem while remaining tractable on large instances. Computational experiments show that our formulation with value-function inequalities consistently outperforms classical models, reducing optimality gaps by up to 85%. The full branch-and-cut algorithm achieves near-optimal solutions, with average gaps below 1.5% even under tight capacity constraints and large scenario sets. A value-of-information analysis further highlights the benefits of probing, showing that limited investments in information can substantially improve cost efficiency and service levels.

Overall, our findings demonstrate the importance of integrating endogenous information acquisition into stochastic lot-sizing. By combining structured reformulation, valid inequalities, and exact algorithms, we provide scalable tools for production planning under uncertainty, while illustrating the broader role of proactive information management in stochastic optimization.

Continuity of Bellman functions in stochastic optimization and control

Teemu Pennanen[1] Ari-Pekka Perkkiö[2]

[1]King's College London, UK, teemu.pennanen@kcl.ac.uk
[2]LMU Munich, Germany, perkkioe@math.lmu.de,

Keywords: Stochastic optimization, Stochastic control, Dynamic programming

We study continuity properties of solutions to Bellman equations (Bellman functions) in discrete-time stochastic optimization and control. Continuity of Bellman functions is important e.g. in the sensitivity analysis of the Bellman equations with respect to changes of the underlying probability measure. Continuity is instrumental also for the convergence of various numerical schemes such as traditional state space discretizations as well as cutting plance approximations that employ convexity to avoid discretizations which tend to be very sensitive to the dimension of the state space. Sufficient conditions are given for lower semicontinuity, continuity and Lipschitz continuity, respectively, of the Bellman functions with respect to their effective domains. Requiring continuity only with respect to the effective domains allows for extended real-valued Bellman functions and more general problem formulations than earlier results on the topic. In particular, our results allow random constraints which are common in most industrial applications of stochastic optimization and control.

Asymptotic behavior of penalty dynamics for constrained variational inequalities

Juan Peypouquet[1] Siqi Qu[2] Mathias Staudigl[2]

[1]Rijksuniversiteit Groningen, Netherlands, j.g.peypouquet@rug.nl

[2]University of Mannheim, Germany, qu.siqi@uni-mannheim.de, m.staudigl@uni-mannheim.de

Keywords: Constrained Variational Inequalities, Dynamical Systems, Penalty Dynamics, Tikhonov regularization

This paper is concerned with the monotone inclusion problem

$$0 \in A(x) + D(x) + \mathsf{NC}_C(x) \tag{1}$$

where $A : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximally monotone operator on a real Hilbert space $D : \mathcal{H} \to \mathcal{H}$ is monotone, and the set $\mathcal{C} \subseteq \mathcal{H}$ is the set of zeroes of a cocoercive operator $B : \mathcal{H} \to \mathcal{H}$. This is a three-operator formulation of a general class of variational problems, where a constrained equilibrium of the sum of two maximally monotone operators is requested over a domain, which admits a representation of the set of zeroes of another single-valued monotone operator . This abstract formulation has many applications in optimal control and optimization, in particular those of a hierarchical nature. In particular, convex simple bilevel problems belong to the above setting, a challenging class of optimization which has received a lot of attention recently. This talk is based on two parts. In the first, we associate with problem (1) a continuous-time dynamical system, designed for solving auxiliary problems governed by the inclusion

$$0 \in A(x) + D(x) + \epsilon x + \beta B(x).$$

Depending on whether is cocoercive or not, we establish strong convergence results of the trajectory to the least norm solution of the original problem (1) by leveraging a technical tracking argument using Lyapunov ideas.

In the second part, we present new ideas how to approach the resolution of inclusions of the form (1) in presence of stochastic data. We present a class of stochastic differential inclusions with semimartingale noise, and study its asymptotic convergence. Discrete-time implications and numerical results close this talk.

References

[1] Staudigl, M., & Qu, S. Tikhonov regularized exterior penalty dynamics for constrained variational inequalities. *IEEE Control Systems Letters*, 8:622-627, 2024.

[2] Qu, S., Staudigl, M., & Peypouquet, J. Asymptotic behavior of penalty dynamics for constrained variational inequalities. (2025). arXiv:2503.03902.

Extended mean field control: a global numerical solution via finite-dimensional approximation

Athena Picarelli[1] Marco Scaratti[1] Jonathan Tam[1,2]

[1]University of Verona, Italy, athena.picarelli@univr.it, marco.scaratti@univr.it
[2]University of Oxford, UK, tam@maths.ox.ac.uk

Keywords: Mean field control, Master Bellman equation, Numerical approximation

We present a finite-dimensional global numerical approximation for a class of extended mean field control problems. Our algorithm learns the value function on the whole Wasserstein domain, as opposed to a fixed initial condition. We leverage the approximation of the mean field problem by a finite-player cooperative optimisation problem, due to the propagation of chaos, together with the usage of finite-dimensional solvers. This avoids the need to directly approximate functions on an infinite-dimensional domain, and allows for more efficient memory usage and faster computation.

References

[1] A. Picarelli, M. Scaratti, J. Tam, Extended mean field control: a global numerical solution via finite-dimensional approximation, arXiv:2503.20510.

A robust method for complementarity problems using a geometric viewpoint

Jean-Pierre Dussault[1] Jean Charles Gilbert[2] Baptiste Plaquevent-Jourdain[2]

[1]Université de Sherbrooke, Québec, Canada, jean-pierre.dussault@USherbrooke.ca
[2]Université de Sherbrooke, Québec, Canada & Inria Paris, France,
jean-charles.gilbert@inria.fr, baptiste.plaquevent-jourdain@USherbrooke.ca

Keywords: complementarity problems, minimum C-function, nonsmoothness, nonconvexity, zonotopes

This talk deals with algorithms to solve complementarity problems (CPs) of the form

$$0 \leq F(x) \perp G(x) \geq 0 \tag{1}$$

where F and G are smooth. The reformulation by C-functions transforms (1) into a system $H(x) = 0$. These approaches benefit from good local convergence properties (one linear system per iteration). Global convergence has been the topic of many contributions, particularly those using the Fischer C-function. Indeed, the associated merit functions are often differentiable, which allows the algorithms to use their gradient if/when needed. Another C-function is the minimum, though its nonsmoothness hinders its globalization. However, its properties (finite termination for LCPs, smaller differential, good performance) motivate its study. [1, 2, 4]

In general, simply solving a linear system may not yield a descent direction of the least-squares merit function when far from a solution. To ensure this, [3] proposes to find a direction in a polyhedron instead of solving a linear system. This method can only converge to points verifying some regularity conditions, as the standard Newton method. To prevent this drawback, we propose a weighted Levenberg-Marquardt approach which reaches a weakly stationary point. Naturally, since the considered problem is nonsmooth nonconvex and without assumptions, such result requires some combinatorial operations, generally of small size − thus they should not be a major difficulty in practice. These are related to the combinatorial geometry of zonotopes. [5]

References

[1] Alexey F. Izmailov, Mikhail V. Solodov, Newton-Type Methods for Optimization and Variational Problems, *Springer Series in Operations Research and Financial Engineering*, 2014.

[2] Francisco Facchinei, Jong-Shi Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, *Springer Series in Operations Research and Financial Engineering*, 2003.

[3] Jean-Pierre Dussault, Mathieu Frappier, Jean Charles Gilbert, Polyhedral Newton-min algorithms for complementarity problems, *Mathematical Programming*, 2025.

[4] Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain, On the B-differential of the componentwise minimum of two affine vector functions, *Mathematical Programming Computation*, 2025.

[5] Baptiste Plaquevent-Jourdain, A Robust Linearization Method for Complementarity Problem – A Detour Through Hyperplane Arrangements, *PhD thesis*, 2025.

Bridging classical optimization and reinforcement learning in power systems: a value-based approach

Rodrigo Porteiro [0000-0001-7793-1645]
Administración Nacional de Usinas y Transmisiones Eléctricas (UTE), Uruguay
rporteiro@ute.com.uy

Keywords: Reinforcement Learning, Hydrothermal Dispatch, Opportunity Cost of Storage

Abstract. Classical hydrothermal dispatch models determine optimal operation policies through Stochastic Dynamic Programming (SDP), where the Bellman value function represents the long-term opportunity cost of storable resources such as water or energy. The marginal value of storage is typically estimated via numerical differentiation of the value function with respect to storage levels, which requires explicit discretization of the state space. As power systems expand, adding reservoirs, batteries, and multi-area interactions, this discretization becomes increasingly burdensome and limits model scalability, particularly under realistic technical constraints and temporal coupling [1].

An alternative formulation based on Reinforcement Learning (RL) is presented, in which the agent is not trained to learn optimal dispatch actions directly. Instead, the agent is trained to estimate the value of storable resources, such as the opportunity cost of water in hydro reservoirs or energy in batteries. These learned value functions are subsequently used as input in deterministic optimization problems solved at each decision stage, ensuring feasibility and consistency with system constraints. Gradient-based policy algorithms, specifically Actor-Critic and Proximal Policy Optimization (PPO), are employed and adapted to focus on state value estimation rather than direct policy learning [2].

This hybrid approach retains the interpretability and robustness of classical optimization while incorporating the scalability and adaptability of modern learning techniques. Numerical experiments in medium-scale power systems confirm the feasibility of this value learning strategy and demonstrate its potential to significantly reduce computational requirements while maintaining high-quality operational outcomes.

This methodology is particularly suitable for systems of intermediate scale, such as the Uruguayan power system, where validation against exact methods remains feasible. As a key contribution, the proposed approach has been implemented and compared to traditional SDP-based methods using a realistic model of the Uruguayan system. The results highlight its capacity to replicate the economic signals of classical models while offering a more scalable and flexible framework. Furthermore, the proposed methodology provides a foundation for extensions in long-term planning, investment analysis that involves storage technologies, and operational strategies in non-deterministic renewable generation patterns [3].

References

[1] M. V. F. Pereira and L. M. V. G. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical Programming*, vol. 52, no. 1–3, pp. 359–375, 1991.

[2] D. Bertsekas, *Reinforcement Learning and Optimal Control*. Athena Scientific, 2019.

[3] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Problems*. Springer, 2013.

Convex Optimization Methods for Nonlinear Control of Synchronous Motors in Transportation Electrification

Matthias Preindl, Department of Electrical Engineering, Columbia University, USA

Synchronous machines, a dominant technology in electric vehicle and industrial drives, exhibit strong nonlinearities due to magnetic saturation and cross-coupling between current components. These effects significantly impact efficiency, torque estimation, and control stability, but their accurate representation leads to optimization problems that are inherently non-convex. This talk presents two convex optimization methodologies that address these challenges, focusing on both modeling and control.

Convex piecewise affine flux maps with energy conservation. Nonlinear current–flux relations $\lambda = \varphi(i)$ obtained from finite element analysis or experiments can be approximated by piecewise affine (PWA) functions of the form $\lambda = L_j i + \psi_j$, $i \in I_j$, where $\{I_j\}$ is a valid simplicial complex. The construction problem is posed as a least-squares fitting subject to semidefinite constraints enforcing $L_j \succ 0$ for all regions, which guarantees monotonicity, invertibility of the map, and conservation of magnetic energy. The corresponding piecewise quadratic energy function is strictly pseudoconvex, ensuring global stability properties and passivity of the state-space model. The resulting convex program is an SDP with quadratic cost and linear matrix inequality (LMI) constraints. Compared to linear flux models, the PWA approach reduces flux estimation error by up to 90% while preserving physical structure and real-time implementability.

Optimal reference generation under nonlinear torque constraints. Efficiency-optimized current setpoints are determined by minimizing copper and iron losses under machine voltage and torque constraints. The problem can be written as

$$\min_{i,\lambda} \; i^\top R i + \omega^2 \lambda^\top G \lambda \quad \text{s.t.} \;\; \tau(i,\lambda) = T_p,\; \|i\| \le i_{\max},\; \|\lambda\| \le \lambda_{\max},\; \lambda = Li + \psi,$$

which has a convex quadratic cost but a non-convex feasible set due to the bilinear torque equation. Nonlinear solvers can only provide local optima, with convergence and runtime strongly dependent on initialization.

Semidefinite relaxation of ORG. Introducing the lifted variable $X = xx^\top$ with $x = [i_r, i_d, i_q, 1]^\top$ yields a semidefinite program where all quadratic constraints become affine in X via the trace operator, e.g., $x^\top A_k x = \operatorname{tr}(A_k X)$. The relaxed formulation,

$$\min_{X \succeq 0} \; \operatorname{tr}(A_0 X) \quad \text{s.t.} \;\; \operatorname{tr}(A_k X) \le b_k,\;\; \operatorname{tr}(A_5 X) = T_p,$$

is convex except for the dropped rank constraint $\operatorname{rank}(X) = 1$. In practice, across the full torque–speed operating region of a commercial WRSM, all feasible solutions were rank-one, ensuring that the relaxation recovers globally optimal setpoints. Solver runtimes are reduced by more than 800% relative to nonlinear programming while providing theoretical optimality guarantees.

Conclusion. Together, these methodologies show how semidefinite programming and convex regression with physical constraints can systematically transform non-convex electric machine control problems into tractable convex formulations. Beyond synchronous machines, the techniques illustrate a general paradigm for applying convex optimization to nonlinear control and energy system problems.

Green Hydrogen Viability in the Transition to a Fully-Renewable Energy Grid

Bárbara Rodrigues[1,2] Daniel Kopisitskiy[3] Denise Cariaga Sandoval[2,4] Miguel F. Anjos[2,5]

[1]INRIA Lille Nord-Europe, Lille, France, barbara.cotrim-rodrigues@inria.fr
[2]University of Edinburgh, Edinburgh, UK,
[3]Université Sorbonne Paris Nord, Paris, France, kopisitskiy@lipn.univ-paris13.fr
[4]Pontifical Catholic University of Chile, Santiago, Chile, dccariaga@uc.cl
[5]GERAD, Montreal, Canada, anjos@stanfordalumni.org

Keywords: Green Hydrogen, Hydrogen Energy Storage, Fully-renewable Grid, Energy Transition, Stochastic Optimization, Economic Analysis

With the transition to a fully renewable energy grid arises the need for a green source of stability and baseload support, which classical renewable generation such as wind and solar cannot offer due to their uncertain and highly-variable generation. In this paper, we study whether green hydrogen can close this gap as a source of supplemental generation and storage. We design a two-stage mixed-integer stochastic optimization model that accounts for uncertainties in renewable generation. Our model considers the investment in renewable plants and hydrogen storage, as well as the operational decisions for running the hydrogen storage systems. For the data considered, we observe that a fully renewable network driven by green hydrogen has a greater potential to succeed when wind generation is high. In fact, the main investment priorities revealed by the model are in wind generation and in liquid hydrogen storage. This long-term storage is more valuable for taking full advantage of hydrogen than shorter-term intraday hydrogen gas storage. In addition, we note that the main driver for the potential and profitability of green hydrogen lies in the electricity demand and prices, as opposed to those for gas. Our model and the investment solutions proposed are robust with respect to changes in the investment costs. All in all, our results show that there is potential for green hydrogen as a source of baseload support in the transition to a fully renewable-powered energy grid.

This work was developed in the context of the 16th AIMMS-MOPTA Optimization Modeling Competition at which it was awarded the First Prize. The full paper can be found at [1].

References

[1] Rodrigues, B., Kopisitskiy, D., Cariaga Sandoval, D. et Anjos, M.F., Green hydrogen viability in the transition to a fully-renewable energy grid. *Optimization and Engineering* (2025). 10.1007/s11081-025-10032-x

A Benders decomposition approach for the EV charging facility location-and-operation problem

Luis Rojo-González[1,2,4] Miguel F. Anjos[1,3,5] Luce Brotcorne[2] Michel Gendreau[1,4]

[1]Département de Mathématiques et de Génie industriel, Polytechnique Montréal
[2]Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL, Lille, F-59000
[3]School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh
[4]Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport
[5]GERAD – HEC Montréal, Montréal

Keywords: Bilevel programming, Linking continuous variables, Benders decomposition, Charging facility location-and-operation problem.

Introduction Electric vehicle (EV) adoption depends on timely charging infrastructure revealing a chicken-and-egg dilemma. A formulation is given by the *Charging Facility Location-and-Operation Problem* as a bilevel program with continuous linking variables. These variables hinder the existence or attainability of the solution (Moore and Bard, 1990) and tailored resolution methods often exclude them (Bolusani and Ralphs, 2022).

Methods This work proposes a two-phase multi-cut Benders decomposition (McDaniel and Devine, 1977) to solve an equivalent MILP of the bilevel program (Kleinert et al., 2021). More precisely, the proposed approach deal with a master problem accounting for primal variables while the subproblem provides bilevel feasibility cuts.

Results Computational experiments show that two-phase variants consistently outperform alternatives with lower medians, upper quartiles, and reduced variability. The effect is most evident on larger networks, where gaps shrink significantly (e.g., from 14.86% to 7.14%).

Conclusion This work solves the charging facility location-and-operation problem using a two-phase Benders decomposition, which consistently outperforms strong duality and relaxation methods. Results show systematically lower optimality gaps, reduced dispersion across instances, efficient master problem control through cut clean-up, and robust performance under varying granularity and route expansion.

References

Bolusani, S. and Ralphs, T. K. (2022). "A framework for generalized Benders decomposition and its application to multilevel optimization". In: *Mathematical Programming* 196.1, pp. 389–426.

Kleinert, T. et al. (2021). "A survey on mixed-integer programming techniques in bilevel optimization". In: *EURO Journal on Computational Optimization* 9, p. 100007.

McDaniel, D. and Devine, M. (1977). "A modified Benders' partitioning algorithm for mixed integer programming". In: *Management Science* 24.3, pp. 312–319.

Moore, J. T. and Bard, J. F. (1990). "The mixed integer linear bilevel programming problem". In: *Operations research* 38.5, pp. 911–921.

Relaxed and Inertial Nonlinear Forward–Backward with Momentum

Fernando Roldán[1] Cristian Vega[2]

[1]Departamento de Ingeniería Matemática, Universidad de Concepción, Chile, fernandoroldan@udec.cl
[2]Instituto de Alta investigación (IAI), Universidad de Tarapacá, Chile, cristianvegacereno6@gmail.com

Keywords: Operator splitting, Monotone operators, Monotone inclusion, Inertial methods, Convex optimization.

In this article, we study inertial algorithms for numerically solving monotone inclusions involving the sum of a maximally monotone and a cocoercive operator. In particular, we analyze the convergence of inertial and relaxed versions of the nonlinear forward–backward with momentum (NFBM). We propose an inertial version of NFBM including a relaxation step and a second version considering a doubleinertial step with additional momentum. By applying NFBM to specific monotone inclusions, we derive inertial and relaxed versions of algorithms such as forward–backward, forward-half-reflected-backward (FHRB), Chambolle–Pock, Condat–Vũ, among others, thereby recovering and extending previous results from the literature for solving monotone inclusions involving maximally monotone, cocoercive, monotone and Lipschitz, and linear bounded operators. We also present numerical experiments on image restoration, comparing the proposed inertial and relaxed algorithms. In particular, we compare the inertial and relaxed FHRB with its non-inertial and momentum versions. Additionally, we compare the numerical convergence for larger step-sizes versus relaxation parameters and introduce a restart strategy that incorporates larger step-sizes and inertial steps to further enhance numerical convergence.

References

[1] Roldán, F., Vega, C. Relaxed and Inertial Nonlinear Forward–Backward with Momentum. J Optim Theory Appl 206, 23 (2025). https://doi.org/10.1007/s10957-025-02694-5

Symmetry-breaking techniques in MILP and application to the Unit Commitment Problem

Cécile Rottner[1]

[1]EDF R&D, cecile.rottner@edf.fr

For a given integer linear program (ILP), a symmetry is a variable permutation which transforms any solution to another solution with same cost. Symmetries arising in ILP can impair the solution process, in particular when symmetric solutions lead to an excessively large Branch and Bound (B&B) search tree. Various techniques, so called symmetry-breaking techniques, are available to handle symmetries in ILP. In this presentation, we focus on so-called ßtructured" symmetries, where all column permutations of the solution matrix are symmetries. Many practical problems related to energy management feature such structured symmetry. We first review existing techniques to handle structured symmetries. We also propose new techniques, bases on variable fixing during B&B. We also introduce the concept of "sub-symmetries", which are symmetries arising in subproblems of the considered ILP. This extends the concept of symmetry and enables to capture more symmetries of the considered ILP. Sub-symmetry breaking techniques are also proposed. Numerical experiments on the Unit Commitment Problem show the efficiency of the proposed techniques.

Trilevel Optimization for Demand Response in Electric Vehicle Charging Networks

Rita Safi[1,2] Yezekael Hayel[2] Raphaël Payen[1] Jean-Baptiste Bréal[1]
Tania Jiménez[2] Riadh Zorgati[1]

[1]EDF R&D, Palaiseau, rita.safi@edf.fr, raphael.payen@edf.fr, jean-baptiste.breal@edf.fr, riadh.zorgati@edf.fr

[2]Laboratoire Informatique d'Avignon, yezekael.hayel@univ-avignon.fr,, tania.jimenez@univ-avignon.fr

Keywords: trilevel optimization, electric vehicle, demand response

This work presents a hierarchical decision-making framework for Demand Response (DR) involving three agents: the Electric Vehicle Aggregator (EVA), the Charging Station Operator (CSO), and the Electric Vehicle (EV). The EVA incentivizes the CSOs to participate in DR, i.e., to adjust their consumption during specific periods in response to an incentive payment, while the EVA ensures that the aggregated contribution of all CSOs in DR meets the requirement imposed by the distribution system operator. To adjust their consumption, CSOs design a price menu that encourages EVs to remain longer at the charging station at lower charging prices, thereby creating flexibility for load adjustment.

The problem is formulated as a trilevel optimization framework with a coupling constraint among the CSOs. The lower and intermediate levels are reformulated into a single Mixed-Integer Quadratic Program (MIQP), transforming the trilevel into a single-leader multi-follower game. Finally, solution approaches based on Stackelberg game algorithms are explored to address this problem, as in [1] and [2].

References

[1] Y. Wan, et al., Stackelberg-Nash game approach for price-based demand response in retail electricity trading, *International Journal of Electrical Power & Energy Systems*, 155 (2024): 109577.

[2] D. Han, et al., Stackelberg Game-Based EV Pricing and Routing Framework to Enhance Local Flexibility, *IEEE Transactions on Industry Applications*, (2025).

Optimal Control of the Coagulation-Fragmentation Equation

Enrico Sartor[1]

[1]Paris-Saclay University, France, enrico00.sartor@gmail.com

Keywords: Coagulation-fragmentation equation, Optimal control, Pontryagin maximum principle, Adjoint-state method

We study optimal control for the coagulation–fragmentation equation, which models the evolution of particle populations undergoing aggregation and break-up. If $f(t,x)$ denotes the number density of particles of size x at time t, the equation reads

$$\partial_t f(t,x) = \mathcal{C}f(t,x) + \mathcal{F}f(t,x),$$

with

$$\mathcal{C}f(t,x) = \tfrac{1}{2}\int_0^x K(y,x-y)f(t,y)f(t,x-y)\,dy - f(t,x)\int_0^\infty K(x,y)f(t,y)\,dy,$$

$$\mathcal{F}f(t,x) = -\alpha(x)f(t,x) + \int_x^\infty \alpha(y)b(x,y)f(t,y)\,dy,$$

where K is the coagulation kernel, α the break-up rate, and b the daughter distribution. Motivated by aerosol fines suppression, clot regulation, and polymer size control, we consider a scalar control $u \in L^2\big([0,T];[u_{\min},u_{\max}]\big)$ that multiplicatively scales coagulation, $K \mapsto u(t)K$, i.e.

$$\partial_t f(t,x) = u(t)\,\mathcal{C}f(t,x) + \mathcal{F}f(t,x).$$

We minimize

$$J(u) = \tfrac{w}{2}\int_0^T (u(t)-1)^2\,dt + \psi\big(f(T)\big), \qquad \psi(f) = \pm\int_{x_{\min}}^{x_{\max}} f(x)\,dx,$$

balancing control effort against a terminal performance criterion (e.g. number of particles of a prescribed size). Working in a weighted L^1 setting, we prove existence of optimal controls via weak-to-weak continuity of the CF dynamics despite the trilinear structure induced by the multiplicative control and bilinear coagulation operator. For first-order conditions, linearization along an optimal pair (f^*,u^*) yields the adjoint

$$-\partial_t\varphi^*(t) = u^*(t)\,D\mathcal{C}[f^*(t)]^*\varphi^*(t) + \mathcal{F}^*\varphi^*(t), \quad \varphi^*(T) = \nabla\psi\big(f^*(T)\big),$$

and the pointwise characterization

$$u^*(t) = P_{[u_{\min},u_{\max}]}\Big(1 - \tfrac{1}{w}\,\langle \mathcal{C}f^*(t), \varphi^*(t)\rangle\Big) \quad \text{a.e. on } [0,T],$$

where $P_{[u_{\min},u_{\max}]}$ denotes projection onto the interval $[u_{\min},u_{\max}]$. Numerically, a finite-volume discretization together with projected gradient and Armijo backtracking illustrates fines suppression and mass concentration in a chosen interval.

QUBO for the Min-Up/Min-Down UCP: Penalty Tuning and Decomposition Approach

Enzo Soler[1] Pascale Bendotti[1] Rodolphe Griset[1*]

[1]EDF R&D, OSIRIS Department, 91120 Palaiseau.

pascale.bendotti@edf.fr

Keywords: UCP, QUBO, Variational algorithms, Decomposition techniques, Column generation

In the context of the energy transition and the growing complexity of power systems, the Unit Commitment Problem (UCP) remains a central challenge in operational planning. It involves determining the optimal schedule for activating power generation units to meet the electricity demand. while minimizing operational costs. Each unit must satisfy minimum uptime (resp. downtime) constraints. The Min-up/min-down UCP (MUCP) is strongly NP hard [3]. Some more technical constraints have also to be taken into account, e.g. ramp constraints. In industrial practice, the UCP is typically addressed using Lagrangian relaxation techniques to decompose the problem into independent subproblems, thus significantly improving computational performance and enabling parallel computing.

This work investigates the potential of quantum variational algorithms to address the MUCP through decomposition techniques, leveraging Quadratic Unconstrained Binary Optimization (QUBO) formulations to solve subproblems efficiently. First we conduct a comparative analysis of penalty coefficient determination methods from the literature [1, 2]. Then we propose a novel hierarchical and normalized penalization strategy tailored to MUCP instances, extending techniques to handle multiple types of constraints [4]. A column generation is considered as an alternative decomposition strategy for solving the MUCP. Integrating quantum computing into this framework raises several interesting questions. One the one hand, quantum solvers may prematurely terminate the column generation process due to suboptimal column selection. On the other hand, their Hamiltonian-based design inherently promotes solution diversity, thus yielding a pool of columns that may still benefit the master problem. A sensitivity analysis of quantum annealing to pricing signals reveals threshold behaviors that characterize the solver's response to variations in dual values. Understanding these behaviors is crucial for designing efficient hybrid decomposition strategies.

References

[1] A. Verma, M. Lewis. Penalty and partitioning techniques to improve performance of QUBO solvers, *Discrete Optimization*, 44:100594, 2022.

[2] M. Ayodele. Penalty Weights in QUBO Formulations: Permutation Problems arXiv:2206.11040, 2022.

[3] P. Bendotti, P. Fouilhoux, C. Rottner. On the complexity of the Unit Commitment Problem, *Annals of Operations Research*, 274(1):119-130, 2019.

[4] M. C. Braun, T. Decker, N. Hegemann, S. F. Kerstan, F. Lorenz. Towards optimization under uncertainty for fundamental models in energy markets using quantum computers, arxiv:2301.01108, 2023.

*Now with Hopia, rgriset@hopia.eu

Real-time pricing for station-based one-way vehicle sharing systems

Rose Sossou Edou, Frédéric Meunier

rose.sossou-edou@enpc.fr, frederic.meunier@enpc.fr
École nationale des ponts et chaussées, France

Keywords: Vehicle sharing systems, pricing, real-time optimization

In this work, we consider a station-based one-way vehicle sharing system, like, e.g., Vélib, whose rebalancing is exclusively performed by the users. This rebalancing is achieved by inciting the users to sometimes modify their destination stations, so as to reroute the flow of vehicles towards stations where demand is high. Dynamic prices attached to the stations are used for the incitation: the amount to pay for a ride is the price attached to the destination station. A user arriving at an origin station may not perform a ride only for two reasons: either there is no available vehicle at the station when the user arrives, or the current prices attached to the stations make the user prefer to walk.

We assume a discrete choice model to describe the users' decisions. Based on this framework, we address the problem of maximizing the proportion of users who eventually take a ride, when prices can be updated according to the state of the system. We propose two heuristics. The first one is a simple rule where the price depends only on the current state of the station. The second one relies on strong duality in the theory of optimal transport [1]. We also propose upper bounds based on mixed integer linear programming, close to those proposed by Waserhole and Jost [2]. The quality of our heuristics is assessed via simulation on synthetic data, with the help of these upper bounds, and with respect to "static" strategies, where the prices cannot change over time. Experiments are currently conducted. In addition, special cases where exact optimal solutions can be determined are also identified.

References

[1] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser, 2015.

[2] A. Waserhole and V. Jost, Pricing in vehicle sharing systems: Optimization in queuing networks with product forms, *EURO Journal on Transportation and Logistics*, vol. 5, pp. 293–320, 2016.

Separable Representations of Optimal Value Functions via Neural Networks

Mario Sperl[1] Luca Saluzzi[2] Dante Kalise[3] Lars Grüne[1]

[1]University of Bayreuth, Germany, mario.sperl@uni-bayreuth.de, lars.gruene@uni-bayreuth.de
[2]Sapienza, University of Rome, Italy, luca.saluzzi@uniroma1.it
[3]Imperial College London, United Kingdom, d.kalise-balza@imperial.ac.uk

Keywords: Separable approximations, decaying sensitivity, neural networks, optimal control

In this talk, we discuss how separable structures provide an effective approach to approximating high-dimensional optimal value functions. The key structural property that enables such approximations is a decaying sensitivity between subsystems, meaning that the influence of one state variable on another diminishes with their graph-based spatial distance. This property makes it possible to construct separable approximations of the optimal value function as a sum of localized contributions. We further demonstrate that these separable approximations admit efficient neural network representations, where the number of parameters grows only polynomially with the state space dimension. These results highlight how structural properties of the problem can be leveraged to obtain scalable neural network representations, thereby mitigating the curse of dimensionality in optimal control. This talk is based on the works [1] and [2].

References

[1] M. Sperl, L. Saluzzi, D. Kalise, and L. Grüne, Separable approximations of optimal value functions and their representation by neural networks, arXiv:2502.08559, 2025.

[2] M. Sperl, L. Saluzzi, L. Grüne, and D. Kalise, Separable approximations of optimal value functions under a decaying sensitivity assumption, 2023 62nd IEEE Conference on Decision and Control (CDC), Singapore, pp. 259-264, 2023.

Model-Based Reinforcement Learning for Exact Combinatorial Optimization

Paul Strang[1,2] Zacharie Ales[2] Come Bissuel[1] Olivier Juan[1]
Safia Kedad-Sidhoum[3] Emmanuel Rachelson [4]

[1]EDF, France, {paul.strang, come.bissuel, olivier.juan}@edf.fr
[2]ENSTA IP Paris, France, zacharie.ales@ensta.fr
[3]CNAM, France, safia.kead_sidhoum@cnam.fr
[4]ISAE-SUPAERO, France, emmanuel.rachelson@isae-supaero.fr

Keywords: Mixed-integer linear programming; Model-based reinforcement learning; Branch-and-bound

Modern mixed-integer linear programming (MILP) solvers are built upon the branch-and-bound (B&B) paradigm. Since the 1980s, considerable research and engineering effort has gone into refining these solvers, resulting in highly optimized systems driven by expert-designed heuristics tuned over large benchmarks. However, in operational settings where structurally similar problems are solved repeatedly, adapting solver heuristics to the distribution of encountered MILPs can lead to substantial gains in efficiency, beyond what static, hand-crafted heuristics can offer. Recent research has thus turned to machine learning (ML) to design efficient, data-driven B&B heuristics tailored to specific instance distributions.

The variable selection heuristic, or branching heuristic, plays a particularly critical role in B&B overall computational efficiency, as it governs the selection of variables along which the search space is recursively split. An early breakthrough came from training a neural network to replicate the behaviour of a greedy branching expert at lower computational cost, thereby surpassing the performance of human-expert heuristics. While subsequent works succeeded in learning efficient branching strategies by reinforcement, none have yet matched the performance achieved by imitation learning (IL) approaches. This trend extends beyond MILPs to combinatorial optimization (CO) problems at large, as reinforcement learning (RL) baselines consistently underperfom both handcrafted heuristics and IL methods trained to replicate expert strategies. Yet, if the performance of IL heuristics are capped by that of the experts they learn from, the performance of RL agents are, in theory, only bounded by the maximum score achievable.

RL agents have only managed to surpass human expert heuristics in a restricted subset of combinatorial problems: board games. In fact, AlphaZero reached superhuman performance in go, chess and shogi by leveraging environment simulators to perform model-based planning. In this work, we seek to extend the use of AlphaZero-like agents beyond board games to exact combinatorial optimization. To that end, we introduce Plan-and-Branch-and-Bound (PlanB&B), a model-based reinforcement learning (MBRL) agent that leverages an internal model of the B&B dynamics to learn improved variable selection strategies. Our computational results suggest that the branching dynamics in B&B can be approximated with sufficient fidelity to enable policy improvement through planning over a learned model, opening the door to broader applications of MBRL to mixed-integer linear programming.

Learning Data-Driven Uncertainty Set Partitions for Robust and Adaptive Forecasting with Missing Data

Akylas Stratigakos[1,2] Panagiotis Andrianesis[3]

[1] UCL Energy Institute, University College London, UK, a.stratigakos@ucl.ac.uk
[2] Department of Electrical and Electronic Engineering, Imperial College London, UK
[3]PERSEE Center, Mines Paris - PSL University, Sophia Antipolis, France,
panagiotis.andrianesis@minesparis.psl.eu

Keywords: adaptive robust optimization, missing data, data-driven uncertainty set partitioning, adversarial learning, energy forecasting.

Short-term forecasting models typically assume the availability of input data (features) when they are deployed and in use. However, equipment failures, disruptions, cyberattacks, may lead to missing features when such models are used operationally, which could negatively affect forecast accuracy, and result in suboptimal operational decisions. In this work [1], we use adaptive robust optimization and adversarial machine learning to develop forecasting models that seamlessly handle missing data operationally. We propose linear- and neural network-based forecasting models with parameters that adapt to available features, combining linear adaptation with a novel algorithm for learning data-driven uncertainty set partitions. The proposed adaptive models do not rely on the identification of historical missing data patterns and are suitable for real-time operations under stringent time constraints. Extensive numerical experiments on an application of short-term wind power forecasting considering horizons from 15 minutes to 4 hours ahead illustrate that our proposed adaptive models are on par with imputation when data are missing for very short periods (e.g., when only the latest measurement is missing), whereas they significantly outperform imputation when data are missing for longer periods. We further provide insights by showcasing how linear adaptation and data-driven partitions (even with a few subsets) approach the performance of the optimal, yet impractical, method of retraining for every possible realization of missing data.

References

[1] A. Stratigakos and P. Andrianesis, "Learning Data-Driven Uncertainty Set Partitions for Robust and Adaptive Energy Forecasting with Missing Data," arXiv:2503.20410v2.

This work was supported in part by the Leverhulme Trust International Professorship (LIP-2020-002) and in part by the European Research Council (ERC) Starting Grant CRISP under Agreement 101165616.

Lissajous varieties

Simon Telen[1]

[1]Max Planck Institute for Mathematics in the Sciences, Germany, simon.telen@mis.mpg.de

Keywords: Nonlinear algebra, Kuramoto model, Convex optimization

We study affine algebraic varieties parametrized by sine and cosine functions, generalizing algebraic Lissajous figures in the plane. We show that, up to a combinatorial factor, the degree of these varieties equals the volume of a polytope. We deduce defining equations from rank constraints on a matrix with polynomial entries. We discuss applications of Lissajous varieties in dynamical systems, in particular the Kuramoto model. This leads us to study connections with convex optimization and Lissajous discriminants. Joint work with Francesco Maria Mascarin.

Irregular Examples of Indecomposable Steiner trees with Infinite Number of Branching Points

Yana Teplitskaya
Paris-Saclay University, France, yana.teplitskaya@universite-paris-saclay.fr,

Keywords: Steiner tree; minimal network; one-dimensional shape optimisation

Steiner tree problem is a problem of connecting the given compact set by the shortest way. I will consider the problem in Euclidean space.

By indecomposable Steiner tree we name the solution of the Steiner tree problem without vertices of degree 2. I will talk about examples of indecomposable Steiner trees with infinite number of branching points (vertices of degree 3).

I will provide different examples of Steiner trees with infinite number of branching points: small (connecting countable number of the terminal points as in [3]), middle (connecting an uncountable set of zero Hausdorff dimension, see [2]) and large (connecting totally disconnected set of positive Hausdorff dimension, as in [1]).

Based on joint works with D. Cherkashin, E.Paolini and E. Stepanov.

References

[1] Cherkashin, D., Teplitskaya, Y. A self-similar infinite binary tree is a solution to the Steiner problem. *Fractal and Fractional*, 7(5), 414 , 2023.

[2] Paolini, E., Stepanov, E., Teplitskaya, Y. An example of an infinite Steiner tree connecting an uncountable set. *Advances in Calculus of Variations*, 8(3), 267-290, 2015.

[3] Cherkashin, D., Paolini, E., Teplitskaya, Y.. Steiner trees with infinitely many terminals on the sides of an angle. Journal of Combinatorial Optimization, 49(3), 1-28, 2025.

Modelling Dynamical Constraints and Maintenance within Aggregated Unit Commitment Problem using Mixed Integer Linear Programming

Hugo Thomas[1] Hélène Arvis[2] Olivier Beaude[2]

[1]École des Mines de Paris, France, hugo.thomas@etu.minesparis.psl.eu
[2]EDF Lab Paris-Saclay, France, helene.arvis@edf.fr, olivier.beaude@edf.fr

Keywords: Unit commitment, Aggregation, Dynamic constraints, Mixed integer linear programming, ERAA data

In this presentation, a long-term dispatch problem is formulated as a unit commitment model that explicitly embeds the dynamic constraints of thermal power units – minimum operating power, minimum up/down times, start-up costs and ramp limits – which determine short-term flexibility [1]. To contain the combinatorial growth of binary decisions over a week at hourly resolution and on a European spatial scale, similar units are grouped into thermal aggregates and the problem is expressed with aggregated variables [2, 3]. Assuming homogeneity and the absence of maintenance windows, the aggregated schedules can be disaggregated to unit-level solutions using the integer decomposability property. Our analysis pinpoints the main sources of non-disaggregability in realistic models – ramp interactions, maintenance windows and heterogeneity – and suggests tractable approximation strategies.

A numerical study is proposed, which uses the European Resource Adequacy Assessment (ERAA) dataset (ENTSO-E harmonised projections) [4]. This publicly available dataset provides 10-year (typically 2025-2035 for current editions) projections of demand, installed capacities, renewable capacity factors and interconnection capacities, alongside multiple historical climate scenarios. It makes it possible to assess the future adequacy between the electricity supply and demand, to carry out studies on the flexibility of the European electricity system under meteorological variability, and to size the interconnections between European countries.

We built a Python platform implementing multi-zone aggregated unit commitment over one year. The platform is designed to serve as a compact test bench to evaluate aggregation/disaggregation schemes and could be used to address other research questions.

References

[1] D. Rajan, S. Takriti, Minimum Up/Down Polytopes of the Unit Commitment Problem with Start-Up Costs, *IBM Research Report*, 2005.

[2] B. Knueven, J. Ostrowski, J.-P. Watson, Exploiting Identical Generators in Unit Commitment, *IEEE Transactions on Power Systems*, 33(4):4496–4507, 2018.

[3] N. Langrené, W. Van Ackooij, F. Bréant, Dynamic Constraints for Aggregated Units : Formulation and Application, *IEEE Transactions on Power Systems*, 26(3):1349–1356, 2011.

[4] European Network of Transmission System Operators for Electricity (ENTSO-E), ERAA dataset, 2023. URL: https://www.entsoe.eu/eraa.

Topology Optimization with Local Perimeter Constraints and Non-Linear Filtering for Minimum Length and Overhang Control

Torres, J.[1] Otero, F.[2,3] Ferrer, A.[1,3]

[1]Universidad Politécnica de Cataluña, ESEIAAT, Spain, jatorres@cimne.upc.edu, aferrer@cimne.upc.edu

[2]Universidad Politécnica de Cataluña, FNB, Spain, fotero@cimne.upc.edu

[3]Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE), Spain

Keywords: topology optimization, additive manufacturing, local perimeter, non-linear filter

A major challenge in the transport industry is the design of lightweight structures that contribute to lower fuel consumption and to reduce the environmental impact. Topology optimization provides an effective approach to achieve such weight reduction while also minimizing design time and cost. However, the resulting geometries are often highly complex and typically require additive manufacturing for production.

Therefore, it is crucial to incorporate manufacturing constraints directly into the optimization process, such as enforcing minimum length scale and controlling overhangs, to avoid impractical shapes and volumes. The present work explores the integration of these constraints into topology optimization. Specifically, we propose the use of non-linear filtering techniques to penalize overhangs in the additive manufacturing context, while simultaneously enforcing a minimum length scale. In addition, this filtering approach is combined with a local perimeter measure, enabling the constraints to be applied in a more spatially localized manner.

Numerical results indicate that the proposed method successfully removes small-scale bar-like features and achieves vertically aligned structural elements, thereby enhancing manufacturability.

State-Constrained Optimal Control on Wasserstein Spaces of Riemannian Manifolds

Ernesto Treumun[1] Hasnaa Zidani[2]

[1]ENSTA - Institut Polytechnique de Paris, France, ernesto.treumun@ensta.fr
[2]INSA Rouen Normandie, France, hasnaa.zidani@insa-rouen.fr

Keywords: Optimal Control, Wasserstein Spaces, Riemannian Manifolds, Hamilton-Jacobi, State Constraints

In this talk we study optimal control problems with state constraints in Mayer form over the Wasserstein space $\mathscr{P}_2(M)$ of a not-necessarily compact Riemannian manifold M. We focus on the Hamilton-Jacobi (HJ) equation on $\mathscr{P}_2(M)$ using the viscosity theory, proving a general comparison principle for the HJ equation, and thus establishing the uniqueness of solutions.

Finally, using classical tools from optimal control theory, we prove that the value function of the state-constrained optimal control problem is a solution of a certain, but known, HJ equation. An additional result -and a cornerstone of this work- is that, under suitable curvature assumptions on the manifold M, we show that the squared Wasserstein distance is directionally differentiable.

References

[1] Luigi Ambrosio. *Gradient Flows: In Metric Spaces and in the Space of Probability Measures.* Lectures in Mathematics. ETH Zürich Ser. Springer Basel AG, Basel, 1st ed edition, 2005.

[2] Averil Aussedat, Othmane Jerhaoui, and Hasnaa Zidani. Viscosity solutions of centralized control problems in measure spaces. *ESAIM: Control, Optimisation and Calculus of Variations*, 30:91, 2024.

[3] Frédéric Jean, Othmane Jerhaoui, and Hasnaa Zidani. Deterministic Optimal Control on Riemannian Manifolds Under Probability Knowledge of the Initial Condition. *SIAM Journal on Mathematical Analysis*, 56(3):3326–3356, June 2024.

Distributionally Robust Optimization for Optimal Control problems with Pontryagin's Maximum Principle

Ange Valli[1] Abdel Lisser[1,2] Sihem Tebbani[1]

[1]Laboratoire des Signaux et Systèmes, Université Paris-Saclay, CNRS, CentraleSupélec, France
[2]Fédération de Mathématiques de CentraleSupélec, Université Paris-Saclay, CNRS, CentraleSupélec, France
{ange.valli, abdel.lisser, sihem.tebbani}@l2s.centalesupelec.fr

Keywords: Pontryagin's Maximum Principle, Distributionally Robust Control, Wasserstein distance

Distributionally Robust Optimization (DRO) is a robust framework based on optimal transport theory for solving optimization problems under uncertainty. Recent research [1] provides tractable formulations based on the Wasserstein distance between probability distributions. For the 1-Wasserstein distance, strong duality results hold under convexity and continuity assumptions. Those approaches have proven their interest in deriving distributionally robust counterparts for various problems, such as regressions or supervised learning [2]. We investigate an approach for solving optimal control problems under uncertainty by including distributionally robust formulations. We derive a system of ordinary differential equations leading to a two-point boundary value problem thanks to Pontryagin's Maximum Principle [3] which determines the optimal control from the initial conditions of the costates of the problem.

References

[1] Kuhn, D., Shafiee, S., & Wiesemann, W. (2025). Distributionally robust optimization. Acta Numerica, 34, 579-804.

[2] Kuhn, D., Esfahani, P. M., Nguyen, V. A., & Shafieezadeh-Abadeh, S. (2019). Wasserstein distributionally robust optimization: Theory and applications in machine learning. In Operations research & management science in the age of analytics (pp. 130-166). Informs.

[3] Bryson, A. E. (2018). Applied optimal control: optimization, estimation and control. Routledge.

Dynamic Algorithm Configuration: Challenges and Opportunities

Diederick Vermetten[1] Carola Doerr[2]

Sorbonne Université, CNRS, LIP6

[1] diederick.vermetten@lip6.fr [2] carola.doerr@lip6.fr

Keywords: Black-box optimization, Algorithm Selection, Chaining

When faced with an optimization problem, we want to utilize the most suitable optimizer for this particular problem. However, in the black-box scenario, this choice can be difficult to make as no or little information can be assumed about the problem at hand. State-of-the-art *algorithm selection* approaches sacrifice a fraction of their evaluation budget to collect information on the problem landscape, which is then used to select the appropriate solver for the given instance [1].

In *per-run algorithm selection,* this cost is reduced by running an initial (default) algorithm for a short evaluation budget, and then using its trajectory to calculate landscape features [2]. This way, the traditional random sampling step of feature-based algorithm selection is replaced by an optimization algorithm.

Extending this per-run principle to a method that performs the selection at any stage of the optimization process, we can create a *fully dynamic algorithm selection (DynAS)* procedure.

In this talk, we will outline the general principles behind DynAS, starting from a data-driven motivation of its potential [4]. From there, we will showcase how efforts to achieve this potential have had mixed results and outline several challenges towards achieving robust DynAS pipelines, among them (1) the choice of appropriate features to characterize the problem landscapes [3] and (2) the design of effective *warm-starting* methods to transfer information from one algorithm to the next.

References

[1] P. Kerschke, H. H. Hoos, F. Neumann, and H. Trautmann. Automated algorithm selection: Survey and perspectives. *Evolutionary computation*, 27(1):3–45, 2019.

[2] A. Kostovska, A. Jankovic, D. Vermetten, J. de Nobel, H. Wang, T. Eftimov, and C. Doerr. Per-run algorithm selection with warm-starting using trajectory-based features. In *PPSN*.

[3] Q. Renau and E. Hart. On the utility of probing trajectories for algorithm-selection. In *EvoApps*, pages 98–114. Springer, 2024.

[4] D. Vermetten, H. Wang, T. Bäck, and C. Doerr. Towards dynamic algorithm selection for numerical black-box optimization: investigating BBOB as a use case. In *GECCO 2020*.

On the Wasserstein Geodesic Principal Component Analysis of probability measures

Nina Vesseron[1] Elsa Cazelles[2] Alice Le Brigant[3] Thierry Klein[4]

[1]CREST-ENSAE, IP Paris, nina.vesseron@ensae.fr
[2]CNRS, IRIT, Université de Toulouse, elsa.cazelles@irit.fr
[3]Université Paris 1 Panthéon Sorbonne, alice.le-brigant@univ-paris1.fr
[4]ENAC, IMT, Université de Toulouse, thierry.klein@math.univ-toulouse.fr

Keywords: Geodesic PCA, Otto-Wasserstein geometry, Bures-Wasserstein geometry

In this talk we discuss how to perform Principal Component Analysis (PCA) of a dataset whose elements are probability distributions, compared using the Wasserstein distance. The goal is to identify geodesic curves in the space of probability measures that best capture the modes of variation of the underlying dataset. To achieve this, we leverage the Riemannian interpretation of the Wasserstein metric. We first address the case of a collection of centered Gaussian distributions, and show how to lift the computations in the space of invertible linear maps using Bures-Wasserstein geometry. For the more general setting of absolutely continuous probability measures, we leverage Otto-Wasserstein geometry and neural networks to parameterize geodesics in Wasserstein space. Finally, we compare to classical tangent PCA through various examples and provide illustrations on real-world datasets. The work described in this talk corresponds to [1].

References

[1] N. Vesseron, E. Cazelles, A. Le Brigant and T. Klein, On the Wasserstein geodesic principal component analysis of probability measures. *arXiv preprint arXiv:2506.04480*, 2025.

Thriving despite toxicity:

in boom-and-bust environments, microbial pacifists often win even against optimal-control-theory-aided antagonists.

MingYi Wang[1] Alexander Vladimirsky[1] Andrea Giometto[1]

[1]Cornell University, USA, mw929@cornell.edu, vladimirsky@cornell.edu, giometto@cornell.edu

Keywords: microbial competition, Piecewise-Deterministic Markov Processes, stochastic optimal control, dynamic programming

In this recent project, our goal was to resolve a fascinating biological puzzle: why toxin-producing microbes almost always win against toxin-sensitive microbes in idealized environments such as a lab, but the situation is often reversed in the real world? Finding an answer required a combination of wet lab experiments, mathematical models, and optimal control theory. It turns out, frequent environmental disruptions such as "dilutions" (e.g., flushing of the gut, or tooth-brushing, or dry soils experiencing heavy rains) give non-antagonistic microbes a good chance to win. This is the case for a fairly broad range of model parameters and remains true whether the dilutions are periodic or randomly timed, and even if antagonistic microbes can use population sensing to produce toxin selectively. We verified the latter by solving non-local Hamilton-Jacobi-Bellman PDEs, encoding the optimal control of toxin production for antagonistic microbes, with dilutions modeled as discontinuous jumps in population trajectories. More broadly, our results also offer some caution about limitations of dynamic models with time-averaged rates.

References

[1] M. Wang, A. Vladimirsky, & A. Giometto, Overcoming toxicity: how nonantagonistic microbes manage to thrive in boom-and-bust environments, *Proc. Natl. Acad. Sci. U.S.A.*, 122 (26) e2424372122, (2025).

A Dynamical Approach to Solving Smooth Saddle Point Problems

Zepeng Wang and Juan Peypouquet

University of Groningen, The Netherlands

zepeng.wang@rug.nl, j.g.peypouquet@rug.nl

Keywords: Primal dual dynamics, saddle point problems, constrained convex optimization

We develop two inertial primal dual dynamical systems to solve smooth saddle point problems with bilinear coupling. The systems are developed based on Nesterov's method, and include both time scaling and Hessian-driven damping. We derive a fast sublinear convergence rate for convex-concave functions, and an accelerated linear convergence rate for strongly convex-strongly concave functions, in terms of the primal-dual gap. We also apply our results to a constrained convex optimization problem with linear equality constraints. These results may provide insights into the design of inertial primal dual algorithms.

MINimum CYcle bases with maximum INtersection in TEmporal GRaphs: ALgorithms and complexITY

Cédric Bentz[1,2] Christophe Picouleau[1,2] Dimitri Watel[1,3]

[1] Laboratoire CEDRIC du CNAM, France
[2] CNAM, Paris, France cedric.bentz@cnam.fr, christophe.picouleau@cnam.fr
[3] ensIIE, Evry, France dimitri.watel@ensiie.fr

Keywords: Temporal graphs, Minimum cycle basis, Matroids intersection problem, Complexity

This work is supported by the FMJH through the PGMO project MINCY-INTEGRALITY.

A common problem in chemo- and bio-informatics consists in studying the positions of atoms constituting a molecule, which form a sequence of molecular graphs derived from these positions at discrete time steps. These graphs all share the same vertex set (which represents the set of atoms), but may vary in some edges, namely the ones associated with hydrogen bonds (which may appear or disappear over time), while others (such as covalent bonds) are persistent. Thus, such graphs may be represented by temporal graphs. As the structure of a molecule is essentially related to the interactions among elementary cycles within its associated graph, such a structure is often characterized by a minimum cycle basis of the graph, which is a concise representation of cycles. So, in order to evaluate the conservation of the molecular structure over time, a natural problem is to look for minimum cycle bases (one for each time step) that have as many cycles in common as possible [1]. Some preliminary results concerning the approximability and parameterized complexity of this problem have already been obtained [2], and the goal of this project is to complement them.

In this presentation, we introduce the problem, the existing results and the questions we aim to answer.

References

[1] Aboulfath, Y., Bougueroua, S., Cimas, A., Barth, D., and Gaigeot, M.-P. Time-resolved graphs of polymorphic cycles for h-bonded network identification in flexible biomolecules. *Journal of Chemical Theory and Computation 20*, 3 (2024), 1019–1035.

[2] Aboulfath, Y., Watel, D., Weisser, M.-A., Mautor, T., and Barth, D. Maximizing minimum cycle bases intersection. In *International Workshop on Combinatorial Algorithms* (2024), Springer, pp. 55–68.

Finite-time Stabilization via Optimal Control Techniques

Weihao Weng[1] Yacine Chitour[1] Paolo Mason[1]

[1]L2S, CNRS, University Paris-Saclay, France,

weihao.weng@centralesupelec.fr, yacine.chitour@centralesupelec.fr, paolo.mason@centralesupelec.fr

Keywords: Optimal control, linear system, finite-time stabilization, Hamilton-Jacobi-Bellman equation

Given a linear control system $\dot{x} = Ax + Bu$, every asymptotically stabilizing feedback controller $u = Kx$ may be obtained by solving an optimal control problem with quadratic integral cost, i.e., the matrix K is such that $A + BK$ is Hurwitz if and only if the feedback $u = Kx$ solves a linear-quadratic problem associated with the linear control system.

On the other hand, (nonlinear) finite-time stabilizing feedback controllers for linear control systems have been obtained in the framework of sliding mode control. The expression of these feedbacks relies on scalar gains which are not explicitly known (see e.g. [1]). The present work is an attempt to determine these gains through an inverse optimal control approach in analogy with the linear-quadratic case. For this purpose, consider the optimal control problem in $\mathbb{R}^n$

$$\text{(OCP)} \qquad \min_{u \in L^{1+q}([0,\infty[,\mathbb{R})} \int_0^\infty \left(\frac{|u(t)|^{1+q}}{1+q} + F(x(t)) \right) dt$$

$$\text{s.t.} \quad \begin{cases} \dot{x}_i(t) = x_{i+1}(t), & i = 1, \dots, n-1 \\ \dot{x}_n(t) = u(t), \\ x(0) = x_0, \end{cases}$$

where $q > 0$ and F satisfies a homogeneity assumption in the sense of [3]. Existence (and uniqueness) of a minimizer u^* for (OCP) leads to the characterization of the associated optimal trajectory x^* using the PMP [2]. We next leverage the homogeneity of the cost function and by extension, of the associated value function given by

$$V_q(x_0) = \int_0^\infty \left(\frac{|u^*(t)|^{1+q}}{1+q} + F(x^*(t)) \right) dt,$$

to establish the finite-time convergence of x^* toward the origin. Finally, after showing regularity properties for V_q we prove that it satisfies a Hamilton-Jacobi-Bellman equation on $\mathbb{R}^n$.

References

[1] Yiguang Hong. Finite-time stabilization and stabilizability of a class of controllable systems. *Systems & control letters*, 46(4):231–236, 2002.

[2] Ernest Bruce Lee and Lawrence Markus. Foundations of optimal control theory. 1967.

[3] Andrey Polyakov. *Generalized homogeneity in systems and control*. Springer, 2020.

Probabilistic Robustness for Compressor Controls in Transient Pipeline Networks

Michael Schuster[1] Elisa Strauch[2] Hendrik Wilka[2]
Jens Lang[2] Martin Gugat[1]

[1]Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany,
michi.schuster@fau.de, martin.gugat@fau.de
[2]Technical University of Darmstadt, Germany,
strauch@mathematik.tu-darmstadt.de, wilka@mathematik.tu-darmstadt.de,
lang@mathematik.tu-darmstadt.de

Keywords: Probabilistic Robustness, Gas Network Control, Probabilistic Constrained Optimization, Kernel Density Estimation, Stochastic Collocation

Uncertainty quantification plays a central role in the modeling and optimization of complex processes such as gas transport simulations. We present a novel strategy to measure the probabilistic robustness of deterministically computed controls in a pipeline network with uncertain gas demand.

Our approach starts by computing optimal controls for a deterministic gas transport problem, where we minimize the total control cost subject to box constraints on the pressure. We then evaluate the probability that the a priori calculated deterministic optimal control satisfies the uncertain pressure bounds in a scenario with randomly perturbed gas demand. To address this, we introduce buffer zones that tighten the pressure bounds in the deterministic scenario. The effect of these stricter bounds is then analyzed in the uncertain setting.

For the computation of the probability, we utilize kernel density estimation[2] on random samples of the uncertain pressure. To reduce the computational effort, we combine this technique with sparse-grid-based stochastic collocation[3] , where only a limited number of gas transport simulations are needed at selected points in the stochastic space.

We present numerical results of our approach applied to a gas network from a public gas library and discuss the effect of buffer zones.

References

[1] M. Schuster, E. Strauch, H. Wilka, J. Lang, M. Gugat, Probabilistic Robustness for Compressor Controls in Transient Pipeline Networks, *Submitted* (2025), Preprint available online

[2] M. Schuster, E. Strauch, M. Gugat, J. Lang, Probabilistic constrained optimization on flow networks, *Optim Eng* 23, 1–50 (2022), https://doi.org/10.1007/s11081-021-09619-x.

[3] J. Lang, P. Domschke, E. Strauch, Adaptive Single- and Multilevel Stochastic Collocation Methods for Uncertain Gas Transport in Large-Scale Networks, in *Mesh Generation and Adaptation*, SEMA SIMAI Springer Series, vol 30. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-92540-6_6.

Optimization of Maintenance Scheduling and Task Clustering in Gas-Insulated Substations

Wenbo Wu[1] Rim Kaddah[1] Miguel F. Anjos[2] Paul-Henri Langlois[3] Sihem Tebbani[4]

[1]Technological Research Institute SystemX, France, {wu.wenbo, rim.kaddah}@irt-systemx.fr
[2]School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Montréal (Québec) H3T 2A7, Canada, anjos@stanfordalumni.org
[3]Réseau de Transport d'Electricité, France, paul-henri.langlois@rte-france.com
[4]Université Paris-Saclay, CentraleSupélec, CNRS, Laboratoire des Signaux et Systèmes (L2S), France, sihem.tebbani@centralesupelec.fr

Keywords: Predictive maintenance, Mixed integer linear programming, Optimization, Gas insulated substations (GIS)

Gas-Insulated Substations (GIS) are compact and highly reliable electrical substations widely deployed in modern power systems, where key high-voltage components are enclosed within sealed metal compartments filled with sulfur hexafluoride (SF6) gas [1]. Despite their operational advantages, GIS maintenance poses significant challenges due to the hazardous environmental impact of SF6 leakage and and the significant cost of maintenance.

An integrated optimization framework for GIS maintenance planning is proposed, which jointly addresses the optimal scheduling of maintenance actions and the temporal grouping of maintenance tasks to enhance operational efficiency. The scheduling problem is formulated as a Mixed Integer Linear Programming (MILP) model that aims to minimize maintenance costs while constraining environmental impacts, ensuring system reliability [2]. Complementing this, a sliding-window based Integer Linear Programming (ILP) model clusters temporally proximate maintenance tasks, consolidating interventions and reducing operational inefficiencies such as frequent site visits and redundant resource deployment [3]. Numerical experiments on various scenarios validate the effectiveness of the proposed framework, demonstrating improved scheduling scalability, enhanced environmental compliance, and operational efficiency gains.

References

[1] CIGRE, CIGRE Position Paper on the Application of SF_6 in *Transmission and Distribution Networks, 2014*

[2] W. Wu, R. Kaddah, M. F. Anjos, P.-H. Langlois, and S. Tebbani, "Environmental impact-aware and cost-efficient optimal predictive maintenance scheduling for Gas Insulated Substations," in *Proceedings of the Joint IFAC Conference: SSSC – TDS – COSY 2025.*

[3] W. Wu, R. Kaddah, M. F. Anjos, P.-H. Langlois, and S. Tebbani, "Task Grouping Optimization for Gas Insulated Substations Predictive Maintenance: Trade-off between cost and environmental impact," in *Proceedings of the 13th IMA International Conference on Modelling in Industrial Maintenance and Reliability*, 2025.

Enhancing the Economic Performance and Resilience of Telecommunications Networks through Integrated Multi-Battery System Solutions

Julien Khamphousone Mustapha Bouhtou Matthieu Chardy Youssouf Hadhbi

Orange Research, Chatillon, France,
youssouf.hadhbi@orange.com

Keywords: Multi-battery systems, telecommunications networks, retail, curtailment, ILP, complexity, greedy heuristic, metaheuristic, reinforcement learning.

Today, telecom networks face increasing economic pressures due to rising energy costs and the demand for higher operational efficiency. Flexible energy storage systems, especially batteries, are essential for balancing supply and demand fluctuations, maintaining network stability during peak loads, and providing reliable backup power during outages or emergencies.

This work offers key insights into battery management strategies that enhance energy efficiency and profitability in multisite telecom networks within retail and curtailment markets. Building on Silva et al. [1, 2], we aim to develop an efficient optimization framework that incorporates additional safety rules and market constraints governing battery usage and energy trading. To address this, we first formulate the multi-battery management as a combinatorial optimization problem that ensures regulatory and safety compliance, while aiming to minimize energy costs and maximize profits from curtailment. We then introduce a novel ILP formulation for the problem and demonstrate its NP-hardness. Using Orange's radio access sites data, results show that the problem becomes more challenging and optimal solutions are out of reach for large-scale instances.

To deal with this, we develop a greedy heuristic that quickly produces good quality solutions. These serve as warm starts for the ILP solver, improving convergence and solution quality. We enhance this approach by incorporating metaheuristics and multiple neighborhood strategies to escape local optima and explore the solution space more effectively. Reinforcement learning is used to adaptively select the most promising moves based on past feedbacks, further boosting search efficiency and solution quality. This hybrid method combines heuristics, metaheuristics, machine learning, and ILP to deliver high-quality solutions suitable for real-world instances.

References

[1] Silva, I. F., Bentz, C. , Bouhtou, M. , Chardy, M., and Kedad-Sidhoum, S. (2020). Battery energy management of a telecommunications company to participate in the curtailment market and reduce the total energy cost. In: 2020 IEEE 8th international conference on smart energy grid engineering (SEGE). IEEE, pp. 121-127.

[2] Silva, I.F., Bentz, C., Bouhtou, M. (2024). Optimization of battery management in telecommunications networks under energy market incentives. Ann Oper Res 332, pp. 949-988.

Gibbs Gradient Descent: A Langevin Approach To Optimization

Oussama Zekri[1] Anna Korba[1] Nicolas Boullé[2]

[1]CREST, ENSAE, France, oussama.zekri@ensae.fr, anna.korba@ensae.fr
[2]Imperial College London, UK, n.boulle@imperial.ac.uk

Keywords: Stochastic sampling, optimization, Gibbs gradient flow

Optimizing over parametric probability distributions underpins sampling tasks in modern generative modeling applications. Gibbs gradient descent (GGD) [1] alleviates costly sampling acquisitions arising from the use of first-order optimization techniques by introducing a single-loop, two-timescale optimization algorithm to optimize through sampling. We provide upper bounds for Gibbs gradient descent in the continuous-time setting with finite and infinite number of particles. Along the way, we introduce an exponential moving average (EMA) algorithm, called EMA-GGD, which naturally extends the method into a three-timescale regime, enabling variance reduction. We illustrate our convergence bounds on several numerical examples.

References

[1] Marion, Pierre, et al. "Implicit diffusion: Efficient optimization through stochastic sampling." AISTATS (2025).

A time-dependent symplectic network for non-convex path planning problems with linear and nonlinear dynamics

Zen Zhang[1] Chenye Wang[2] Shanquin Liu[3] Jérôme Darbon[4] George Em Karniadakis[5]

[1]Division of Applied Mathematics, Brown university, zhen_zhang1@brown.edu, chenye_wang@brown.edu, shanquin_liu@brown.edu, jerome_darbon@brown.edu, george_karniadakis@brown.edu,

Keywords: Deep neural networks, Optimal control, Path planning, Physics Informed Learning

We propose a neural network architecture to address high–dimensional optimal control problems with linear and nonlinear dynamics [1]. An important application of this method is to solve path planning problems of multi-agent vehicles in real time. The new method extends our previous SympOCNet framework by introducing a time-dependent symplectic network into the architecture. In addition, we propose a more general latent representation, which greatly improves model expressivity based on the universal approximation theorem. We demonstrate the efficiency of our approach for path planning problems with obstacle and collision avoidance, including systems with Newtonian dynamics and non-convex environments, up to dimension 512.

References

[1] Zhang, Zhen and Wang, Chenye and Liu, Shanqing and Darbon, Jérôme and Karniadakis, George Em, A Time-Dependent Symplectic Network for Nonconvex Path Planning Problems with Linear and Nonlinear Dynamics, *SIAM Journal on Scientific Computing*, 47(4):C769-C794, 2025.

Author Index