Wednesday, 8:30 - 8:45

■ WA-01

Wednesday, 8:30 - 8:45 - Room: M:A

Opening session

Stream: Opening session
Invited session
Chair: *Pontus Giselsson*

Wednesday, 8:45 - 9:35

■ WB-01

Wednesday, 8:45 - 9:35 - Room: M:A

Plenary I - Gabriel Peyré

Stream: Plenaries
Plenary session
Chair: *Pontus Giselsson*

1 - Conservation laws for gradient flows
Gabriel Peyré

Understanding the geometric properties of gradient descent dynamics is a key ingredient in deciphering the recent success of very large machine learning models. A striking observation is that trained over-parameterized models retain some properties of the optimization initialization. This "implicit bias" is believed to be responsible for some favorable properties of the trained models and could explain their good generalization properties. In this talk I will first rigorously expose the definition and basic properties of "conservation laws", which are maximal sets of independent quantities conserved during gradient flows of a given model (e.g. of a ReLU network with a given architecture) with any training data and any loss. Then I will explain how to find the exact number of these quantities by performing finite-dimensional algebraic manipulations on the Lie algebra generated by the Jacobian of the model. In the specific case of linear and ReLu networks, this procedure recovers the conservation laws known in the literature, and prove that there are no other laws. This is a joint work with Sibylle Marcotte and Rémi Gribonval.

Wednesday, 10:05 - 11:20

■ WC-02

Wednesday, 10:05 - 11:20 - Room: M:O

Conic and Semidefinite Optimization

Stream: Conic optimization: theory, algorithms and applications
Invited session
Chair: *Miguel Anjos*

1 - Learning to Relax Nonconvex Quadratically Constrained Quadratic Programs
Burak Kocuk, Buket Özen

Nonconvex quadratically constrained quadratic programs are NP-Hard to solve in general. Literature primarily employs either semidefinite or linear programming to relax them, which are usually effective for distinct problem types, and a holistic understanding of which relaxation should be preferred over the other for a given instance is lacking. In this research, we present a learning-based approach to predict whether a semidefinite or linear programming relaxation would produce a stronger bound for a given instance by examining spectral properties and sparsity patterns of the data matrices.

2 - Beyond Traditional PCA: The Two-Step-SDP Algorithm for Data Analysis
Eloisa Macedo

Analysing real-world data is fundamental for informed decision-making in various fields, from business and science, to healthcare and government. Due to ever-growing advances in sensing data, e.g., in the context of smart cities, large databases of objects and attributes can be difficult to analyse and extract meaningful information. Many data analysis frameworks for extracting hidden patterns in data have been proposed in the literature. Some methodologies allow to reduce the dimension of both objects and variables. In this context, the Two-Step-SDP methodology is relevant due to its light approach that incorporates a clustering on both objects and attributes and reveals good computational performance when compared to similar methodologies. First, the Two-Step-SDP algorithm focuses on solving two relaxed Semidefinite Programming (SDP) clustering models. These solutions provide starting points (centroids) for the clustering process of both sets of objects and attributes. Then, the algorithm iteratively refines these initial centroids to find optimal clusters, with the main objective of maximizing the between cluster deviance in a lower-dimensional space. The obtained solutions allow to unveil the underlying structure of the data through nonoverlapping clusters, which is important to ease data interpretation. The algorithm is applied to real road traffic-related data and unlocks the true potential of extracting information behind data.

3 - Semidefinite liftings for the complex cut polytope
Miguel Anjos, Lennart Sinjorgo, Renata Sotirov

We consider the complex cut polytope: the convex hull of Hermitian rank-one matrices xx*, where the elements of the n-dimensional vector x are complex m-th unit roots. These polytopes find applications in MAX-3-CUT, digital communication, and more generally, complex quadratic programming. For m = 2, the complex cut polytope corresponds to the well-known real cut polytope. We provide an exact description of the complex cut polytope for m = n = 3 and investigate second order semidefinite liftings of the complex cut polytope. For such second order liftings, we show a method for reducing the size of the matrix, without weakening the approximation. We support our theoretical findings with numerical experiments.

■ WC-03

Wednesday, 10:05 - 11:20 - Room: M:J

Optimization in neural architectures I

Stream: Optimization in neural architectures: convergence and solution characterization
Invited session
Chair: *Manish Krishan Lal*
Chair: *Maria-Luiza Vladarean*

1 - Optimization dynamics of equivariant neural networks
Axel Flinth

A symmetry in a learning task is an inductive bias, that one should take advantage of. This is the main tenet behind the development of so called geometric deep learning, which objective is to develop and analyze methods for capturing such symmetries. There are frameworks available which allow oneself to design network architectures manifestly equivariant to almost any given group symmetry. A different approach is to simply train a standard architecture on symmetric (augmented) data. Despite the natural relations between the two strategies, systematic comparisons are far and few. In this talk, we will present initial some results about the relation of the training dynamics of the two approaches. We will discuss conditions which guarantee that the critical points of both training dynamics coincide. We will see that in these cases, the critical points may still have different stability properties for the two strategies.

2 - Understanding neural architectures via projection onto sets of generalized bilinear constraints
Manish Krishan Lal

The geometry associated with matrix-matrix products can be described with generalized sets of bilinear forms. This paves the way to develop a tensorization approach to the data space in many learning problems. Starting from simple deep neural networks, we exploit this product structure in different architectures such as CNNs, Autoencoders, GANs, and transformers and provide many alternative frameworks to train these networks. The theory lends its support from closed-form projections onto sets of bilinear constraints, hidden convexity, and SDP duality arising in the nonconvex projection problems, and sampling in lifted spaces.

■ WC-04

Wednesday, 10:05 - 11:20 - Room: M:M

Optimization in regression, classification and learning I

Stream: Optimization in regression, classification and learning
Invited session
Chair: *Paula Amaral*

1 - Machine learning outcome prediction model-based radiotherapy treatment plan optimization using the open-source toolkit pyanno4rt

Tim Ortkamp, Martin Frank, Oliver Jäkel, Niklas Wahl

Inverse radiotherapy treatment plan optimization problems can be described as continuous, multi-objective, nonlinear, potentially nonconvex and large-scale (up to hundreds of thousands of variables). Conventionally, these problems are formed by translating clinical criteria into a set of mathematical objectives, which rely on empirical dose prescription and tolerance parameters rather than directly optimizing for the treatment outcome as quantified by normal tissue complication probability (NTCP) and tumor control probability (TCP). Classical outcome models like the Lyman-Kutcher-Burman model are technically feasible for optimization, but questionable with regards to accuracy and usability. Our contribution therefore lies in the development of a multivariate machine learning (ML) model-based optimization framework. To this end, we implemented pyanno4rt, an open-source Python package featuring different optimization methods using zero-, first- and second-order solvers, along with various strategies for model integration, efficient calculation via JIT compilation, and automatic differentiation. We internally fitted three ML models (logistic regression, neural networks, support vector machine) on two head-and-neck (N)TCP datasets, then optimized and evaluated conventional, radiobiological and ML model-based treatment plans for an example patient. Our results show that applying ML models can yield acceptable dose distributions with improved outcome predictions.

2 - Use of Machine Learning techniques in predicting the course of relapsing-remitting MS in individual patients

Raffaele Mariosa, Laura Palagi

We aim to use ML tools to derive a robust and accurate prognostic tool to personalize the treatment of multiple sclerosis (MS). MS is the leading cause of progressive neurological disability in young people, having a high impact on social and economic costs. An early prediction of disease course would allow to differentiate treatment strategies based on the expected severity of the disease. Most ML studies use unconventional data, hindering real-world applicability regardless of performance. Here, we applied ML on data collected by the Italian National MS Registry (IMSR) to predict medium-term disease course. We employed traditional supervised ML algorithms (XG-Boost, Support Vector Machine, Balanced Random Forest Classifiers), ensembles and auto-ML tools to predict the state of an RR patient after time T (180, 360 and 720 days) from the current visit. Moreover, to address the critical need for model interpretability and explainability, we use optimal tree mathematical programming models to extract the main rules for classification. In particular, we use MIRET [1] that offers also a hierarchical visualization tool based on a heatmap representation of the tree ensemble's feature use. Our model of active collaboration between doctors and data analyst has been successful in optimizing the analytical workflow, emphasizing the importance of data quality and suggesting strategies for enhancing both database management and analysis.

3 - White box models in classification

Paula Amaral, Rui Malha, Tiago Dias

Machine learning models have been successfully applied in many situations, often without a critical perspective on these mechanisms. Black box models make it very difficult, if not impossible, to perform counterfactual analysis. Exploring other, more transparent machine learning techniques can be beneficial in situations where scrutinizing the model behind the results is important. In this presentation, we will discuss two perspectives: one based on the analysis of symbolic data and another that introduces a new classification model based on a cloud of spheres.

■ WC-05

Wednesday, 10:05 - 11:20 - Room: M:N

Optimization for learning I

Stream: Optimization for learning
Invited session
Chair: *Manu Upadhyaya*

1 - Incorporating History and Deviations in Forward-Backward Splitting

Pontus Giselsson

We propose a variation of the forward-backward splitting method for solving structured monotone inclusions. Our method integrates past iterates and two deviation vectors into the update equations. These deviation vectors bring flexibility to the algorithm and can be chosen arbitrarily as long as they together satisfy a norm condition. We present special cases where the deviation vectors, selected as predetermined linear combinations of previous iterates, always meet the norm condition. Notably, we introduce an algorithm employing a scalar parameter to interpolate between the conventional forward–backward splitting scheme and an accelerated O(1/n2)-convergent forward-backward method that encompasses both the accelerated proximal point method and the Halpern iteration as special cases. The existing methods correspond to the two extremes of the allowed scalar parameter range. By choosing the interpolation scalar near the midpoint of the permissible range, our algorithm significantly outperforms these previously known methods when addressing a basic monotone inclusion problem stemming from minimax optimization.

2 - Optimal Acceleration for Minimax and Fixed-Point Problems is Not Unique

TaeHo Yoon, Jaeyeon Kim, Jaewook Suh, Ernest Ryu

Recently, accelerated algorithms using the anchoring mechanism for minimax optimization and fixed-point problems have been proposed, and matching complexity lower bounds establish their optimality. In this work, we present the surprising observation that the optimal acceleration mechanism in minimax optimization and fixed-point problems is not unique. Our new algorithms achieve exactly the same worst-case convergence rates as existing anchor-based methods while using materially different acceleration mechanisms. Specifically, these new algorithms are dual to the prior anchor-based accelerated methods in the sense of H-duality. This finding opens a new avenue of research on accelerated algorithms since we now have a family of methods that empirically exhibit varied characteristics while having the same optimal worst-case guarantee.

3 - Accelerated Algorithms For Nonlinear Matrix Decomposition With The Relu Function
Giovanni Seraghiti, Arnaud Vandaele, Margherita Porcelli, Nicolas Gillis

In this contribution I propose a new problem in low-rank matrix factorization, that is the Nonlinear Matrix Decomposition (NMD): given a sparse nonnegative matrix, find a low-rank approximation, that recovers the original matrix by the application of an element-wise nonlinear function. I will focus on the so-called ReLu-NMD, where the nonlinear function is the rectified unit (ReLu) non-linear activation.

At first, I will provide a brief overview of the motivations and possible interpretations of the model, supported by theoretical examples. I will explain the idea that stands behind ReLU-NMD and how nonlinearity can be exploited to get low-rank approximation of given data.

Then, I will stress the connection with neural networks and I will present some of the the existing approaches developed to tackle ReLu-NMD.

Furthermore, I will introduce two new algorithms: (1)Aggressive Accelerated NMD (A-NMD) which uses an adaptive Nesterov extrapolation to accelerate an existing algorithm, and (2)Three-Block NMD (3B-NMD) which parametrizes the low-rank approximation in two factors and leads to a significant reduction in the computational cost.

Finally, I will illustrate the effectiveness of the proposed algorithms on synthetic and real-world data sets, providing some possible applications.

■ WC-06

Wednesday, 10:05 - 11:20 - Room: M:H

Advances in monotone inclusions and related methods

Stream: Methods for non-/monotone inclusions and their applications
Invited session
Chair: *Dimitri Papadimitriou*

1 - Warped proximal iterations for solving nonmonotone inclusions and applications
Dimitri Papadimitriou, Cong Bang Vu

In a real Hilbert space H, the monotone inclusion problem aims at finding a zero of a set-valued maximally monotone operator A. The term warped proximal iteration was recently introduced as generalization of the proximal point algorithm for finding a zero point of a maximally monotone operator A acting on H. Nevertheless, the maximal monotonicity of A restricts its applicability to the class of convex optimization problems as well as operator splitting methods for composite monotone inclusions. The solving of general nonmonotone inclusion, i.e., the inclusion where the operator A is nonmonotone, is an open and challenging research problem. For this purpose, the notion of r-(co)hypomonotonicity has been introduced to guarantee the convergence of the generated sequence. From this perspective, our first objective is to extend the definition of r-hypomotonicity. The second is to investigate the weak convergence property of the warped proximal iteration as well as its various applications to (constrained) nonconvex optimization problems. In particular, we place our attention to the finding of KKT points for a class of nonconvex quadratic programming problems with equality constraints.

2 - Revisiting Inexact Fixed-Point Iterations for Min-Max Problems: Stochasticity and Structured Nonconvexity
Ahmet Alacaoglu, Donghwan Kim, Stephen Wright

In this talk, we revisit the analysis of inexact Halpern and Krasnosel'skii-Mann (KM) iterations for solving constrained and stochastic min-max problems. First, we relax the inexactness requirement on the computation of the resolvent in Halpern iteration. Second, we analyze the stochastic KM iteration with the multilevel Monte-Carlo estimator which allows obtaining almost-unbiased samples of the resolvent. We present the consequences of these results in the case of constrained and stochastic convex-concave min-max problems. Then, we apply our developments to solve constrained nonconvex-nonconcave min-max problems either satisfying (i) rho-cohypomonotonicity or (ii) the assumption of the existence of a solution to the rho weak Minty Variational inequality. Our results help go beyond the existing barrier of $1/(2L)$ for the upper bound of rho parameter (that roughly determines the level of nonconvexity in the problem) in the literature of first-order algorithms for these problems. We extend this upper bound to $1/L$ while preserving the optimal first-order complexity for both deterministic and stochastic versions of the problem.

3 - Set-valued IFS for the Analysis of Iterative Methods in Non-Convex Optimization
Allahkaram Shafiei

In the domain of non-convex optimization, dissecting the behavior and convergence characteristics of iterative methods is essential yet challenging due to the complexity of the objective functions. This study presents a novel strategy employing set-valued iterated function systems (IFS) to examine such methods. Set-valued IFS offers a structured framework to model the dynamic behavior of iterative algorithms, effectively tracking the evolution of solution sets throughout iterations. Let us consider a stochastic dynamical system in the form of a ground set, a family of maps from elements of the ground set to subsets of the ground set, and a probability function that suggests the probability of each map within the family. This generalizes the iterated function systems (IFS) to set-valued maps. In this setting, we present several novel conditions for the existence of a unique invariant measure of convergence of iterated random compositions of such maps. Within this paper, we examine innovative stochastic reformulations of the non-convex feasibility problem aimed at enhancing the advancement of novel algorithmic methodologies. So, by this method, we are inspired to solve non-convex feasibility problems refer to a class of optimization problems where the goal is to find a point (or set of points) that satisfies a given set of constraints, but the feasible region defined by these constraints is non-convex.

■ WC-07

Wednesday, 10:05 - 11:20 - Room: M:I

Optimization applications I

Stream: Optimization applications
Invited session
Chair: *Jacopo Maria Ricci*

1 - Derivative-Free Optimization Applied to Hydraulic Coefficient Estimation
Fabio Fortunato Filho, José Mario Martínez

To comprehend and contribute to decision-making in situations of environmental disasters related to rivers, hydraulic studies play a crucial role. In the case of dam ruptures, for example, it is vital to consider both short-term and long-term challenges. For this study, it is necessary to have access to the maximum amount of available information in order to apply mathematical models capable of accurately representing reality. In this context, the Saint-Venant equations, a widely used one-dimensional model, play a fundamental role. One of the most challenging parameters to measure is the Manning roughness coefficient, whose precise determination is essential as it considers the fluid friction with different surfaces. Estimating this parameter is particularly challenging due to possible variations along the channel. In this scenario, aiming to estimate the Manning coefficient, we propose an approach that utilizes the Augmented Lagrange Method, aiming to minimize errors in solving differential equations at different points, without extrapolating the coefficient variations. To address this issue, we will use the PRIMA (BOBYQA) software with interpolation-based space reduction. This model was applied to the East Fork river, seeking to predict information at different points, considering boundary conditions and estimated parameters.

2 - MAD risk parity portfolios
Jacopo Maria Ricci, Cagin Ararat, Francesco Cesarone, Mustafa Pinar

In this paper, we investigate the features and the performance of the risk parity (RP) portfolios using the mean absolute deviation (MAD) as a risk measure. The RP model is a recent strategy for asset allocation that aims at equally sharing the global portfolio risk among all the assets of an investment universe. We discuss here some existing and new results about the properties of MAD that are useful for the RP approach. We propose several formulations for finding MAD-RP portfolios computationally, and compare them in terms of accuracy and efficiency. Furthermore, we provide extensive empirical analysis based on three real-world datasets, showing that the performances of the RP approaches generally tend to place both in terms of risk and profitability between those obtained from the minimum risk and the Equally Weighted strategies.

3 - Efficient computation of convex hull prices with level and subgradient methods: a computational comparison of dual methods
Sofiane Tanji, Yassine Kamri, François Glineur, Mehdi Madani

Convex Hull (CH) Pricing, used on US electricity markets and raising interest in Europe, is a pricing rule designed to handle markets in the presence of non-convexities such as startup costs and minimum uptimes. In such markets, the market operator makes side payments to generators to cover lost opportunity costs, and CH prices are those who minimize such side payments. These prices can also be obtained by solving a (partial) Lagrangian dual of the original mixed-integer program, where power balance constraints are dualized. Computing CH prices then amounts to minimizing a sum of nonsmooth convex objective functions, where each term depends only on a single generator. The subgradient of each of those terms can be obtained independently by solving smaller mixed-integer programs. In this work, we benchmark a large panel of first-order methods to solve the above dual CH pricing problem. We test eight algorithms, namely the bundle level method and a proximal variant, subgradient methods with various stepsize strategies, two recent parameter-free methods and an accelerated gradient method combined with smoothing. We compare those methods on two representative sets of real-world medium-scale instances, and include a recently proposed purely primal column generation method for reference. Our numerical experiments show that the bundle proximal level method and two variants of the subgradient method perform most favorably compared to other methods recently proposed.

Wednesday, 11:25 - 12:40

■ WD-02

Wednesday, 11:25 - 12:40 - Room: M:O

Conic and polynomial optimization

Stream: Conic optimization: theory, algorithms and applications
Invited session
Chair: *Immanuel Bomze*

1 - **Uncertain standard quadratic optimization under distributional assumptions: a chance-constrained epigraphic approach**
Immanuel Bomze, Daniel de Vicente

The standard quadratic optimization problem (StQP) consists of minimizing a quadratic form over the standard simplex. Without convexity or concavity of the quadratic form, the StQP is NP-hard. This problem has many interesting applications. Sometimes, the data matrix is uncertain. We investigate models where the distribution of the data matrix is known but where both the StQP after realization of the data matrix and the here-and-now problem are indefinite.

2 - **On Tractable Convex Relaxations of Standard Quadratic Optimization Problems under Sparsity Constraints**
Bo Peng, Immanuel Bomze, Yuzhou Qiu, E. Alper Yildirim

Standard quadratic optimization problems (StQPs) provide a versatile modelling tool in various applications. In this paper, we consider StQPs with a hard sparsity constraint, referred to as sparse StQPs. We focus on various tractable convex relaxations of sparse StQPs arising from a mixed-binary quadratic formulation, namely, the linear optimization relaxation given by the reformulation-linearization technique, the Shor relaxation, and the relaxation resulting from their combination. We establish several structural properties of these relaxations in relation to the corresponding relaxations of StQPs without any sparsity constraints, and pay particular attention to the rank-one feasible solutions retained by these relaxations. We then utilize these relations to establish several results about the quality of the lower bounds arising from different relaxations. We also present several conditions that ensure the exactness of each relaxation.

3 - **New results for sparse conic reformulations**
Markus Gabl

Positive semidefinite and copositive optimization offers powerful tools for nonlinear optimization but often suffer from poor scalability due to the matrix variable's high dimension and the conic constraints' computational demands. One remedy is to exploit sparsity patterns in the problem data via sparse reformulations where only submatrices are subjected to conic constraints, while the rest of the variables are dropped from the problem. The theoretical tools from linear algebra that enable these reformulations come from the theory of positive semidefinite and completely positive (cp) matrix completion. Due to the limitations of that theory, exact reformulations are only available for specific sparsity patterns, unless one considers additional problem structure. The latter has hardly been explored so far. We present new results that extend existing literature in several ways. We present new cp completion results that lie outside known sufficient conditions and are based on copositive optimization theory and recent results on the uniqueness of cp-factorizations. Further, we discuss a strategy for sparse copositive reformulations of QPs where only the objective, but not the constraints, exhibit sparsity.

■ WD-03

Wednesday, 11:25 - 12:40 - Room: M:J

Optimization in neural architectures II

Stream: Optimization in neural architectures: convergence and solution characterization
Invited session
Chair: *Maria-Luiza Vladarean*
Chair: *Manish Krishan Lal*

1 - **Vanishing Gradients in Reinforcement Finetuning of Language Models**
Noam Razin

Pretrained language models are commonly aligned with human preferences and downstream tasks via reinforcement finetuning (RFT), which refers to maximizing a (possibly learned) reward function using policy gradient algorithms. In this talk, I will present a recent work identifying a fundamental optimization obstacle in RFT. Namely, the expected gradient for an input vanishes when its reward standard deviation under the model is small, even if the expected reward is far from optimal. Through experiments on an RFT benchmark and controlled environments, as well as a theoretical analysis, we demonstrate that vanishing gradients due to small reward standard deviation are prevalent and detrimental, leading to extremely slow reward maximization. Lastly, we explore ways to overcome vanishing gradients in RFT. We find the common practice of an initial supervised finetuning (SFT) phase to be the most promising candidate, which sheds light on its importance in an RFT pipeline. Moreover, we show that a relatively small number of SFT optimization steps on as few as 1% of the input samples can suffice, indicating that the initial SFT phase need not be expensive in terms of compute and data labeling efforts. Overall, our results emphasize that being mindful of inputs whose expected gradient vanishes, as measured by the reward standard deviation, is crucial for the successful execution of RFT.

2 - **A phase transition between positional and semantic learning in a solvable model of dot-product attention**
Hugo Cui, Freya Behrens, Florent Krzakala, Lenka Zdeborová

We investigate how a dot-product attention layer learns a positional attention matrix (with tokens attending to each other based on their respective positions) and a semantic attention matrix (with tokens attending to each other based on their meaning). For an algorithmic task, we experimentally show how the same simple architecture can learn to implement a solution using either the positional or semantic mechanism. On the theoretical side, we study the learning of a non-linear self-attention layer with trainable tied and low-rank query and key matrices. In the asymptotic limit of high-dimensional data and a comparably large number of training samples, we provide a closed-form characterization of the global minimum of the non-convex empirical loss landscape. We show that this minimum corresponds to either a positional or a semantic mechanism and evidence an emergent phase transition from the former to the latter with increasing sample complexity.

3 - On the spectral bias of two-layer linear networks
Aditya Varre

In this talk, we study the behaviour of two-layer fully connected networks with linear activations trained with gradient flow on the square loss. We show how the optimization process carries an implicit bias on the parameters that depends on the scale of its initialization. The main result of the paper is a variational characterization of the loss minimizers retrieved by the gradient flow for a specific initialization shape. This characterization reveals that, in the small-scale initialization regime, the linear neural network's hidden layer is biased toward having a low-rank structure. To complement our results, we showcase a hidden mirror flow that tracks the dynamics of the singular values of the weights matrices and describe their time evolution. We support our findings with numerical experiments illustrating the phenomena.

■ WD-04

Wednesday, 11:25 - 12:40 - Room: M:M

BARON Tutorial

Stream: Tutorials
Invited session
Chair: *Yi Zhang*

1 - BARON Tutorial
Yi Zhang

This tutorial provides a comprehensive guide to the global MINLP solver BARON, covering details from installation to advanced features. Attendees will learn how to use BARON via different ways and utilize its core functionalities for expert uses. The session will delve into practical examples for illustrating advanced features, such as finding an irreducible inconsistent set for infeasible problems, using BARON as a local solver, exploiting parallelization for MINLP, enforcing convexity and branching priority on problems. Additionally, we will explain customized options, enabling users to tailor the software to their specific needs. This tutorial is designed for both new users looking to get started and experienced users aiming to deepen their understanding of BARON.

This tutorial will last for about 25-30 mins.

■ WD-05

Wednesday, 11:25 - 12:40 - Room: M:N

Optimization for learning II

Stream: Optimization for learning
Invited session
Chair: *Manu Upadhyaya*

1 - Compressed Gradient Descent with Matrix Stepsizes for Non-Convex Optimization
Hanmin Li, Avetik Karagulyan, Peter Richtarik

This paper introduces a new method for minimizing matrix-smooth non-convex objectives through the use of novel Compressed Gradient Descent (CGD) algorithms enhanced with a matrix-valued stepsize. The proposed algorithms are theoretically analyzed first in the single-node and subsequently in the distributed settings. Our theoretical results reveal that the matrix stepsize in CGD can capture the objective's structure and lead to faster convergence compared to a scalar stepsize. As a byproduct of our general results, we emphasize the importance of selecting the compression mechanism and the matrix stepsize in a layer-wise manner, taking advantage of model structure. Moreover, we provide theoretical guarantees for free compression, by designing specific layer-wise compressors for the non-convex matrix smooth objectives. Our findings are supported with empirical evidence.

2 - Optimization flows landing on the Stiefel manifold: continuous-time flows, deterministic and stochastic algorithms
Bin Gao, P.-a. Absil

We study a continuous-time system that solves optimization problems over the set of orthonormal matrices, which is also known as the Stiefel manifold. The resulting optimization flow follows a path that is not always on the manifold but asymptotically lands on the manifold. We introduce a generalized Stiefel manifold to which we extend the canonical metric of the Stiefel manifold. We show that the vector field of the proposed flow can be interpreted as the sum of a Riemannian gradient on a generalized Stiefel manifold and a normal vector. Moreover, we prove that the proposed flow globally converges to the set of critical points, and any local minimum and isolated critical point is asymptotically stable.

3 - Is maze-solving parallelizable?
Romain Cosson

Can you find a short path, without a map? Is maze-solving parallelizable? These intriguing questions can be cast formally into the lens of competitive analysis: a versatile toolkit that has evolved since the 1990s and that has recently been invigorated by continuous optimization techniques. The talk will use these problems as a case study of the connections between online and distributed algorithms, and of the unsuspected role that convex optimization plays in modern approaches for competitive analysis. The talk will be based on published and ongoing works with co-authors. The two problems in the title are known as " layered graph traversal " (introduced by the literature on online algorithms) and " collective tree exploration " (introduced by the literature on distributed algorithms).

■ WD-06

Wednesday, 11:25 - 12:40 - Room: M:H

Linear Programming

Stream: Methods for non-/monotone inclusions and their applications
Invited session
Chair: *Dimitri Papadimitriou*

1 - Fundamental properties of absolute value linear programming problems
Milan Hladik

The area of absolute value linear programming is quite recent and intensively developing. The basic problem has the form of a linear program, but the constraints have an additional absolute value term, which makes them nonlinear and nonsmooth and the problem itself hard to deal with. Even the simple-formulated problem Ax+|x|=b, called the absolute value equations by Mangasarian, is intractable; the classical NP-hard problems such as the Set-Partitioning problem or the standard linear complementarity problem can easily be reduced to this form.

In our presentation, we explore basic properties of absolute value linear programs. In particular, we inspect the geometric shape of the solution set and study conditions for convexity, connectedness and boundedness. We also analyse the corresponding complexity issues, showing that many basic questions are NP-hard to answer. The problem also admits a certain type of duality, based on the theory of interval linear programming. Since one may ask for integral solutions, we investigate conditions ensuring integrality of the vertices of the feasible set, extending the classical results on unimodular matrices in linear programming.

2 - Intensity Modulated Radiotherapy Planning through Linear Programming under uncertainties
Nicole Cristina Cassimiro de Oliveira, Aurelio Oliveira

Radiotherapy is a form of cancer treatment that consists of using a source of ionizing radiation that destroys the tumor. In the context of Radiotherapy Planning, these ionizing particles can vary the intensity of the fluence, achieving a dose distribution with superior compliance. A treatment plan comprises information on how the dose and probability of physical damage from irradiation is distributed within the patient. In order to define the target volumes and organs at risk for the patient, CT scans are performed that allow the oncologist to prescribe doses. The main objective is to deliver enough dose to the tumor for healing, while minimizing the unavoidable dose to healthy organs. However, the prescription may be inaccurate based on the physician's experience, patient positioning or internal organ movement. In view of this, to determine an efficient and effective treatment, optimization techniques are used in order to improve the total dose of radiation applied to the patient. The present work proposes a comparison between the treatment plans obtained by Primal-Dual Interior Point Method using different fuzzy numbers with the Surprise Functions approach. The results obtained were compared by means of histograms, average dose analysis for each structure and verification of dose surface. In conclusion, both proposed fuzzy numbers produced viable treatment plans, inferring that the approached model is an important tool in decision making in the planning of treatment.

3 - New techniques for finding MIP formulations for combinatorial disjunctive constraints
Peter Dobrovoczki, Tamas Kis

Computationally efficient MIP formulations for combinatorial disjunctive constraints (CDC-s) is still a challenging task. A combinatorial disjunctive constraint is a union of simplices with vertices pertaining to a finite set of points in the n-dimensional Euclidean space. With a combinatorial disjunctive constraint we can associate a so-called conflict (hyper) graph. If the hyper graph is an ordinary graph, then we can represent the CDC with a so-called pairwise independent branching scheme. Findig a pairwise independent branching scheme boils down to finding a biclique cover of the conflict graph. On the one hand, the size of the biclique cover has a strong impact on the size of resulting MIP formulation. On the other hand, not all combinatorial disjunctive constraints have a pairwise independent branching scheme representation, e.g., the conflict graph is not an ordinary graph.

We propose a novel method for finding a MIP formulation for arbitrary combinatorial disjunctive constraints. Our method first works on the subgraph of the conflict graph which is an ordinary graph, and finds a biclique cover on it with a new technique. Then, we construct an auxiliary graph to resolve the remaining conflicts. The coloring of this graph yields the other half of the MIP formulation.

We demonstrate our method on a couple of problems from the literature. Our technique for finding a biclique cover of an ordinary graph is new and may be of interest on its own right.

■ WD-07

Wednesday, 11:25 - 12:40 - Room: M:I

Optimization applications II

Stream: Optimization applications
Invited session
Chair: *Edoardo Cesaroni*

1 - SIMP-Based Topology Optimization Of 3D Magnetic Circuits With Mechanical Constraints

Zakaria Houta, Nicolas Lebbe, Thomas Huguet, Frédéric Messine

A topology optimization problem is a problem where the variables allow to define optimally the structure of certain parts of an object. This type of optimization is essential in many applications, such as the design of the structure of aircraft wings or electric motors. In our case, we are interested in optimizing the design of a 3D magnetic circuit which resembles a hall effect thruster. This circuit contains a variable area, which is descretized into small elements, and the optimization has to decide if each of these elements is vacuum or iron. The objective function to be optimized is defined from the magnetic field calculated in a target area. This magnetic field must correspond as well as possible to a fixed objective field, to ensure proper thruster operation.

One of the difficulties encountered in manufacturing topologically optimized magnetostatic structures found by solving the problem described above, is that they are not necessarily mechanically stable. In order to take this mechanical constraint into account, we have developed a SIMP-based topology optimization algorithm which relies on gradient information and on numerical simulations of both the mechanical deformation and the magnetostatic behavior of the structure. By comparing the designs obtained with those from magnetostatic optimization alone, our approach proves effective in obtaining efficient and robust designs.

2 - Enhancing Cost-Optimal Operation of Offshore Hydrogen-Based Power-to-X System

Nikola Mößner

The H2Mare subproject PtX-Wind supported by german government is developing economically feasible methods for creating methanol and other products offshore and independent of the power grid. There is a focus on optimizing operations amidst natural wind power fluctuations. The approach involves a two-stage mixed integer optimization process, starting with an economic analysis to optimize investments in the individual subsystems (components) using the availability of wind power during a type year. The system's operation schedule is optimized in a rolling 2-weeks horizon considering flexibility options for each plant component, such as changing discrete operating points to partial load, scheduled manufacturing using interim product storage and the usage of a battery storage to cover short lull. A demand-side-management approach is thus applied to process engineering. Accuracy of transmission function descriptions as well as material interconnections between system components are studied to improve overall optimization results and performance. In the further work the optimization is to be applied by means of a software-in-the-loop and expanded to a stochastic optimization. The research contributes to renewable energy integration by establishing the feasibility of cost-optimized, grid-independent operation for hydrogen-based Power-to-X systems. The findings will inform future energy policy and industrial uses.

3 - Formulation of an ML-based model for the Assessment of Maximum Sprint Capability in Elite Soccer Players

Edoardo Cesaroni

In recent years, machine learning techniques have gained widespread popularity, fueled by the availability of vast amounts of data across various industries. The sports industry, including football, has also embraced these techniques to gain a competitive edge. In our work, we focus on estimating and analyzing the maximum acceleration capability of football players using tracking data, which defines the player's position on the field at each frame during a match. This information can contribute to more precise tactical preparation and personalized training strategies. We then develop a regression-based model that reconstructs a profile per player, capturing the relationship between acceleration and speed during a maximum sprint when the athlete performs at their peak capability. By applying this model to data sampled at regular intervals during a match, we can identify variations in this profile, which reflect changes in the player's physical condition. It also allows us to evaluate how the time it takes for a player to cover specific distances at a certain initial speed varies over 90 minutes. Furthermore, we employ clustering, to identify players with similar athletic characteristics and performance. This approach deepens our understanding of player capabilities and enables data-driven decision-making in sports analysis and training. It provides a deeper level of insight into player performance, facilitating player evaluation.

Wednesday, 14:10 - 15:50

■ WE-02

Wednesday, 14:10 - 15:50 - Room: M:O

Recent advances in computer-aided analyses of optimization algorithms I

Stream: Conic optimization: theory, algorithms and applications
Invited session
Chair: *Adrien Taylor*
Chair: *Manu Upadhyaya*

1 - Last-Iterate Convergence of Extragradient-based Methods
Eduard Gorbunov, Adrien Taylor, Samuel Horvath, Nicolas Loizou, Gauthier Gidel

Extragradient (EG) and Optimistic Gradient (OG) methods are among the most popular methods for solving monotone Lipschitz variational inequalities. However, despite the long history of these methods, the questions about their last-iterate convergence were resolved very recently. In this talk, we will discuss several results about the last-iterate convergence of EG and OG under Lipschitzness and the monotonicity of the operator. In particular, we will focus on the proofs of these results that were obtained via computer. Extensions to the case of negative comonotone operators will also be considered.

2 - Automated tight Lyapunov analysis for first-order methods
Manu Upadhyaya, Sebastian Banert, Adrien Taylor, Pontus Giselsson

We present a methodology for establishing the existence of quadratic Lyapunov inequalities for a wide range of first-order methods used to solve convex optimization problems. In particular, we consider (i) classes of optimization problems of finite-sum form with (possibly strongly) convex and possibly smooth functional components, (ii) first-order methods that can be written as a linear system on state-space form in feedback interconnection with the subdifferentials of the functional components of the objective function, and (iii) quadratic Lyapunov inequalities that can be used to draw convergence conclusions. We present a necessary and sufficient condition for the existence of a quadratic Lyapunov inequality within a predefined class of Lyapunov inequalities, which amounts to solving a small-sized semidefinite program. We showcase our methodology on several first-order methods that fit the framework. Most notably, our methodology allows us to significantly extend the region of parameter choices that allow for duality gap convergence in the Chambolle-Pock method.

3 - Second-order interpolation conditions for univariate functions, towards a tight analysis of second-order optimization methods
Anne Rubbens, Nizar Bousselmi, Julien Hendrickx, François Glineur

We focus in this work on the derivation of interpolation constraints for given function classes, that is necessary and sufficient conditions that a data set must satisfy to ensure the existence of a function of the class of interest, defined on the whole space and interpolating the data.

Interpolation constraints allow to manipulate only data sets instead of continuous functions, and are essential to obtain tight guarantees on the worst-case performance of optimization methods on given function classes, where one manipulates a finite set of iterates (and associated evaluations of a function and its derivatives). While interpolation constraints are available for a wide range of first-order function classes (smooth, convex,..), allowing to tightly analyze first-order method, this is not the case for second-order function classes. We thus propose interpolation conditions for (convex) univariate function with Lipschitz second derivative, with the objective of tightly analyzing second-order methods.

4 - Analysis of Second-Order Methods via non-convex Performance Estimation
Nizar Bousselmi, Anne Rubbens, Julien Hendrickx, François Glineur

The "Performance Estimation Problem framework" computes numerically the exact worst-case performance of a given optimization method on a given function class, whenever interpolation constraints for this class are available. It allowed to determine the convergence rate of numerous first-order methods and also to design optimal methods. The convex formulation of the Performance Estimation Problems can be efficiently solved but it restricted the analysis to certain types of methods. Recently, some works have been proposed to tackle a non-convex formulation of Performance Estimation Problems thanks to recent global branch and bound solvers. It allows more flexibility in the analysis but at a computational cost. In this work, we propose to exploit the non-convex Performance Estimation Problem and recently derived second-order interpolation conditions, to analyze second-order optimization methods.

■ WE-03

Wednesday, 14:10 - 15:50 - Room: M:J

Structured sparse optimization

Stream: Variational analysis: theory and algorithms
Invited session
Chair: *Xiaoqi Yang*

1 - Mix Sparse Optimization: Theory and Applications
Yaohua Hu

Structured sparse optimization has been extensively applied in the modelling of many important problems in various disciplines. The mix sparse structure is inherited in a wide class of practical applications, namely, the sparse structure appears as the inter-group and intra-group manners simultaneously. In this talk, we will discuss the nonconvex regularization method for mix sparse optimization problem, as well as a first-order iterative algorithm, and present its consistency theory, asymptotic theory and convergence theory. Applications of mix sparse optimization method to gene regulatory networks and differential optical absorption spectroscopy will be presented.

2 - Efficient Low-rank Identification via Accelerated Iteratively Reweighted Nuclear Norm Minimization
Hao Wang

This paper considers the problem of minimizing the sum of a smooth function and the Schatten-p norm of the matrix. Our contribution involves proposing accelerated iteratively reweighted nuclear norm methods designed for solving the nonconvex low-rank minimization problem. Two major novelties characterize our approach. Firstly, the proposed method possesses a rank identification property, enabling the provable identification of the "correct" rank of the stationary point within a finite number of iterations. Secondly, we introduce an adaptive updating strategy for smoothing parameters. This strategy automatically fixes parameters associated with zero singular values as constants upon detecting the "correct" rank while quickly driving the rest parameters to zero. This adaptive behavior transforms the algorithm into one that effectively solves smooth problems after a few iterations, setting our work apart from existing iteratively reweighted methods for low-rank optimization. We prove the global convergence of the proposed algorithm, guaranteeing that every limit point of the iterates is a critical point. Furthermore, a local convergence rate analysis is provided under the Kurdyka-Lojasiewicz property. We conduct numerical experiments using both synthetic and real data to showcase our algorithm's efficiency and superiority over existing methods.

3 - An Inexact Projected Regularized Newton Method for Fused Zero-norms Regularization Problems
Yuqia Wu, Shaohua Pan, Xiaoqi Yang

In this talk, we are concerned with structured ell_0-norms regularization problems, with a twice continuously differentiable loss function and a box constraint. This class of problems have a wide range of applications in statistics, machine learning and image processing. To the best of our knowledge, there is no effective algorithm in the literature for solving them. In this work, we first obtain a polynomial-time algorithm to find a point in the proximal mapping of the fused ell_0-norms with a box constraint based on dynamic programming principle. We then propose a hybrid algorithm of proximal gradient method and inexact projected regularized Newton method to solve structured ell_0-norms regularization problems. The whole sequence generated by the algorithm is shown to be convergent by virtue of a non-degeneracy condition, a curvature condition and a Kurdyka-Lojasiewicz property. A superlinear convergence rate of the iterates is established under a locally H"olderian error bound condition on a second-order stationary point set, without requiring the local optimality of the limit point. Finally, numerical results highlight the features of our considered model, and the superiority of our proposed algorithm.

4 - Projectional coderivatives with applications
Xiaoqi Yang

The Lipschitz-like property relative to a set is important in applications. In this talk we introduce a projectional coderivative of set-valued mappings and present its calculations in some special cases. We apply this coderivative to obtain a complete characterization for a set-valued mapping to have the Lipschitz-property relative to a closed and convex set. For an extended real-valued function, we apply the obtained results to investigate its Lipschitz continuity relative to a closed and convex set and the Lipschitz-like property of a level-set mapping relative to a half line. We apply our results to study the Lipschitz-like property of the solution mapping of a parametric affine variational inequality problem.

■ WE-04

Wednesday, 14:10 - 15:50 - Room: M:M

Large-scale optimization I

Stream: Large-scale optimization
Invited session
Chair: *Max Nilsson*

1 - A divergence-based condition to ensure quantile improvement in black-box global optimization
Thomas Guilmeau, Emilie Chouzenoux, Víctor Elvira

Black-box global optimization aims at seeking for the minimizers of an objective function whose analytical form is not known. To do so, many state-of-the-art methods rely on sampling-based strategies, where sampling distributions are built in an iterative fashion, so that their mass concentrate where the objective function is low. Despite empirical success, the convergence of these methods remains difficult to show theoretically. In this work, we introduce a new framework, based on divergence-decrease conditions, to study and design black-box global optimization algorithms. We show that the information-geometric optimization approach fits within our framework, which yields a new proof for its convergence analysis. We also establish a quantile improvement result for two novel algorithms, one related with the cross-entropy approach with mixture models, and another using heavy-tailed sampling distributions.

2 - Low rank of the matrix LASSO under RIP with consequences for fast large-scale algorithms
Andrew McRae

A popular and efficient algorithmic approach to large-scale low-rank matrix recovery is optimization over matrices constrained to be low rank. However, this approach has limited theoretical guarantees due to nonconvexity. We address this problem for the popular “matrix LASSO” estimator. In general, computing this estimator requires many full singular value decompositions of large matrices. To get around this, first, we show that the LASSO estimator is exactly low-rank under similar conditions to those in classic low-rank matrix sensing error bounds. Second, this low solution rank guarantees the success of several low-rank algorithms. Specifically, we show that if (a) the ground truth is low-rank, (b) the (linear) sensing operator satisfies the matrix restricted isometry property (RIP), and (c) the measurement error is small enough relative to the amount of regularization, then the (unique) LASSO estimate has rank (approximately) bounded by that of the ground truth. From this, we conclude (a) that a low-rank–projected proximal algorithm (which can take advantage of fast partial SVD algorithms), will converge linearly to the global solution from any initialization, and (b) that the nonconvex landscape of a low-rank Burer-Monteiro–factorized problem formulation is benign in the sense that all second-order critical points yield the global solution.

3 - Computing Augustin Information via Hybrid Geodesically Convex Optimization
Guan-Ren Wang

We propose a Riemannian gradient descent with the Poincaré metric to compute the Augustin information, a widely used quantity for characterizing exponential error behaviors in information theory. We provide a non-asymptotic optimization error guarantee for this algorithm. Our result is based on a novel hybrid analysis of Riemannian gradient descent for functions that are geodesically convex in a Riemannian metric and geodesically smooth in another. Numerical experimental results demonstrate the empirical efficiency of the algorithm.

■ WE-05

Wednesday, 14:10 - 15:50 - Room: M:N

Randomized optimization algorithms part 1/2

Stream: Randomized optimization algorithms
Invited session
Chair: *Laurent Condat*

1 - An Inexact Restoration based algorithm with random models for unconstrained noisy optimization
Simone Rebegoldi, Benedetta Morini

In this talk, we focus on unconstrained differentiable optimization problems where the evaluation of the objective function and its gradient is noisy. First, we consider a constrained reformulation of the original problem based on the Inexact Restoration (IR) approach where y is the noise level of the evaluation, h(y) is a non-negative value measuring its accuracy in probability, and the constraint h(y) = 0 represents the ideal case where the estimates are evaluated exactly. Such a reformulation is viable whenever one uses sample average approximations as the noisy evaluations of both function and gradient. Then, we propose a trust-region algorithm with first-order random models that leverages the IR constrained reformulation of the problem. At each iteration, the proposed algorithm enforces the sufficient accuracy in probability of the function and gradient estimates, and then employs an acceptance test involving both the noisy function and the infeasibility measure. We prove an iteration complexity result that bounds the expected number of iterations needed to achieve an approximate first-order optimality point. The numerical experiments on some least squares problems show that the proposed algorithm tends to compute larger trust-region radii than other previously proposed trust-region algorithms with random models, and compares well with them.

2 - An Optimal Structured Zeroth-order Algorithm for Non-smooth Optimization
Cesare Molinari

Finite-difference methods are a class of algorithms designed to solve black-box optimization problems by approximating a gradient of the target function on a set of directions. In black-box optimization, the non-smooth setting is particularly relevant since, in practice, differentiability and smoothness assumptions cannot be verified. To cope with nonsmoothness, several authors use a smooth approximation of the target function and show that finite difference methods approximate its gradient. Recently, it has been proved that imposing a structure in the directions allows improving performance. However, only the smooth setting was considered. To close this gap, we introduce and analyze O-ZD, the first structured finite-difference algorithm for non-smooth black-box optimization. Our method exploits a smooth approximation of the target function and we prove that it approximates its gradient on a subset of random orthogonal directions. We analyze the convergence of O-ZD under different assumptions. For non-smooth convex functions, we obtain the optimal complexity. In the non-smooth non-convex setting, we characterize the number of iterations needed to bound the expected norm of the smoothed gradient. For smooth functions, our analysis recovers existing results for structured zeroth-order methods for the convex case and extends them to the non-convex setting. We conclude with numerical simulations, observing that our algorithm has very good practical performance.

3 - Random Newton-type Iterations with Application to Electronic Structure Determination
Titus Pinta

We combine two separate theories in the context of molecular electronic structure determination. On the one hand, we present a general framework of Newton-type methods for root-finding problems, with constraints, that characterizes classes of functions for which superlinear or better convergence is guaranteed. On the other hand we present a theory of random function iterations, where the functions are randomly selected mappings generated from familiar optimization problems. Putting these two ideas together yields a Markov operator whose kernel is given by the Newton-type mapping of a random variable to the critical point set of a randomly generated optimization problem. The "fixed points" that we seek are invariant measures of the corresponding Markov operator. In the context of molecular electronic structure determination, the optimization problem is a maximum liklihood estimation problem. We present the foundations of both theoretical threads and preliminary results of their combination toward reconstructing molecular electronic structures from X-ray free electron laser measurements.

4 - A variable metric proximal stochastic gradient method with dynamical variance reduction
Andrea Sebastiani, Pasquale Cascarano, Giorgia Franchini, Erich Kobler, Federica Porta

Training deep learning models typically involves minimizing the empirical risk over large datasets, dealing with a potentially non-differentiable regularization. In this work, we present a stochastic gradient method tailored for classification problems, that are ubiquitous in the scientific field. The variance of the objective's gradients is controlled using an automatic sample size selection, along with a variable metric to precondition the stochastic gradient directions. Additionally, a non-monotone line search is employed for the step size selection. The convergence of this first-order algorithm can be derived for both convex and non-convex objective functions. The numerical experiments suggest that the proposed approach performs comparably to state-of-the-art methods in training non only standard statical models for binary classification but also artificial neural networks for multi-class image classification.

■ WE-06

Wednesday, 14:10 - 15:50 - Room: M:H

Higher-order Methods in Mathematical Programming I

Stream: Challenges in nonlinear programming
Invited session
Chair: *Mathias Staudigl*

1 - Spectral Preconditioning for Gradient Methods on Graded Non-convex Functions
Nikita Doikov

The performance of optimization methods is often tied to the spectrum of the objective Hessian. Yet, conventional assumptions, such as smoothness, do often not enable us to make finely-grained convergence statements – particularly not for non-convex problems. Striving for a more intricate characterization of complexity, we introduce a unique concept termed graded non-convexity. This allows to partition the class of non-convex problems into a nested chain of subclasses. Interestingly, many traditional non-convex objectives, including partially convex problems, matrix factorizations, and neural networks, fall within these subclasses. Then, we propose gradient methods with spectral preconditioning, which employ inexact top eigenvectors of the Hessian to address the ill-conditioning of the problem, contingent on the grade. Our analysis reveals that these new methods provide provably superior convergence rates compared to basic gradient descent on applicable problem classes, particularly when large gaps exist between the top eigenvalues of the Hessian. Joint work with Sebastian U. Stich and Martin Jaggi.

2 - Barrier Algorithms for Constrained Non-Convex Optimization
Pavel Dvurechensky, Mathias Staudigl

A key problem in mathematical imaging, signal processing and computational statistics is the minimization of non-convex objective functions that may be non-differentiable at the relative boundary of the feasible set. This paper proposes a new family of first- and second-order interior-point methods for non-convex optimization problems with linear and set constraints, combining self-concordant barriers with quadratic and cubic regularization respectively. Our approach is based on a potential-reduction mechanism and, under the Lipschitz continuity of the corresponding derivative with respect to the local barrier-induced norm, attains a suitably defined class of approximate first- or second-order KKT points with worst-case iteration complexity typical for unconstrained optimization. Based on these findings, we develop also new path-following schemes attaining the same complexity, modulo adjusting constants.

3 - Accelerated cubic regularized quasi-newton methods
Dmitry Kamzolov

In this paper, we propose the first Quasi-Newton method with a global convergence rate of $O(fracL_1R2k)$ for general convex functions. Quasi-Newton methods, such as BFGS, SR-1, are well-known for their impressive practical performance. However, they are theoretically slower than gradient descent for general convex functions. This gap between impressive practical performance and poor theoretical guarantees was an open question for a long period of time. In this paper, we make a significant step to close this gap. We improve upon the existing rate and propose the Cubic Regularized Quasi-Newton Method with a convergence rate of $O(fracL_1R2k)$. The key to achieving this improvement is to use the Cubic Regularized Newton Method over the Damped Newton Method as an outer method, where the Quasi-Newton update is an inexact Hessian approximation. Using this approach, we propose the first Accelerated Quasi-Newton method with a global convergence rate of $O(fracL_1R2k2)$ for general convex functions. In special cases where we have access to additional computations, for example Hessian-vector products, we can improve the inexact Hessian approximation and achieve a global convergence rate of $O(fracL_2R3k3)$, which make it intermediate second-order method. To make these methods practical, we introduce the Adaptive Inexact Cubic Regularized Newton Method and its accelerated version, which provide real-time control of the approximation error. We show that the p

4 - Relaxation Approaches for Nonlinear Sparse Optimization Problems
Steffensen Sonja

In many applications, sparse solutions are favoured over non-sparse solutions with comparable objective value. One approach to induce sparsity relies on the ell_0 norm in the objective. Often this semicontinuous function is approximated using the continuous and convex ell_1-norm instead. However, this approximation can lead to suboptimal results with respect to the sparsity of the solution. We will present alternative exact reformulations (with respect to the ell_0 norm) and relaxations leading to standard nonlinear but nonconvex programs. In our talk we will discuss and relate the relations between the different reformulations in particular with respect to the original problem. Furthermore, we accompany the theoretical results by some numerical tests using randomly generated data sets and give a brief outlook on the application of our approach to sparse optimal control problems.

■ WE-07

Wednesday, 14:10 - 15:50 - Room: M:I

Optimization applications III

Stream: Optimization applications
Invited session
Chair: *Atanu Maji*

1 - A homotopy continuation method for flowsheet optimization
David Mogalle, Tobias Seidel, Michael Bortz, Karl-Heinz Küfer

Distillation processes are a major part of the chemical industry. In order to find an optimal operating point for these processes, an optimization problem with a large number of variables and a large number of highly nonlinear equality constraints - the so-called MESH equations - has to be solved. To deal with the high dimensionality, one can reduce the problem into a much smaller one by solving most of the MESH equations implicitly. However, this reduction method is only guaranteed to work for feasible points of the original problem. Thus, we get an optimization problem where the remaining constraints can no longer be evaluated at every point. To mitigate this issue, we propose a homotopy continuation method. First, we solve a simplified problem where we can guarantee the evaluation of the constraint functions on the whole domain. Then, we gradually transform our simplified problem back to the original one. The step size of this transformation is chosen in a way such that we stay in the region where evaluation of the constraints is possible. Finally, we prove that the homotopy continuation method converges under reasonable assumptions on the parameters that characterize the distillation process.

2 - An integrated optimization-simulation system for the repair sequence problem of electricity distribution networks

Atanu Maji, Anna Livia Croella, Laura Palagi, Alberto Tofani

Electricity distribution networks are complex systems designed to deliver electricity from power stations to consumers efficiently. When disruptions occur, effective strategies are crucial to minimize the number of disconnected electrical stations and users. While some stations can swiftly reconnect via pre-installed automated switches or tele-controlled devices, others necessitate manual intervention from operator crews. We propose a Repair Sequence Optimization (RSO) approach for large power distribution grids that integrates an optimization model within the simulator RecSIM, developed by ENEA of Italy in collaboration with an electricity distributor Areti S.p.A. The RSO employs a mathematical model for the Repair Sequence problem, which can be formulated as scheduling for parallel identical machines.

We tested our approach on a realistic power distribution grid of a large metropolitan city, which shows the efficiency of the optimization procedure. Each simulation was performed considering a randomly generated scenario, whose results were subsequently analysed and verified to ensure consistency with actual restoration operations.

3 - Financial optimization routine for hybrid distributed power generation with battery storage system

Paulo Rotella Junior, Arthur Leandro Guerra Pires, Luiz Celio Souza Rocha, Karel Janda

Intermittency is one of the main challenges associated with using renewable energy sources, but it can be mitigated through hybrid generation and battery storage systems. This study proposes a financial optimization routine to identify the optimal configuration for residential hybrid distributed generation systems (photovoltaic-wind), including battery energy storage, aiming to optimize the economic indicators Net Present Value and Levelized Cost of Energy simultaneously. Initially, the Design of Experiments approach, specifically the Response Surface Methodology, was adopted to model the objective functions, utilizing four input variables: X1: wind energy representativeness in the project (in %), X2: Demand Level, in kWh, X3: Battery Type: Lead-acid or Lithium-ion, and X4: Scenario Type: Peak/Intermediate or Total. The Desirability method was chosen for multi-objective optimization. The results that simultaneously optimize both responses are: X1 = 7.6%, X2 = 407.1 kWh, X3 = Lead-acid battery, and X4 = Total (Scenario Type). Finally, the findings emphasize the importance of incentives to promote residential wind power generation, as it is currently less cost-competitive than photovoltaic energy. Moreover, the higher cost of wind energy and the expense of batteries can render residential power generation projects unviable under certain circumstances.

4 - The Bin Packing Problem with Setups: properties and formulations

Fabio Ciccarelli, Roberto Baldacci, Stefano Coniglio, Fabio Furini

In this work, we introduce a novel extension to the classical Bin Packing Problem (BPP), the Bin Packing Problem with Setups (BPPS), in which items are partitioned into classes with associated setup weights and costs. The objective is therefore to find an optimal packing, i.e., a solution that minimizes the sum of bins and setup costs. The BPPS presents a complex extension of the traditional Bin Packing Problem, and it holds significant implications across various logistical and industrial domains.

We explore different possible formulations for the BPPS, investigating their theoretical properties and conducting an analysis of the bounds provided by their linear relaxations. By doing this, we show that it is possible to improve the tightness of the natural ILP formulation of the problem by adding specific constraints or modifying its structure with symmetry breaking approaches. Specifically, we adapt the Asymmetric Representative Formulation for the Vertex Coloring Problem (VCP) to our case, in order to mitigate the symmetry present in natural BPPS formulations.

The theoretical results are confirmed by our computational experiments, which show that the performances of the modified formulations are better than the ones of the standard alternative.

Our findings contribute to a deeper understanding of the BPPS landscape, its theoretical properties and practical implications.

Wednesday, 16:20 - 18:00

■ WF-02

Wednesday, 16:20 - 18:00 - Room: M:O

Recent advances in computer-aided analyses of optimization algorithms II

Stream: Conic optimization: theory, algorithms and applications
Invited session
Chair: *Adrien Taylor*
Chair: *Manu Upadhyaya*

1 - Exact worst-case convergence rates of gradient descent: a complete analysis for all constant stepsizes over nonconvex and convex functions
Teodor Rotaru, François Glineur, Panagiotis Patrinos

We derive exact worst-case convergence rates on the minimum gradient norm of the gradient descent iterates with constant stepsizes. Our analysis covers all possible stepsizes and arbitrary upper/lower bounds on the objective function's curvature, thus including convex, strongly convex, and weakly convex (hypoconvex) objective functions. Unlike prior approaches, relying solely on inequalities connecting consecutive iterations, our analysis employs inequalities involving an iterate and its two predecessors. While this complicates the proofs to some extent, it enables to achieve, for the first time, an exact full-range analysis of gradient descent for any constant stepsizes, (covering, in particular, normalized stepsizes greater than one), whereas the literature contained only conjectured rates of this type. Our analysis accommodates arbitrary bounds on both the upper and lower curvatures of the smooth objective function. In the nonconvex case, this extends existing partial results that are valid only for gradient Lipschitz functions (i.e., where lower and upper bounds on curvature are equal), leading to improved rates for weakly convex functions. From our exact rates, we deduce the optimal constant stepsize for gradient descent. Leveraging our analysis, we furthermore introduce a new variant of gradient descent based on a unique, fixed sequence of variable stepsizes, demonstrating its superiority over any constant stepsize schedule in the (strongly) convex case.

2 - A Linear-Quadratic Program for Estimating Performance of Convex Optimization Algorithm
Ashkan Panahi

Performance estimation programs (PEP) are a powerful approach for analyzing and designing optimization algorithms using computer programs. However, their superior performance often comes with the price of a rapidly increasing amount of computation. In this talk, we discuss certain relaxations of the standard PEP formulation that lead to a quadratic optimization problem with linear constraints, hence referred to as Linear-Quadratic PEP (LQ-PEP). We show that LQ-PEP can provide both asymptotic and non asymptotic bounds on convergence with a fixed amount of computation. In certain scenarios, we show that LQ-PEP can be tied to the well-known linear-quadratic control problem. In this way, LQ-PEP may employ tools from the theory of control, especially spectral theory, and harmonic analysis, to forego traditional methods such as potential functions. We showcase LQ-PEP in a few case studies. We look at the problem of decentralized, asynchronous optimization with delays, under a directed communication graph, and with constraints. For this difficult scenario, we propose a novel optimization algorithm called Asymptotic Double Averaging and Gradient projection (ASY-DAGP). We present the analysis of ASY-DAGP by the LQ-PEP formalism and obtain the first set of generic, sub-linear speeds of convergence in the above scenario. Our result characterizes the error induced by asynchrony and delay in a parameter called "delay factor".

3 - On the convergence rate of the difference-of-convex algorithm (DCA)
Hadi Abbaszadehpeivasti

In this talk, I present the non-asymptotic convergence rate of the DCA (difference-of-convex algorithm), also known as the convex-concave procedure, with two different termination criteria that are suitable for smooth and non-smooth decompositions, respectively. The DCA is a popular algorithm for difference-of-convex (DC) problems and known to converge to a stationary point of the objective under some assumptions. I derive a worst-case convergence rate of O(1/sqrt(N)) after N iterations of the objective gradient norm for certain classes of DC problems, without assuming strong convexity in the DC decomposition and give an example which shows the convergence rate is exact. I also provide a new convergence rate of O(1/N) for the DCA with the second termination criterion. Moreover, I present a new linear convergence rate result for the DCA under the assumption of the Polyak-Łojasiewicz inequality. The novel aspect of this analysis is that it employs semidefinite programming performance estimation.

4 - Non-expansiveness for frugal resolvent splitting methods, using PEP
Anton Åkerman, Emanuele Naldi, Enis Chenchene, Sebastian Banert, Pontus Giselsson

This talk focuses on frugal resolvent splitting methods with minimal lifting. A PEP-formulation is presented, and used to derive conditions for non-expansiveness for this class of algorithms. Some further results are presented, and a few notes are given on more general conditions for convergence.

■ WF-03

Wednesday, 16:20 - 18:00 - Room: M:J

Splitting algorithms

Stream: Variational analysis: theory and algorithms
Invited session
Chair: *Francisco Javier Aragón Artacho*

1 - A splitting method for computing Wasserstein Barycenters
Welington de Oliveira, Daniel Mimouni, Malisani Paul, Jiamin Zhu

The Wasserstein barycenter (WB) is an important tool for summarizing sets of probability measures. It finds applications in applied probability, clustering, image processing, etc. When the measures' supports are finite, computing a (balanced) WB can be done by solving a linear optimization problem whose dimensions generally exceed standard solvers' capabilities. In the more general setting where measures have different total masses, we propose a convex nonsmooth optimization formulation for the so-called unbalanced WB problem. Due to their colossal dimensions, we introduce a decomposition scheme based on the Douglas-Rachford splitting method that can be applied to both balanced and unbalanced WB problem variants. Our algorithm, which has the interesting interpretation of being built upon averaging marginals, operates a series of simple (and exact) projections that can be parallelized and even randomized, making it suitable for large-scale datasets.

2 - Regularity of sets under a reformulation in a product space of reduced dimension
Rubén Campoy

Different notions on regularity of sets and of collection of sets play an important role in the analysis of the convergence of projection algorithms in nonconvex scenarios. While some projection algorithms can be applied to feasibility problems defined by finitely many sets, some other require the use of a product space reformulation to construct equivalent problems with two sets. In this talk we analyze how some regularity properties are preserved under a reformulation in a product space of reduced dimension. This allows us to establish local linear convergence of parallel projection methods which are constructed through this reformulation.

3 - First-order splitting methods for decentralized optimization
Felipe Atenas, Matthew Tam, Minh Dao ngoc

We consider finitely many agents over a connected network, whose aim is to cooperatively solve a convex optimization problem with composite sum structure, without having direct access to the information of the overall network. We propose primal-dual splitting methods of proximal-gradient type that use no central coordinator, performing forward-backward steps separately for each agent, and rounds of communication among neighbors. The algorithms we propose extend the so-called PG-EXTRA method for decentralized minimization problems, to three-operator splitting and convex-concave min-max problems.

4 - The Boosted Double-Proximal Subgradient Algorithm for Nonconvex Optimization
Francisco Javier Aragón Artacho

In this talk we present a new splitting algorithm that can be used to tackle very general structured nonconvex minimization problems. Specifically, we assume that the objective function can be expressed as the sum of a locally Lipschitz function f that satisfies the descent lemma plus a l.s.c. prox-bounded function g minus a finite sum of compositions of continuous convex functions and differentiable functions with Lipschitz continuous gradients. Our algorithm makes use of subgradients of f, the gradients of the differentiable functions, and proximal steps of the functions g and of the conjugates of the convex functions. As, in addition, it includes a line-search step that allows to improve its performance, we name it Boosted Double-proximal Subgradient Algorithm (BDSA). If time permits, we will conclude the talk with some numerical experiments demonstrating the good performance of BDSA.

This is a joint work with David Torregrosa-Belén and Pedro Pérez-Aros.

■ WF-04

Wednesday, 16:20 - 18:00 - Room: M:M

Multiobjective Optimization I

Stream: Multiobjective optimization
Invited session
Chair: *Radu Strugariu*

1 - On an open problem related to the Gale's example in conic linear programming
Constantin Zalinescu

In the paper [Duality gap function in infinite dimensional linear programming, J. Math. Anal. Appl., 437 (2016), 1-15], the authors, N.T. Vinh, D.S. Kim, N.N. Tam and N.D. Yen, formulate an open problem concerning the perturbed Gale's example. Our aim is to present a solution to that open problem. First, we derive the main duality results for the conic linear programming problems using Rockafellar's perturbation method in convex analysis.

2 - Conic cancellation laws and applications
Marius Durea

We discuss, on finite and infinite dimensional normed vector spaces, some versions of Rådström cancellation law that are suited for applications to set optimization problems. We call our results "conic" variants of the celebrated result of Rådström, since they involve the presence of an ordering cone. Several adaptations to this context of some topological properties of sets are studied and some applications to subdifferential calculus associated to set-valued maps and to necessary optimality conditions for constrained set optimization problems are given.

3 - Sharp solutions for nonsmooth optimization problems under generalized convexity
Radu Strugariu

We present some sufficient optimality conditions for sharp solutions of constrained nonsmooth optimizations problems in terms of several kinds of subgradients. We focus as well on some generalized convexity assumptions on sets (Shapiro property) and for functions (approximate convexity).

4 - Solving vector optimization problems with ADMM
Daniel Hernandez Escobar, Joakim da Silva, Jens Sjölund

We consider the numerical solution to vector optimization problems. We focus mainly on convex problems whose preference order is defined by a generalized inequality. To approximate the set of efficient solutions, we employ the Alternating Direction Method of Multipliers and a parallel strategy. Although our approach may produce duplicate solutions, it can leverage GPUs or TPUs to achieve fast convergence. We illustrate this by solving multi-objective linear programs. Moreover, we outline how to adapt this technique to solve other problem classes, for instance, when a Lorentz cone defines the generalized inequality.

■ WF-05

Wednesday, 16:20 - 18:00 - Room: M:N

Randomized optimization algorithms part 2/2

Stream: Randomized optimization algorithms
Invited session
Chair: *Egor Shulgin*
Chair: *Laurent Condat*

1 - Sketch-and-Project Meets Newton Method: Global O(1/k2) Convergence with Low-Rank Updates
Slavomír Hanzely

In this paper, we propose the first sketch-and-project Newton method with fast O(1/k2) global convergence rate for self-concordant functions. Our method, SGN, can be viewed in three ways: i) as a sketch-and-project algorithm projecting updates of Newton method, ii) as a cubically regularized Newton method in sketched subspaces, and iii) as a damped Newton method in sketched subspaces. SGN inherits best of all three worlds: cheap iteration costs of sketch-and-project methods, state-of-the-art O(1/k2) global convergence rate of full-rank Newton-like methods and the algorithm simplicity of damped Newton methods. Finally, we demonstrate its comparable empirical performance to baseline algorithms.

2 - TAMUNA: Doubly-Accelerated Distributed Optimization with Local Training, Compression, and Partial Participation
Laurent Condat

In distributed optimization, a large number of machines perform computations in parallel and communicate back and forth with a distant server. Communication is the main bottleneck, as it is typically slow and costly. In addition to communication-efficiency, a robust algorithm should allow for partial participation, since some machine may not be able to participate in the process at certain iterations. To reduce the communication load, two strategies are popular: 1) communicate less frequently; 2) compress the communicated vectors. We introduce TAMUNA, the first algorithm that harnesses these two strategies jointly, provably benefits from them, and allows for partial participation. TAMUNA converges linearly to the minimizer of the sum of smooth and strongly convex functions, with double acceleration: its communication complexity exhibits a better dependency on the condition number of the functions and on the dimension.

3 - Unveiling the Power of Adaptive Methods Over SGD: A Parameter-Agnostic Perspective
Xiang Li

Adaptive gradient methods are popular in optimizing modern machine learning models, yet their theoretical benefits over vanilla Stochastic Gradient Descent (SGD) remain unclear. This presentation examines the convergence of SGD and adaptive gradient methods when optimizing stochastic non-convex functions without the need for setting algorithm hyper-parameters based on problem-specific knowledge. First, we explore smooth functions and compare SGD to well-known adaptive methods like AdaGrad, Normalized SGD with Momentum (NSGD-M), and AMSGrad. Our findings reveal that while untuned SGD can reach the optimal convergence rate, it comes at the expense of an unavoidable catastrophic exponential dependence on the smoothness constant. Adaptive methods, on the other hand, eliminate this reliance without needing to know the smoothness constant in advance. We then look at a broader group of functions characterized by (L0, L1) smoothness. Here, SGD is shown to fail without proper tuning. We present the first instance of tuning-free convergence with adaptive methods in this context, specifically with NSGD-M, achieving near-optimal rate despite an exponential dependence on the L1 constant. We also demonstrate that this dependency is unavoidable for a family of normalized momentum methods.

■ WF-06

Wednesday, 16:20 - 18:00 - Room: M:H

Stochastic Gradient Methods: Bridging Theory and Practice

Stream: Challenges in nonlinear programming
Invited session
Chair: *Simon Weissmann*

1 - Stochastic Optimization under Hidden Convexity
Ilyas Fatkhullin, Niao He, Yifan Hu

In this work, we consider constrained stochastic optimization problems under hidden convexity, i.e., those that admit a convex reformulation via non-linear (but invertible) map c(). A number of non-convex problems ranging from optimal control, revenue and inventory management, to convex reinforcement learning all admit such a hidden convex structure. Unfortunately, in the majority of applications considered, the map c() is unavailable or implicit; therefore, directly solving the convex reformulation is not possible. On the other hand, the stochastic gradients with respect to the original variable are often easy to obtain. Motivated by these observations, we examine the basic projected stochastic (sub-) gradient methods for solving such problems under hidden convexity. We provide the first sample complexity guarantees for global convergence in smooth and non-smooth settings. Additionally, in the smooth setting, we improve our results to the last iterate convergence in terms of function value gap using the momentum variant of projected stochastic gradient descent.

2 - On Almost Sure Convergence Rates for Stochastic Gradient Methods
Sara Klein

Stochastic gradient methods are among the most important algorithms in training machine learning problems. While classical assumptions such as strong convexity allow for simple analysis, they are often not satisfied in applications. In recent years, global and local gradient domination properties have been shown to be a sufficient relaxation of strong convexity. They have been proven to hold in diverse settings, such as policy gradient methods. In this talk, we will discuss almost sure convergence rates for SGD (with and without momentum) under global and local gradient domination assumptions. Afterwards, we will apply the results to reinforcement learning, more precisely to softmax parameterized stochastic policy gradient methods.This is joint work with Simon Weissmann and Leif Döring.

3 - Optimal sampling for stochastic and natural gradient descent
Robert Gruhlke, Philipp Trunschke, Anthony Nouy

We consider the problem of optimising the expected value of a loss functional over a nonlinear model class of functions, assuming that we have only access to realisations of the gradient of the loss. This is a classical task in statistics, machine learning and physics-informed machine learning. A straightforward solution is to replace the exact objective with a Monte Carlo estimate before employing standard first-order methods like gradient descent, which yields the classical stochastic gradient descent method. But replacing the true objective with an estimate ensues a “generalisation error”. Rigorous bounds for this error typically require strong compactness and Lipschitz continuity assumptions while providing a very slow decay with sample size. We propose a different optimisation strategy relying on a natural gradient descent in which the true gradient is approximated in local linearisations of the model class via (quasi-)projections based on optimal sampling methods. Under classical assumptions on the loss and the nonlinear model class, we prove that this scheme converges almost surely monotonically to a stationary point of the true objective and we provide convergence rates.

4 - Stochastic gradient methods and tame geometry
Johannes Aspman, Jiri Nemecek, Vyacheslav Kungurtsev, Fabio V. Difonzo, Jakub Marecek

Stochastic differential equations of Langevin-diffusion form have received significant attention, thanks to their foundational role in both Bayesian sampling algorithms and optimization in machine learning. In the latter, they serve as a conceptual model of the stochastic gradient flow in training over-parameterized models. However, the literature typically assumes smoothness of the potential, whose gradient is the drift term. Nevertheless, there are many problems for which the potential function is not continuously differentiable, and hence the drift is not Lipschitz continuous everywhere. This is exemplified by robust losses and Rectified Linear Units in regression problems. In arXiv:2206.11533, we show some foundational results regarding the flow and asymptotic properties of Langevin-type Stochastic Differential Inclusions under assumptions from tame geometry. In arXiv:2302.00709 we extend some of these results for manifold constrained optimization. In arXiv:2311.13544 we further show how to approximate the stratification of tame functions, which makes the above results practical.

■ WF-07

Wednesday, 16:20 - 18:00 - Room: M:I

Regularization methods for Machine Learning and Inverse Problems

Stream: Optimization for Inverse Problems and Machine Learning
Invited session
Chair: *Emanuele Naldi*

1 - On the Bredies-Chenchene-Lorenz-Naldi algorithm
Shambhavi Singh

Monotone inclusion problems occur in many areas of optimization and variational analysis. Splitting methods, which utilize resolvents or proximal mappings of the underlying operators, are often applied to solve these problems. In 2022, Bredies, Chenchene, Lorenz, and Naldi introduced a new elegant algorithmic framework that encompasses various well known algorithms including Douglas-Rachford and Chambolle-Pock. They obtained powerful weak and strong convergence results, where the latter type relies on additional strong monotonicity assumptions. In this paper, we complement the analysis by Bredies et al. by relating the projections of the fixed point sets of the underlying operators that generate the (reduced and original) preconditioned proximal point sequences. We also obtain strong convergence results in the case of linear relations. Various examples are provided to illustrate the applicability of our results.

2 - An optimal structured zeroth-order algorithm for non-smooth optimization
Marco Rando, Cesare Molinari, Lorenzo Rosasco, Silvia Villa

Finite-difference methods are a class of algorithms designed to solve black-box optimization problems by approximating a gradient of the target function on a set of directions. In black-box optimization, the non-smooth setting is particularly relevant since, in practice, differentiability and smoothness assumptions cannot be verified. To cope with nonsmoothness, several authors use a smooth approximation of the target function and show that finite difference methods approximate its gradient. Recently, it has been proved that imposing a structure in the directions allows improving performance. However, only the smooth setting was considered. To close this gap, we introduce and analyze the first structured finite-difference algorithm for non-smooth black-box optimization. Our method exploits a smooth approximation of the target function and we prove that it approximates its gradient on a subset of random orthogonal directions. We analyze the convergence of our procedure under different assumptions. In particular, for non-smooth convex functions, we obtain the optimal complexity. In the non-smooth non-convex setting, we characterize the number of iterations needed to bound the expected norm of the smoothed gradient.

3 - On learning the optimal regularization parameter in inverse problems
Jonathan Chirinos Rodriguez

Selecting the best regularization parameter in inverse problems is a classical and yet challenging problem. Recently, data-driven approaches have become popular to tackle this challenge due to its promising results in many applications. Nevertheless, few theoretical guarantees have been provided up to date.

In this work, we provide a theoretical analysis for this problem by applying classical statistical learning techniques. In particular, we characterize the error performance of this method following an empirical risk minimization approach. We show that, provided with enough data, this approach can reach sharp rates while being essentially adaptive to the noise and smoothness of the problem. Our analysis studies a wide variety of regularization methods, including spectral regularization methods (Tikhonov regularization, Landweber iteration, the -method), non-linear Tikhonov regularization and general convex regularizers such as sparsity inducing norms and Total Variation regularization. Finally, we show some numerical simulations that corroborate and illustrate the theoretical findings.

4 - Adaptive Bregman-Kaczmarz: An Approach to Solve Linear Inverse Problems with Independent Noise Exactly

Lionel Tondji

We consider the block Bregman-Kaczmarz method for finite dimensional linear inverse problems. The block Bregman-Kaczmarz method uses blocks of the linear system and performs iterative steps with these blocks only. We assume a noise model that we call independent noise, i.e. each time the method performs a step for some block, one obtains a noisy sample of the respective part of the right-hand side which is contaminated with new noise that is independent of all previous steps of the method. One can view these noise models as making a fresh noisy measurement of the respective block each time it is used. In this framework, we are able to show that a well-chosen adaptive stepsize of the block Bergman-Kaczmarz method is able to converge to the exact solution of the linear inverse problem. The plain form of this adaptive stepsize relies on unknown quantities (like the Bregman distance to the solution), but we show a way how these quantities can be estimated purely from given data. We illustrate the finding in numerical experiments and confirm that these heuristic estimates lead to effective stepsizes.

Thursday, 8:45 - 9:35

■ TA-01

Thursday, 8:45 - 9:35 - Room: M:A

Plenary II - Amir Beck

Stream: Plenaries
Invited session
Chair: *Giancarlo Bigi*

1 - New results on the multi-dimensional linear discriminant analysis problem
Amir Beck

Fisher linear discriminant analysis (LDA) is a well-known technique for dimensionality reduction and classification. The method was first formulated in 1936 by Fisher. We present three different formulations of the multi-dimensional problem and show why two of them are equivalent. We then prove a rate of convergence of the form $O(qk2))$ for solving the third model. Finally, we consider the max-min LDA problem in which the objective function seeks to maximize the minimum separation among all distances between all classes. We show how this problem can be reduced into a quadratic programming problem with nonconvex polyhedral constraints and describe an effective branch and bound method for its solution. Joint work with Raz Sharon.

Thursday, 10:05 - 11:20

■ TB-02

Thursday, 10:05 - 11:20 - Room: M:O

Solver-based optimization algorithms

Stream: Conic optimization: theory, algorithms and applications
Invited session
Chair: *Yassine Kamri*

1 - Numerical design of optimized first-order methods
Yassine Kamri, Julien Hendrickx, François Glineur

Large-scale optimization is central to many engineering applications such as machine learning, signal processing, and control. First-order optimization methods are a popular choice for solving problems due to their performance and low computational cost. However, in many applications such as machine learning, their performances are highly dependent on the choice of step sizes. Tuning step sizes of first-order methods is a notoriously hard nonconvex problem. In this talk, we present several approaches to optimize the step sizes of first-order methods, relying on recent developments in Performance Estimation Problems. We compare our results to existing optimized step sizes and provide new results, namely numerically optimized step sizes for cyclic coordinate descent and inexact gradient descent, leading to accelerated convergence rate with and without momentum.

2 - Minimization of a sum of pointwise minimum of finite collections of convex functions.
Guillaume Van Dessel, François Glineur

Recently, a few papers tackled the problem of minimizing the sum of a number N of so-called truncated convex functions, i.e. each term of the sum stems as the pointwise minimum between a constant and a convex function. The resulting objective can be expressed as the difference between two convex functions (DC). Hence, celebrated DC algorithms and their variants can be applied. Thanks to a straightforward bi-convex reformulation of the problem, an alternating-minimization (AM) scheme can also provide, hopefully good, local solutions.

In this work, we extend this framework by allowing each term of the sum to be a minimum of a finite number of convex conic representable functions, called components. Accordingly, we propose a family of local methods, dubbed as relaxed alternating minimization (RAM). We provide new lemmas, tailored for the problem at hand, about the inherent concepts of criticality and local optimality and we show, without requiring the strict convexity of the components, that every accumulation point of (RAM) is critical. We show the empirical superiority of (RAM), compared to (AM) and (DC) approaches, on piecewise-linear regression but also multifacility and restricted location problems.

If time allows, we will present our new mixed-integer reformulation for the studied class. In practice, based on the aforementioned lemmas, it can help to produce local optimality certificates or exit directions from neighbourhoods of critical points when using (RAM).

■ TB-03

Thursday, 10:05 - 11:20 - Room: M:J

In memory of Georg Still - part 1

Stream: In memory of Georg Still
Invited session
Chair: *Oliver Stein*

1 - On the weakest constraint qualification for sharp local minimizers
Oliver Stein, Maximilian Volk

A result of Georg Still and Martin Streng characterizes the sharp local minimality of feasible points of nonlinear optimization problems by a strengthened version of the Karush-Kuhn-Tucker conditions, as long as the Mangasarian-Fromovitz constraint qualification holds. This strengthened condition is not easy to check algorithmically since it involves the topological interior of some set. In this article, we derive an algorithmically tractable version of this condition, called strong Karush-Kuhn-Tucker condition. We show that the Guignard constraint qualification is the weakest condition under which a feasible point is a strong Karush-Kuhn-Tucker point for every continuously differentiable objective function possessing the point as a sharp local minimizer. As an application, our results yield an algebraic characterization of strict local minimizers of linear programs with cardinality constraints.

2 - Copositive and semi-infinite optimization
Mirjam Duer

A copositive optimization problem is a problem in matrix variables involving the constraint that the matrix should lie in the copositive cone. Copositive programming can be regarded as a special instance of linear semi-infinite programming. This talk will study these relations and highlight the contributions Georg Still made in this field.

3 - A reflection on the work of Georg Still on semi-infinite linear optimization
Etienne De Klerk

In this talk I will reflect on the work of Georg Still in the area of semi-infinite linear programming, where he made notable contributions. In particular, I will discuss the application in Chebyshev approximation, and some recent trends in this area.

■ TB-04

Thursday, 10:05 - 11:20 - Room: M:M

Optimization in regression, classification and learning II

Stream: Optimization in regression, classification and learning
Invited session
Chair: *Paula Amaral*

1 - Memetic Differential Evolution Methods for Semi-Supervised Clustering

Pierluigi Mansueto, Fabio Schoen

In this work, we deal with semi-supervised Minimum Sum-of-Squares Clustering (MSSC) problems where background knowledge is given in the form of instance-level constraints. In particular, we take into account "must-link" and "cannot-link" constraints, each of which indicates if two dataset points should be associated to the same or to a different cluster. The presence of such constraints makes the problem at least as hard as its unsupervised version: it is no more true that each point is associated to its nearest cluster center, thus requiring some modifications in crucial operations, such as the assignment step. In this scenario, we propose novel memetic methodologies based on the Differential Evolution (DE) paradigm and designed to return a (hopefully) optimal solution satysfing all the constraints. The proposals directly extend state-of-the-art DE-based memetic approaches recently proposed for the unsupervised scenario. The new algorithms are compared with some state-of-the-art algorithms from the literature on a set of well-known datasets, highlighting their effectiveness and efficiency in finding good quality clustering solutions.

2 - Solving Kernel Ridge Regression with Gradient Descent for a Non-Constant Kernel

Oskar Allerbo

Kernel ridge regression, KRR, is a generalization of linear ridge regression that is non-linear in the data, but linear in the parameters. The solution can be obtained either as a closed-form solution, which includes a matrix inversion, or iteratively through gradient descent. Using the iterative approach opens up for changing the kernel during training, something that is investigated in this paper. We theoretically address the effects this has on model complexity and generalization. Based on our findings, we propose an update scheme for the bandwidth of translational-invariant kernels, where we let the bandwidth decrease to zero during training, thus circumventing the need for hyper-parameter selection. We demonstrate on real and synthetic data how decreasing the bandwidth during training outperforms using a constant bandwidth, selected by cross-validation and marginal likelihood maximization. We also show theoretically and empirically that using a decreasing bandwidth, we are able to achieve both zero training error in combination with good generalization, and a double descent behavior, phenomena that do not occur for KRR with constant bandwidth but are known to appear for neural networks.

3 - Matrix-wise L0-constrained Sparse Nonnegative Least Squares

Nicolas Nadisic, Arnaud Vandaele, Jeremy Cohen, Nicolas Gillis

Nonnegative least squares problems with multiple right-hand sides (MNNLS) arise in models that rely on additive linear combinations. In particular, they are at the core of most nonnegative matrix factorization algorithms and have many applications. The nonnegativity constraint is known to naturally favor sparsity, that is, solutions with few non-zero entries. However, it is often useful to further enhance this sparsity, as it improves the interpretability of the results and helps reducing noise, which leads to the sparse MNNLS problem. In this paper, as opposed to most previous works that enforce sparsity column- or row-wise, we first introduce a novel formulation for sparse MNNLS, with a matrix-wise sparsity constraint. Then, we present a two-step algorithm to tackle this problem. The first step divides sparse MNNLS in subproblems, one per column of the original problem. It then uses different algorithms to produce, either exactly or approximately, a Pareto front for each subproblem, that is, to produce a set of solutions representing different tradeoffs between reconstruction error and sparsity. The second step selects solutions among these Pareto fronts in order to build a sparsity-constrained matrix that minimizes the reconstruction error. We perform experiments on facial and hyperspectral images, and we show that our proposed two-step approach provides more accurate results than state-of-the-art sparse coding heuristics applied both column-wise and globally.

■ TB-05

Thursday, 10:05 - 11:20 - Room: M:N

Optimization for learning III

Stream: Optimization for learning
Invited session
Chair: *Max Nilsson*

1 - MAST: Model-Agnostic Sparsified Training

Egor Shulgin, Peter Richtarik

We introduce a novel optimization problem formulation that departs from the conventional way of minimizing machine learning model loss as a black-box function. Unlike traditional formulations, the proposed approach explicitly incorporates an initially pre-trained model and random sketch operators, allowing for sparsification of both the model and gradient during training. We establish the insightful properties of the proposed objective function and highlight its connections to the standard formulation. Furthermore, we present several variants of the Stochastic Gradient Descent (SGD) method adapted to the new problem formulation, including SGD with general sampling, a distributed version, and SGD with variance reduction techniques. We achieve tighter convergence rates and relax assumptions, bridging the gap between theoretical principles and practical applications, covering several important techniques such as Dropout and Sparse training. This work presents promising opportunities to enhance the theoretical understanding of model training through a sparsification-aware optimization approach.

2 - Online Learning and Information Exponents: The Importance of Batch size & Time/Complexity Tradeoffs
Stephan Ludovic

We study the impact of the batch size on the iteration time of training two-layer neural networks with one-pass stochastic gradient descent (SGD) on multi-index target functions of isotropic covariates. We characterize the optimal batch size minimizing the iteration time as a function of the hardness of the target, as characterized by the information exponents. We show that performing gradient updates with large batches minimize the training time without changing the total sample complexity. However, larger batch sizes are detrimental for improving the time complexity of SGD. We provably overcome this fundamental limitation via a different training protocol, Correlation loss SGD, which suppresses the auto-correlation terms in the loss function. We show that one can track the training progress by a system of low dimensional ordinary differential equations (ODEs). Finally, we validate our theoretical results with numerical experiments.

3 - Compressed and distributed least-squares regression: convergence rates with applications to Federated Learning
Constantin Philippenko, Aymeric Dieuleveut

In this paper, we investigate the impact of compression on stochastic gradient algorithms for machine learning, a technique widely used in distributed and federated learning. We underline differences in terms of convergence rates between several unbiased compression operators, that all satisfy the same condition on their variance, thus going beyond the classical worst-case analysis. To do so, we focus on the case of least-squares regression (LSR) and analyze a general stochastic approximation algorithm for minimizing quadratic functions relying on a random field. More particularly, we highlight the impact on the convergence of the covariance of the additive noise induced by the algorithm. We consider weak assumptions on the random field, tailored to the analysis (specifically, expected Hölder regularity), and on the noise covariance, enabling the analysis of various randomizing mechanisms, including compression. We then extend our results to the case of federated learning.

■ TB-06

Thursday, 10:05 - 11:20 - Room: M:H

Advances in Semi-Definite Programming

Stream: Challenges in nonlinear programming
Invited session
Chair: *Pavel Dvurechensky*

1 - A conditional gradient homotopy method with applications to Semidefinite Programming
Mathias Staudigl, Pavel Dvurechensky, Shimrit Shtern

We propose a new homotopy-based conditional gradient method for solving convex opti- mization problems with a large number of simple conic constraints. Instances of this template naturally appear in semidefinite programming problems arising as convex relaxations of com- binatorial optimization problems. Our method is a double-loop algorithm in which the conic constraint is treated via a self-concordant barrier, and the inner loop employs a conditional gradient algorithm to approximate the analytic central path, while the outer loop updates the accuracy imposed on the temporal solution and the homotopy parameter. Our theoretical it- eration complexity is competitive when confronted to state-of-the-art SDP solvers, with the decisive advantage of cheap projection-free subroutines. Preliminary numerical experiments are provided for illustrating the practical performance of the method.

2 - SDP Relaxations for Training ReLU Activation Neural Networks
Karthik Prakhya, Tolga Birdal, Alp Yurtsever

Solving non-convex optimization problems is crucial for training machine learning models, including neural networks. However, non-convexity often leads to less reliable and less robust neural networks with unclear inner workings. While convex formulations have been used for verifying neural network robustness, their application to training neural networks remains relatively unexplored. In this work, we propose a semidefinite programming relaxation for training two-layer ReLU networks in a lifted space, which can be solved in polynomial time. Numerical experiments demonstrate that this SDP formulation provides reasonably tight lower bounds on the training objective across various prediction and classification tasks.

3 - Learning quantum Hamiltonians at any temperature in polynomial time with Chebyshev and bit complexity
Ales Wodecki, Jakub Marecek

We consider the problem of learning local quantum Hamiltonians given copies of their Gibbs state at a known inverse temperature, following Haah et al. [2108.04842] and Bakshi et al. [arXiv:2310.02243]. Our main technical contribution is a new flat polynomial approximation of the exponential function based on the Chebyshev expansion, which enables the formulation of learning quantum Hamiltonians as a polynomial optimization problem. This, in turn, can benefit from the use of moment/SOS relaxations, whose polynomial bit complexity requires careful analysis [O'Donnell, ITCS 2017]. Finally, we show that learning a k-local Hamiltonian, whose dual interaction graph is of bounded degree, runs in polynomial time under mild assumptions.

■ TB-07

Thursday, 10:05 - 11:20 - Room: M:I

Global optimization I

Stream: Global optimization
Invited session
Chair: *Sonia Cafieri*

1 - Improved branch-and-bound for integer D-Optimality

Jon Lee

We develop a branch-and-bound algorithm for the integer D-optimality problem, a central problem in statistical design theory, based on two convex relaxations, employing variable-bound tightening and fast local-search procedures, testing our ideas on randomly-generated test problems. We are able to prove some relationships between various bounds for the related data-fusion problem.

2 - On Model Generation and Decomposition for Global Optimization and Machine Learning using the Generate-and-Solve Approach

Ivo Nowak

We present decomposition-based generate-and-solve methods for generating reformulations of complex optimization problems that can be solved easier than the original problem. The methods are implemented in the open-source frameworks Decogo and Decolearn. Numerical results for nonconvex MINLPs and machine learning problems are presented.

3 - Complex geometrical test for optimality conditions in Interval Branch and Bound method

Boglárka G.-Tóth, Mihály Gencsi

This study focuses on solving constrained nonlinear programming problems. The Interval Branch and Bound (IBB) method is the most widely used approach for obtaining rigorous solutions. However, few IBB implementations utilize the Karush-Kuhn-Tucker or Fritz-John optimality conditions to eliminate non-optimal boxes. The Fritz-John optimality condition involves solving an interval-valued system of equations. When solving these equations, we sometimes cannot reduce the size of the box due to overestimation of the gradient boxes.

This study considers the optimality conditions from a geometric perspective. A new, more complex geometrical optimality test is introduced to precede the Fritz-John optimality condition. The goal is to speed up the IBB method and eliminate unnecessary computations.

The test easily ensures that there is no local optimum in the box, or that we cannot reduce the size of the box by solving the optimality condition system because of the overestimation of the gradient boxes. We will describe the geometric test and its usefulness, as well as present experimental results demonstrating the effectiveness of the complex geometrical test.

Thursday, 11:25 - 12:40

■ TC-02

Thursday, 11:25 - 12:40 - Room: M:O

Developments in interior point methods

Stream: Developments in interior point methods
Invited session
Chair: *Martin Skovgaard Andersen*
Chair: *Utkarsh Detha*

1 - Pushing the limits of interior methods for nonlinear optimization
Anders Forsgren, Pim Heeman

Interior methods form a powerful class of methods for solving nonlinear optimization problems. The main computational work lies in solving the linear equations that arise when Newton's method is applied to solving a set of nonlinear equations. The nonlinear equations may be of primal form, setting the gradient of the log barrier function to zero, or of primal-dual form, formulated as a perturbation of the first-order necessary optimality conditions.

The ability to solve large problems is therefore tied to the ability to solve the large systems of linear equations. An important feature is that the optimal solution to the optimization problem is what is desired, neither the optimal solution to a nonlinear equation for a given value of the barrier parameter nor the solution to the particular linear system of equations. There will be a tradeoff between the accuracy in the solution of the linear equations and the overall computational cost for solving the optimization problem.

In the talk, past and ongoing work in this area, in which the speaker has been involved, will be discussed.

2 - Clarabel: An interior point solver for conic optimization
Yuwen Chen

The homogeneous self-dual embedding (HSDE) is a popular technique that unifies optimality and infeasibility detection of a convex problem and has been widely used in convex interior-point solver. However, it only allows a linear cost in the objective and we need to eliminate the quadratic term in the objective function, replacing it with an epigraphical upper bound and an additional second-order cone constraint in the objective.

We presented the interior-point solver, Clarabel, for conic optimization with quadratic objectives by using the homogeneous embedding for infeasibility detection, which tackled the issue of transforming QPs into SOCPs. The solver is fully written in both Julia and Rust languages. We also support second-order cones, exponential cones, power cones and positive definite cones in Clarabel. Our solver is competitive with the cutting-edge conic solvers on various benchmarks and outperforms them in terms of time and numerical stability on QPs and SOCPs.

3 - A low-rank error correction technique for improving approximate factorisation preconditioners
Andreas Bock

Interior point methods (IPM) require the solution of linear systems of equations for the Newton directions. These linear systems become increasingly ill-conditioned as the IPM progresses, so preconditioning is mandatory when using the conjugate gradient method within an inexact Newton method. We propose a preconditioner based on nearness in the Bregman log determinant divergence. The form of the proposed preconditioner is given as the sum of a user-specified approximate factorisation (e.g. incomplete Cholesky) plus a low-rank correction term. The latter is based on the (in general indefinite) approximate factorisation error, and is generally different from one obtained by a truncated singular value decomposition (TSVD). This gives rise to a new type of truncation which we compare with the TSVD in terms of convergence of the preconditioned conjugate gradient method (PCG). We show numerous examples where PCG converges to a small tolerance using the proposed preconditioner, whereas PCG using a TSVD-based preconditioner fails. We also consider SuiteSparse matrices from IPM from linear programming that do not admit an incomplete Cholesky factorisation by default, and present a robust incomplete Cholesky preconditioner based on the proposed methodology.

■ TC-03

Thursday, 11:25 - 12:40 - Room: M:J

In memory of Georg Still - part 2

Stream: In memory of Georg Still
Invited session
Chair: *Mirjam Duer*

1 - A generic analysis for multi-leader-disjoint followers games: bilevel problems first
Gemayqzel Bouza Allende

Multi-leader-follower Games are models which combine bilevel programming with generalized Nash games. They have attracted more and more attentionover the past few years due to their application in situations where there is competition and hierarchical relations among the individuals. In this contribution, we focus on the case in which several leaders are acting according to a generalized Nash game and each leader has a number of exclusive followers who interact through a standard Nash game between them. These models are a generalization of bilevel and single leader multi-follower games. In this contribution, the generic properties of their solutions are discussed. It should be pointed out that this analysis for bilevel problems was done in [2] and settled foundations of [1]. We also show that these properties will stay stable under small perturbations of the involved functions.

[1] D. Aussel, G. Bouza, S. Dempe and S. Lepaul: Genericity Analysis of Multi-Leader-Disjoint-Followers Game, SIAM Journal on Optimization, 31, 3, (2021), 2055-2079.

[2] G. Bouza-Allende and G. Still, Solving bilevel programs with the KKT-approach, Math. Program., 138 (2013), 309–332.

2 - Genericity and stability in linear conic programming
Bolor Jargalsaikhan, Mirjam Duer, Georg Still

In linear conic programming, we maximize or minimize a linear function over the intersection of an affine space and a convex cone. In this talk, we discuss properties of conic problems such as uniqueness of the optimal solution, nondegeneracy, and strict complementarity. A property is said to be stable at a problem instance if the property still holds under a small perturbation of the problem data. We say that a property is weakly generic if the property holds for almost all problem instances. We start by showing that Slater's condition is weakly generic and stable. It is known that uniqueness of the optimal solution, nondegeneracy, strict complementarity are weakly generic properties in conic programming, so we study the stability of these weakly generic properties. For the semidefinite programs, we show that all these properties are stable. Moreover, we characterize first order optimal solutions in conic programs and give necessary and sufficient conditions for their stability.

3 - Two methods for the maximization of homogeneous polynomials over the simplex
Faizan Ahmed

The paper deals with the numerical solution of the problem P to maximize a homogeneous polynomial over the unit simplex. We discuss the convergence properties of the so-called replicator dynamics for solving P. We further examine an ascent method, which also makes use of the replicator transformation. Numerical experiments with polynomials of different degrees illustrate the theoretical convergence results.

Note: This was the last paper that Georg and I have written. We finished this paper during the COVID time. During the presentation, I will share memories of working with Georg since 2009 as a master's student, a PhD student, and then collaborating with him on writing a few more papers.

■ TC-04

Thursday, 11:25 - 12:40 - Room: M:M

Multiobjective Optimization II

Stream: Multiobjective optimization
Invited session
Chair: *Firdevs Ulus*

1 - Adapting the DMulti-MADS algorithm to mixed-integer multiobjective derivative-free optimization
Ludovic Salomon, Sébastien Le Digabel, Christophe Tribes

The DMulti-MADS method is an extension of the Mesh Adaptive Direct Search (MADS) algorithm for multiobjective derivative-free optimization. It is convergence-based and at the same time has shown good experimental performance. DMulti-MADS was originally designed for continuous variables only. However, many "real-world" engineering applications also have integer variables that need to be considered. In this talk, we describe a simple adaptation of the DMulti-MADS algorithm to consider both continuous and integer variables. Numerical experiments on artificial benchmarks and real-world problems are performed against state-of-the-art algorithms.

2 - Numerical tests of a global solver for optimistic semivectorial bilevel problems
Daniel Hoff, Gabriele Eichfelder

In semivectorial bilevel optimization one considers an optimization problem, the upper-level problem, which contains the weakly efficient set of a multiobjective optimization problem, the lower-level problem, as a constraint. We replace the lower-level weakly efficient set by using its optimal value function and present a branch-and-bound method for optimistic semivectorial bilevel problems. This method improves suitable approximations of the value function until the lower and upper bound on the minimal value are sufficiently close. In this talk the focus is on numerical tests.

3 - Weight Space Decomposition for Multiobjective Linear Programming in the Context of Equitable Optimization
Firdevs Ulus, Ozlem Karsu

We consider equitable linear optimization problems (ELOP), which are multiobjective optimization problems, with each objective representing the benefit that one entity receives. In such problems, the concept of dominance is replaced by equitable dominance. The aim is to find the set of equitably nondominated points, which can be done by solving ordered weighted averaging (OWA) scalarization. Each solution corresponds to a different set of weights that make the solution optimal, hence a different degree of inequity aversion. We discuss a novel use of the parametric simplex algorithm to address ELOP. The algorithm not only provides the set of equitably nondominated solutions but also yields the corresponding weight space decomposition. We also propose an alternative method based on geometric duality to compute the weight space decomposition given the set of nondominated solutions.

■ TC-05

Thursday, 11:25 - 12:40 - Room: M:N

Recent advances in bilevel optimization I

Stream: Bilevel optimization: strategies for complex decision-making
Invited session
Chair: *Alp Yurtsever*

1 - Computing synthetic controls using bilevel optimization
Pekka Malo

The synthetic control method (SCM) represents a notable innovation in estimating the causal effects of policy interventions and programs in a comparative case study setting. In this paper, we demonstrate that the data-driven approach to SCM requires solving a bilevel optimization problem. We show how the original SCM problem can be solved to the global optimum through the introduction of an iterative algorithm rooted in Tykhonov regularization or Karush-Kuhn-Tucker approximations.

2 - Smooth over-parametrized solvers for sparse optimisation
Clarice Poon

Non-smooth regularization is a core ingredient of many imaging or machine learning pipelines. Non-smoothness encodes structural constraints on the solutions, such as sparsity, group sparsity, low-rank and sharp edges. It is also the basis for the definition of robust loss functions such as the square-root lasso. Standard approaches to deal with non-smoothness leverage either proximal splitting or coordinate descent. The effectiveness of their usage typically depend on proper parameter tuning, preconditioning or some sort of support pruning. In this work, we advocate and study a different route. By over-parameterization and marginalising on certain variables (Variable Projection), we show how many popular non-smooth structured problems can be written as smooth optimization problems. The result is that one can then take advantage of quasi-Newton solvers such as L-BFGS and this, in practice, can lead to substantial performance gains. Another interesting aspect of our proposed solver is its efficiency when handling imaging problems that arise from fine discretizations (unlike proximal methods such as ISTA whose convergence is known to have exponential dependency on dimension). On a theoretical level, one can connect gradient descent on our over-parameterized formulation with mirror descent with a varying Hessian metric. This observation can then be used to derive dimension free convergence bounds and explains the efficiency of our method in the fine-grids regime.

■ TC-06

Thursday, 11:25 - 12:40 - Room: M:H

Stochastic methods and applications

Stream: Methods for non-/monotone inclusions and their applications
Invited session
Chair: *Feifei Hu*

1 - Optimization for a car sharing systems using a queueing network with correlated arrival flows and moving servers
Chesoong Kim

Car sharing services have quickly developed for client transportation in many countries. The use of well managed car sharing services is profitable for individual. In addition, the social benefits are provided via efficient use of parking lots, increase throughput of the roads, reduction of a traffic jam and carbon emission. We construct the mathematical model of car sharing as an open queueing network. The expressions for calculating the main performance measures of the network are presented and numerical examples to solve optimization problem based on the obtained results is considered.

2 - Stochastic optimal power flow problem using interior point methods with block structure
Aurelio Oliveira, Catalina Jaramillo Villalba

This work focuses on optimizing power flow in the electrical sector, considering uncertainties in energy demand, and aiming to minimize costs and energy loss within the constraints of the physical network. We use a two-stage stochastic programming building discrete scenarios associated to the uncertainty of demand. The problems are solved using specialized software employing interior point methods for stochastic programming. A mathematical formulation using splitting variables is presented and computational experiments, comparing the new approach with a classical one are performed. The findings indicate that the proposed method yields solutions consistent with other softwares, showing notable improvements in terms of processing time and the number of iterations.

3 - Gradient Descent for Noisy Optimization
Feifei Hu

We study the use of gradient descent with backtracking line search (GD-BLS) to solve the noisy optimization problem, assuming that the objective function is strictly convex. Without making any global assumptions on the objective function and without assuming that the variance of the stochastic gradient is bounded, we show that GD-BLS allows to estimate the minimiser of the objective function with an error bounded by some value proportional to the computational budget to the power of -0.25. We show that we can improve this rate by stopping the optimization process earlier when the gradient of the objective function is sufficiently close to zero and then using the residual computational budget to optimize with GD-BLS a finer approximation of the function. We also show a more general convergence result for GD-BLS under some certain assumption. Beyond knowing one parameter, achieving the aforementioned convergence rates do not require to tune the algorithms parameters according to the specific objective functions at hand, and we exhibit a simple noisy optimization problem for which stochastic gradient is not guaranteed to converge while the algorithms discussed in this work are.

■ TC-07

Thursday, 11:25 - 12:40 - Room: M:I

Global optimization II / MINLP

Stream: Global optimization
Invited session
Chair: *Sonia Cafieri*

1 - Solving MINLPs to Global Optimality with FICO Xpress Solver
Tristan Gally

This talk introduces the global optimization capability within FICO Xpress Solver, which allows to solve general mixed-integer nonlinear problems to global optimality. We will discuss how existing features of the MILP and local NLP solvers are used and explain several different extensions for, e.g., spatial branching and convexification cuts, and their performance implications.

2 - Global optimization of continuous and discrete nonlinear programs with BARON
Yi Zhang, Nikolaos Sahinidis

BARON, under development since the early 1990s, has evolved into a robust computational system for solving continuous and discrete optimization problems to global optimality. The current generation of BARON focuses on enhancements in hybrid relaxations, presolve methods, convexification techniques, heuristics, and robustness improvements. Our experiments underscore the importance of all algorithmic components within the branch-and-reduce framework. Computational comparisons against state-of-the-art global and local solvers highlight BARON's superior performance.

3 - On computing upper bounds in nonlinear problems involving disjunctive constraints
Sonia Cafieri, Marcel Mongeau

The recently introduced continuous quadrant penalty formulation of logical disjunctive constraints constitutes a continuous-optimization alternative to the classical formulations (such as the bigM formulation) based on the introduction of binary variables. We focus on nonlinear problems whose only combinatorial aspect comes from their logical constraints, and build on the continuous quadrant penalty, that yields to a continuous nonconvex problem, to derive an efficient computation of upper bounds to be used within Branch-and-Bound-based multi-step approaches. We show on two problems, respectively from discrete geometry and arising in air traffic management optimization, that our approach is effective at speeding up the computational convergence.

Thursday, 14:10 - 15:50

■ TD-02

Thursday, 14:10 - 15:50 - Room: M:O

Recent advances in computer-aided analyses of optimization algorithms III

Stream: Conic optimization: theory, algorithms and applications
Invited session
Chair: *Adrien Taylor*
Chair: *Manu Upadhyaya*

1 - Provable non-accelerations of the heavy-ball method
Aymeric Dieuleveut, Adrien Taylor, Baptiste Goujaud

We show that the heavy-ball (HB) method provably does not reach an accelerated convergence rate on smooth strongly convex problems. More specifically, we show that for any condition number and any choice of algorithmic parameters, either the worst-case convergence rate of HB on the class of L-smooth and μ-strongly convex quadratic functions is not accelerated (that is, slower than 1 O()), 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 class of functions that also satisfy higher-order regularity conditions.

2 - Provable non-accelerations of the heavy-ball method
Baptiste Goujaud, Adrien Taylor, Aymeric Dieuleveut

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

3 - Exact convergence rates of the last iterate in subgradient methods
François Glineur, Moslem Zamani

We study the convergence of the last iterate in subgradient methods applied to the minimization of a nonsmooth convex function with bounded subgradients.

We first introduce a proof technique that generalizes the standard analysis of subgradient methods. It is based on tracking the distance between the current iterate and a different reference point at each iteration. Using this technique, we obtain the exact worst-case convergence rate for the objective accuracy of the last iterate of the projected subgradient method with either constant step sizes or constant step lengths. Tightness is shown with a worst-case instance matching the established convergence rate. We also derive the value of the optimal constant step size when performing a given number of iterations.

Finally, we introduce a new optimal subgradient method that achieves the best possible last-iterate accuracy after a given number of iterations. Its convergence rate matches exactly the lower bound on the performance of any black-box method on the considered problem class. We also show that no single memory-less method can achieve this optimal rate for any number of iterations.

[paper can be found at https://arxiv.org/abs/2307.11134]

4 - Algorithms with learned deviations
Sebastian Banert

Deviations are a way to modify existing optimization algorithms. In contrast to, e.g., momentum terms parametrized by a single number, they can have the same dimensions as the optimization variables. Traditionally interpreted as unwanted errors, they can be used to tailor an algorithm to a specific class of problems. By choosing the deviations small enough, convergence rates can be guaranteed. Our bounds for the norms of the deviations are based on known quantities, not summability. The performance of specialized algorithms for certain problem classes can be improved by training the deviations with deep learning. This talk is based on joint work with Jevgenija Rudzusika, Ozan Öktem, Jonas Adler, Hamed Sadeghi, and Pontus Giselsson.

■ TD-03

Thursday, 14:10 - 15:50 - Room: M:J

Variational techniques and subdifferentials

Stream: Variational analysis: theory and algorithms
Invited session
Chair: *Abderrahim Hantoute*

1 - Lagrange duality on DC evenly convex optimization problems via a generalized conjugation scheme
Maria Dolores Fajardo, Jose Vidal-Nunez

In this talk, we present how Lagrange duality is connected to optimization problems whose objective function is the difference of two convex functions, briefly called DC problems. We enter two Lagrange dual problems, each of them obtained via a different approach. While one of the duals corresponds to the standard formulation of the Lagrange dual problem, the other is written in terms of conjugate functions. When one of the involved functions in the objective is evenly convex, both problems are equivalent, but this relation is no longer true in the general setting. For this reason, we show conditions ensuring not only weak, but also zero duality and strong duality between the primal and one of the dual problems written using conjugate functions. For the other dual, and due to the fact that weak duality holds by construction, we just develop conditions for zero duality gap and strong duality between the primal DC problem and its (standard) Lagrange dual problem. We apply also the obtained results to characterize weak and strong duality together with zero duality gap between the primal problem and its Fenchel-Lagrange dual.

2 - Smoothing Effect of Epi-convergence and Inf-convolution in Optimization
Mohammed Ibrahim Abdelkayoum Ghitri, Abderrahim Hantoute

We use smoothing processes based on the infimal convolution of convex, proper and lower semi-continuous functions to regularise optimization problems given by convex-composite functions. We show that the proposed regularisation schemes still (epi-)converge to the original data, even if the chosen kernel is an arbitrary convex function. This also allows to derive upper estimates of the subdifferentials of the epi-limits of non-convex functions without using qualifying conditions (namely BCQ-type conditions).

3 - Robust and continuous metric subregularity in the radius of stability context
Jesús Camacho

In this talk we introduce the concepts of robust and continuous metric subregularity. Two new variational properties of the feasible set mapping in the (finite) linear inequality systems setup. The motivation of this study goes back to the seminal work by Dontchev, Lewis, and Rockafellar (2003) on the radius of metric regularity. The implications of the unstable continuity behavior of the metric subregularity modulus are reflected in those definitions. A deep study on the stability of the end-set leads to characterizations of both properties. While a explicit formula is given for the radius of robust metric subregularity, only partial progress has been made concerning the radius of continuous metric subregularity.

4 - Discretization and reduction in finite and infinite convex optimization
Abderrahim Hantoute

We analyze the reduction of the number of constraints in the finite and infinite convex optimization framework, while maintaining the same optimal value. This approach allows a considerable simplification of the study of the duality of such problems under reduced constraints qualifications. We illustrate our result by an explicit characterization of the normal cone to the (effective) domain of a supremum function, necessary to provide more precise optimality conditions.

■ TD-04

Thursday, 14:10 - 15:50 - Room: M:M

Multiobjective Optimization III

Stream: Multiobjective optimization
Invited session
Chair: *Gabriela Kovacova*

1 - Circuits in a Box: Multi-Objective Optimization for Analog Electronic Circuits
Christopher Schneider

We present a method for approximating so-called performance spaces of analog electronic circuits. Performance spaces as representations for attainable image spaces of analog electronic circuits are well-suited for supporting their conventional design process. The design task can be formulated as non-convex multi-objective optimization problem.

Our approach is based on a box-coverage method for Pareto optimization with modifications to get an approximation of the complete image space of the problem. The thereby resulting models provide accurate information about all valid compromise solutions of the competing circuit performances.

Real world examples demonstrate how this approach can help analog system designers to make reliable topology decisions, even without detailed knowledge about the underlying circuits. Furthermore, we present an interactive visualization technique to explore performance spaces of dimension four and higher.

2 - Robust multi-objective stochastic control
Gabriela Kovacova, Igor Cialenco

Model uncertainty is relevant for various dynamic optimization problems. As one example, let us mention the portfolio selection problem – an investor does not know the true distribution of asset return on the market. Various approaches to handling uncertain stochastic control problems have been developed, among them the robust approach with the aim of optimizing under the worst-case scenario.

While model uncertainty and robust optimization are relatively well understood for standard control problems with a single (scalar) objective, this is much less the case for problems with multiple objectives. In recent years, several (dynamic) problems of financial mathematics have been approached through methods of multi-objective and set optimization. Set-valued Bellman's principle, a version of the well known Bellman's principle for problems with multiple or set-valued objectives, has been derived across different problems.

In this work we explore the robust approach to model uncertainty for multi-objective stochastic control problems. Robust multi-objective optimization has been explored in the static but not in the dynamic setting. We are particularly interested in the application of dynamic programming and the impact model uncertainty has on the set-valued Bellman's principle. We show how the set-valued Bellman's principle is replace by certain set relations (or inclusions) under robustness and present assumptions under which equality can be obtained.

3 - Playing the Budget-Constrained Multi-Battle Contest with Randomized Strategies for Maximizing Winning Probabilities
Chien-Hsin Chen, Po-An Chen, Wing-Kai Hon

The result of our previous work on budget-constrained multi-battle contests is basically the optimal budget ratio that guarantees the first-mover to win against an adversarial follower and the corresponding bidding strategy. One may wonder what a player should do when their initial budget is not high enough. Our model and results in this paper provide a more general way to discuss what a player with only a relatively moderate amount of budget would do: our goal is to derive "the leader's mixed strategy that maximizes the winning probability for the whole game, under each rule for tie-breaking", assuming the follower plays adversarially. Moreover, we express such a maximum winning probability in terms of the amount of extra money.

In particular, the winning probabilities when several basic and more complicated rules are applied will "not" be affected by the role of players; only which game rule is used and the amount of extra money of the rich player matters. When the reciprocal of the extra amount of money is integral and the most complicated rule is applied, instead of giving a closed-form solution, the winning probability can be characterized by a recursive formula.

■ TD-05

Thursday, 14:10 - 15:50 - Room: M:N

Nonsmooth optimization algorithms

Stream: Nonsmooth and nonconvex optimization algorithms
Invited session
Chair: *Susan Ghaderi*

1 - A feasible directions method for nonconvex optimization over linear constraints with a nonsmooth concave regularizer
Nadav Hallak, Amir Beck

This talk presents a feasible directions approach for the minimization of a continuous function over linear constraints in which the update directions belong to a predetermined finite set spanning the feasible set. These directions are recurrently investigated in a cyclic semi-random order, where the stepsize of the update is determined via univariate optimization. The method achieves that any accumulation point is a stationary point, and enjoys a sublinear rate of convergence in expectation w.r.t a new optimality measure that acts as a proxy for the stationarity condition.

2 - Spectral and nuclear norms of order three tensors: Complexity and computation
Zhening Li

The recent decade has witnessed a surge of research in modeling and computing from twoway data (matrices) to multiway data (tensors). However, there is a drastic phase transition for most tensor optimization and computation problems when the order of a tensor increases from two to three: Most tensor problems are NP-hard while that for matrices are easy. It triggers a question on where exactly the transition occurs. The work aims to study the kind of question for the spectral norm and the nuclear norm. Although computing the spectral norm for a general ×m×n tensor is NP-hard, we show that it can be computed in polynomial time if is fixed. In another word, the best rank-one approximation of a general ×m×n tensor can be computed in polynomial time for fixed . This is the same for the nuclear norm. While these polynomial-time methods are not currently implementable in practice, we propose fully polynomial-time approximation schemes (FPTAS) for the spectral norm based on polytope approximation and for the nuclear norm with further help of duality theory and semidefinite optimization. Numerical experiments show that our FPTAS can compute these tensor norms for small but large m and n. To the best of our knowledge, this is the first method that can exactly compute the nuclear norm of general asymmetric tensors. Both our polynomial-time algorithms and FPTAS can be extended to higher-order tensors as well.

3 - High-order Moreau envelope in nonsmooth convex setting: L-smoothness and inexact gradient method
Alireza Kabgani, Masoud Ahookhosh

In this talk, we revisit the differentiability and Lipschitz continuity properties of the gradient of the high-order Moreau envelope for nonsmooth convex functions. Leveraging these insights, we propose an inexact gradient method and explore its convergence under mild conditions. Additionally, we establish the global convergence of the method by leveraging the Kurdyka-Lojasiewicz inequality. We illustrate the effectiveness of our approach through both theoretical analyses and numerical results.

4 - Scaled gradient methods under convexity and local smoothness
Susan Ghaderi, Yves Moreau, Masoud Ahookhosh

This paper explores the unconstrained minimization of smooth convex functions with locally Lipschitz gradients, focusing on the Scaled Gradient Method (SGA) and its adaptive version, AdaSGA. We analyze their convergence, complexity under local smoothness, and effectiveness through numerical experiments. Our findings, particularly applied to Poisson linear inverse problems, showcase the potential of these methods in optimization and computational applications.

■ TD-06

Thursday, 14:10 - 15:50 - Room: M:H

Stochastic methods

Stream: Methods for non-/monotone inclusions and their applications
Invited session
Chair: *Tony Stillfjord*

1 - Almost sure convergence of stochastic Hamiltonian descent methods

Måns Williamson

Gradient normalization and soft clipping are two popular techniques for tackling instability issues and improving convergence of stochastic optimization methods. In this talk, we study these types of methods through the lens of dissipative Hamiltonian systems. Gradient normalization and certain types of soft clipping algorithms can be seen as (stochastic) implicit-explicit Euler discretizations of dissipative Hamiltonian systems, where the kinetic energy function determines the type of clipping that is applied. We make use of unified theory from dynamical systems to show that all of these schemes converge almost surely to stationary points of the objective function.

2 - A gradient-free optimisation algorithm

Ettore Fincato

We develop a gradient-free algorithm for the minimisation of strongly lower-semicontinuous, lower-bounded functions. The definition of the algorithm does not require differentiability or convexity of the target function, and the implementation can be carried out by simple or sequential Monte Carlo procedures. We develop supporting theory, prove convergence of the algorithm and provide numerical experiments.

3 - Analysis of a class of stochastic component-wise soft-clipping schemes

Tony Stillfjord, Måns Williamson, Monika Eisenmann

Clipping techniques are often used when solving optimization problems related to the training of machine learning applications. They make gradient-based optimization method more robust by rescaling the gradients such that excessively large steps are avoided. This is particularly useful in stochastic optimization, where the stochasticity leads to extra variability in the size of the gradients. In spite of this, so-called soft-clipping methods have not been analyzed to a large extent in the literature, and a rigorous mathematical analysis is lacking in the general, nonlinear case.

In this talk, I will therefore present recent results on convergence analysis of a large class of stochastic, component-wise soft-clipping schemes. This class covers several existing schemes and suggests a whole range of new schemes to be considered further. Our numerical experiments indicate that many of them perform on par with state-of-the-art methods, without significant tuning.

The main contribution of this work is the unifying convergence analysis. Under standard assumptions such as Lipschitz continuous gradients of the objective function, we give rigorous proofs of convergence in expectation. These include rates for both convex and non-convex problems. In addition, we prove almost sure convergence to a stationary point in the non-convex case.

■ TD-07

Thursday, 14:10 - 15:50 - Room: M:I

Accelerated methods in modern optimization

Stream: Optimization for Inverse Problems and Machine Learning
Invited session
Chair: *Enis Chenchene*

1 - Near-optimal Closed-loop Method via Lyapunov Damping for Convex Optimization

Camille Castera, Severin Maier, Peter Ochs

Optimal rates of convergence for first-order convex optimization are achieved via so-called open-loop damping strategies. In particular, Nesterov's algorithm features the intriguing momentum parameter (k-1)/(k+2) that depends on the iteration index k, making the initial iteration index an hyper-parameter that affects the performance of the algorithm. From a dynamical system perspective, we overcome this issue by designing the momentum (or damping) parameter in a closed-loop manner, using a Lyapunov function. This creates a feedback loop between the momentum and the speed of convergence of the system. We show that the resulting method achieves a convergence rate arbitrarily close to the optimal one while also getting rid of an additional hyper-parameter.

2 - Inertial methods beyond minimizer uniqueness

Hippolyte Labarrière

When considering a convex composite minimization problem, it is well known that momentum allows first-order methods to be accelerated both theoretically and numerically. Since the seminal works of Polyak in 1964 and Nesterov in 1983, inertial methods have been widely studied and we know that for a suitable choice of parameters inertial methods ensure fast convergence rates under additional geometry assumptions such as strong convexity. However, the improved convergence results demonstrated in the literature hold only if the function to minimize has a unique minimizer. This extra assumption is limiting, since some common functions (such as the LASSO function or $L1$ regularized functions in general) can satisfy some growth condition without having a unique minimizer. The question then arises: is this assumption necessary to prove fast convergence properties?

We propose an approach that aims to avoid this hypothesis while still obtaining fast convergence rates. This strategy allows to extend several known convergence results in the continuous setting (i.e. for dynamical systems associated to inertial schemes). We also provide fast convergence guarantees for the iterates of FISTA (introduced by Beck and Teboulle) and V-FISTA (proposed by Beck) in a relaxed setting, showing that this uniqueness assumption is not required for inertial methods to be efficient.

3 - Variance reduction techniques for stochastic proximal point algorithms

Cheik Traoré

In the context of finite sums minimization, variance reduction techniques are widely used to improve the performance of state-of-the-art stochastic gradient methods. Their practical impact is clear, as well as their theoretical properties. Stochastic proximal point algorithms have been studied as an alternative to stochastic gradient algorithms since they are more stable with respect to the choice of the stepsize but its variance reduced version are not as studied as the gradient ones. In this work, we propose the first unified study of variance reduction techniques for stochas- tic proximal point algorithms. We introduce a generic stochastic proximal algorithm that can be specified to give the proximal version of SVRG, SAGA, and some of their variants for smooth and convex functions. We provide several convergence results for the iterates and the objective function values. In addition, under the Polyak-Łojasiewicz (PL) condition, we obtain linear con- vergence rates for the iterates and the function values. Our numerical experiments demonstrate the advantages of the proximal variance reduction methods over their gradient counterparts, es- pecially about the stability with respect to the choice of the stepsize.

Thursday, 16:20 - 17:20

■ TE-01

Thursday, 16:20 - 17:20 - Room: M:A

Plenary III - Gabriele Eichfelder EUROPT Fellow

Stream: Plenaries
Plenary session
Chair: *Giancarlo Bigi*
Chair: *Pontus Giselsson*
Chair: *Oliver Stein*

1 - Multiobjective optimization, uncertainty, and a bit of set optimization
Gabriele Eichfelder

In many applications one has to deal with various difficulties at the same time, like uncertain data or several competing objective functions. For instance, for the integration of neighborhood networks into overarching distributing energy networks, a pure optimization of the neighborhood networks under an externally defined weighting of the relevant targets does not adequately model the problem. Moreover, uncertainties in the form of fluctuations or other disturbances can appear and a found solution has to be robust against that. A robust approach to uncertain multiobjective optimization corresponds to solving a set-valued optimization problem. However, it is a very difficult task to solve these optimization problems even for specific cases. In this talk, we give a short introduction and motivation for multiobjective optimization. In case of uncertainties, a family of parametrized multiobjective optimization problems has to be examined and we discuss approaches on how this is done in the literature. Thereby, basic concepts of set optimization are needed, i.e., of optimization with a set-valued objective function. For that reason, we also provide a short introduction to the main optimality concepts and to new solution approaches used in set optimization. We combine everything for obtaining a solvable reformulation of an uncertain multiobjective problem using a robust approach which extends the classical epigraphical reformulation from single-objective optimization to multiple objectives.

Friday, 8:45 - 9:35

■ FA-01

Friday, 8:45 - 9:35 - Room: M:A

Plenary IV - Sebastian Stich

Stream: Plenaries
Plenary session
Chair: *Alp Yurtsever*

1 - A Universal Framework for Federated (Convex) Optimization
Sebastian Stich

Federated learning has emerged as an important paradigm in modern large-scale machine learning. Unlike traditional centralized learning, where models are trained using large datasets stored on a central server, federated learning keeps the training data distributed across many clients, such as phones, network sensors, hospitals, or other local information sources. In this setting, communication-efficient optimization algorithms are crucial.

We provide a brief introduction to local update methods developed for federated optimization and discuss their worst-case complexity. Surprisingly, these methods often perform much better in practice than predicted by theoretical analyses using classical assumptions. Recent years have revealed that their performance can be better described using refined notions that capture the similarity among client objectives.

In this talk, we introduce a generic framework based on a distributed proximal point algorithm, which consolidates many of our insights and allows for the adaptation of arbitrary centralized optimization algorithms to the convex federated setting (even with acceleration). Our theoretical analysis shows that the derived methods enjoy faster convergence if the similarity among clients is high.

Based on joint work with Xiaowen Jiang and Anton Rodomanov.

Friday, 10:05 - 11:20

■ FB-02

Friday, 10:05 - 11:20 - Room: M:O

First-order methods for multi-level and multi-objective optimization

Stream: Advances in first-order optimization
Invited session
Chair: *Brecht Evens*

1 - Adaptivity in convex optimization beyond minimization
Puya Latafat, Andreas Themelis, Silvia Villa, Panagiotis Patrinos

Bilevel optimization is a comprehensive framework that bridges single- and multi-objective optimization. It encompasses many general formulations, such as, but not limited to, standard nonlinear programs. This work demonstrates how elementary proximal gradient iterations can be used to solve a wide class of convex bilevel optimization problems without involving subroutines. Compared to and improving upon existing methods, ours (1) can handle a much wider class of problems, including both constraints and nonsmooth terms, (2) requires mere local Lipschitz gradient continuity of the differentiable terms without imposing any strong convexity assumptions, and (3) provides a systematic adaptive stepsize selection strategy that leads to large and nonmonotonic stepsize sequences while being insensitive to the choice of parameters.

2 - Relying on first-order methods to deal with utility functions-based scalarization in multi-objective optimization
Lorenzo Lampariello, Simone Sagratella, Valerio Giuseppe Sasso, Vladimir Shikhman

We study a general scalarization approach via utility functions in multi-objective optimization. It consists of maximizing utility that is obtained from the objectives' bargaining with regard to a disagreement reference point. We propose a gradient-based numerical scheme to solve utility-dependent single-objective optimization problems. Here, the main difficulty comes from the necessity to address constraints which are associated with a disagreement reference point. Our crucial observation is that the explicit treatment of these additional constraints may be avoided. This is the case if the Slater condition is satisfied and the utility function under consideration has the so-called barrier property. Under these assumptions, we prove the convergence of our scheme to Pareto optimal points. Preliminary numerical experiments on real-world financial datasets in a portfolio selection context confirm the efficiency of our approach.

■ FB-03

Friday, 10:05 - 11:20 - Room: M:J

In memory of Georg Still - part 3

Stream: In memory of Georg Still
Invited session
Chair: *Oliver Stein*

1 - Parametric optimisation applied to a fitting problem in finance
Ralf Werner, Dirk Banholzer, Joerg Fliege

The Nelson-Siegel and the Svensson model are two of the most widely used models for the term structure of interest rates. Even though the models are quite simple and intuitive, fitting them to market data is numerically challenging and various difficulties have been reported. For this reason a novel mathematical analysis of the fitting problem based on parametric optimisation is carried out. It is based on the known observation that the fitting problem can be formulated as a separable nonlinear least-squares problem, in which the linear parameters can be eliminated. We specifically provide a thorough discussion on the conditioning of the inner part of the reformulated problem and show that many of the reported difficulties encountered when solving it are inherent to the problem formulation and cannot be tackled by choosing a particular optimisation algorithm. Our stability analysis provides novel insights that we then use to show that some of the ill-conditioning of the problem can be avoided, and that a suitably chosen penalty approach can be used to take care of the remaining ill-conditioning. As our numerical results indicate, this approach has indeed the expected impact, while being independent of the choice of a particular optimisation algorithm. As a side benefit, we establish smoothness and differentiability properties of the reduced objective function, which for the first time puts global optimisation methods for the reduced problem on a sound mathematical basis

2 - Parametric stationary point sets need MFCQ for stably being topological manifolds
Harald Guenzel, Daniel Hernandez Escobar, Jan-J Ruckmann

We consider a parametric standard optimization problem with smooth problem data and discuss the question under which condition its stationary point set stably constitutes a topological manifold having the same dimension as that of the parameter space. In other words, we ask, when this set is locally a manifold of that dimension and it also stays such a manifold after small perturbations of the problem data.

We show that this question can only be answered affirmatively if the Mangasarian-Fromovitz Constraint Qualification holds at the stationary point under consideration.

3 - Global aspects of the continuous reformulation for cardinality-constrained optimization problems
Vladimir Shikhman, Sebastian Lämmel

The main goal of this paper is to relate the topologically relevant stationary points of a cardinality-constrained optimization problem and its continuous reformulation up to their type. For that, we focus on the non-degenerate M- and T-stationary points, respectively. Their M-index and T-index, respectively, which uniquely determine the global and local structure of optimization problems under consideration in algebraic terms, are traced. As novelty, we suggest to regularize the continuous reformulation for this purpose. The main consequence of our analysis is that the number of saddle points of the regularized continuous reformulation grows exponentially as compared to that of the initial cardinality-constrained optimization problem. This growth appears to be inherent and is reproduced in other relaxation attempts.

■ FB-04

Friday, 10:05 - 11:20 - Room: M:M

Large-scale optimization II

Stream: Large-scale optimization
Invited session
Chair: *Anton Åkerman*

1 - An interior proximal gradient method for nonconvex optimization
Alberto De Marchi, Andreas Themelis

We consider structured minimization problems subject to smooth inequality constraints and present a flexible algorithm that combines interior point (IP) and proximal gradient schemes. While traditional IP methods cannot cope with nonsmooth objective functions and proximal algorithms cannot handle complicated constraints, their combined usage is shown to successfully compensate the respective shortcomings. We provide a theoretical characterization of the algorithm and its asymptotic properties, deriving convergence results for fully nonconvex problems, thus bridging the gap with previous works that successfully addressed the convex case. Our interior proximal gradient algorithm benefits from warm starting, generates strictly feasible iterates with decreasing objective value, and returns after finitely many iterations a primal-dual pair approximately satisfying suitable optimality conditions. As a byproduct of our analysis of proximal gradient iterations we demonstrate that a slight refinement of traditional backtracking techniques waives the need for upper bounding the stepsize sequence, as required in existing results for the nonconvex setting.

2 - Krasnoselskii-Mann Iterations: Inertia, Perturbations and Approximation
Daniel Cortild, Juan Peypouquet

Krasnoselskii-Mann iterations constitute a class of fixed point iterations combined with relaxations, employed to approximate fixed points of (quasi-)nonexpansive operators. We present a study of a family of such iterations combining different inertial principles into a single framework. We provide a systematic, unified and insightful analysis of the hypotheses that ensure their weak, strong and linear convergence, either matching or improving previous results obtained by analysing particular cases separately. We also show that these methods are robust with respect to different kinds of perturbations–which may come from computational errors, intentional deviations, as well as regularisation or approximation schemes–under surprisingly weak assumptions on the magnitude of the perturbations. Although we mostly focus on theoretical aspects, a numerical illustration based on the image inpainting problem reveals possible trends in the behaviour of these types of methods.

The talk is based on joint work with Juan Peypouquet.

3 - A Fast Optimistic Method for Monotone Variational Inequalities
Michael Sedlmayer, Dang-Khoa Nguyen, Radu Ioan Bot

We study monotone variational inequalities that can arise as optimality conditions for constrained convex optimization or convex-concave minimax problems and propose a novel algorithm that uses only one gradient/operator evaluation and one projection onto the constraint set per iteration. The algorithm, which we call fOGDA-VI, achieves a o(1/k) rate of convergence in terms of the restricted gap function as well as the natural residual for the last iterate. Moreover, we provide a convergence guarantee for the sequence of iterates to a solution of the variational inequality. These are the best theoretical convergence results for numerical methods for (only) monotone variational inequalities reported in the literature.

■ FB-05

Friday, 10:05 - 11:20 - Room: M:N

Recent advances in bilevel optimization II

Stream: Bilevel optimization: strategies for complex decision-making
Invited session
Chair: *Alp Yurtsever*
Chair: *Ahmet Alacaoglu*

1 - Interior point method based on the optimal value function for bilevel optimization problems
Yasushi Narushima, Seima Yamamoto

The bilevel optimization problem has upper-level and lower-level optimization problems. First, the upper-level problem is solved and then the lower-level problem is solved; the two problems are interrelated. To develop an efficient algorithm for solving the bilevel optimization problem, we apply the interior point method to the bilevel optimization problem. Because the method is effective for nonlinear optimization problems, we expect that the interior point method for the bilevel optimization problem will also be efficient. First, we derive the optimality conditions using the optimal value function because more manageable optimality conditions can be obtained than those that use the KKT conditions of the lower-level problem. Based on these optimality conditions, we propose an interior point method that guarantees local convergence. Additionally, by modifying this algorithm, we propose an interior point method that ensures global convergence. In this algorithm, we use the squared L2-norm of the shifted KKT condition as a merit function because the values of the optimal value function cannot be expressed explicitly. Moreover, by refining the calculation of step sizes, we ensure interior points and guarantee global convergence. Finally, some numerical results are given to confirm the effectiveness of the proposed algorithm.

2 - Asymmetric data-driven interdiction problems with cost uncertainty: a distributionally robust optimization approach
Sergei Ketkov

In this study, we consider a class of bilevel interdiction problems, where the cost coefficients are subject to uncertainty. More precisely, in our problem setting, the cost vector is treated as a random vector, whose probability distribution can only be observed through a finite training data set. Furthermore, both the leader and the follower are assumed to possess limited and potentially asymmetric information about the nominal (true) distribution. In order to address the cost uncertainty, we formulate a bilevel Wasserstein distributionally robust optimization (BDRO) problem, where both decision-makers attempt to hedge themselves against the worst-case distribution and realization of the random cost vector. In the case where the leader has full information about the follower's data set, BDRO with properly defined Wasserstein ambiguity sets and objective criteria is shown to admit a single linear mixed-integer programming (MILP) reformulation. On the other hand, when the leader has only partial or no information about the follower's data, we propose two alternative formulations of BDRO, based, respectively, on oversampling from the leader's data set and a conservative robust approximation. Finally, we explore numerically how the quality of data and decision-maker's risk preferences affect the model's out-of-sample performance.

3 - Hyperparameter optimization for kernel-regularized system identification
Lujing Chen, Martin Skovgaard Andersen, Tianshi Chen

System identification is an example of an inverse problem in which one seeks to estimate a mathematical model of a system based on knowledge of its input and output. In the past decade, significant advances have been made within kernel-regularized methods where prior knowledge is included into the estimation problem in the form of a kernel function with small a number of hyperparameters. The problem of finding a suitable set of hyperparameters can be formulated as an optimization problem. This is generally a difficult nonconvex problem with an objective function that is expensive to evaluate when dealing with complex models and large data sets.

To address the computational challenges associated with the hyperparameter optimization problem, we explore the use of cheap, approximate function evaluations combined with Bayesian optimization. Specifically, we employ Krylov subspace methods and stochastic trace estimation with randomized preconditioning to compute an approximate function value. Our methodology yields promising results, as demonstrated by numerical experiments.

■ FB-06

Friday, 10:05 - 11:20 - Room: M:H

Higher-order Methods in Mathematical Programming II

Stream: Challenges in nonlinear programming
Invited session
Chair: *Marianna E.-Nagy*

1 - Analysis of the primal-dual central path for nonlinear semidefinite optimization without the nondegeneracy condition
Takayuki Okuno

We study properties of the central path underlying a nonlinear semidefinite optimization problem, called NSDP for short. The latest radical work on this topic was contributed by Yamashita and Yabe (2012): they proved that the Jacobian of a certain equation-system derived from the Karush-Kuhn-Tucker (KKT) conditions of the NSDP is nonsingular at a KKT point under the second-order sufficient condition (SOSC), the strict complementarity condition (SC), and the nondegeneracy condition (NC). This yields uniqueness and existence of the central path through the implicit functiontheorem. In this paper, we consider the following three assumptions on aKKT point: the enhanced SOSC, the SC, and the Mangasarian-Fromovitzconstraint qualification. Under the absence of the NC, the Lagrangemultiplier set is not necessarily a singleton and the nonsingularity of the above-mentioned Jacobian is no longer valid. Nonetheless, we establish that the central path exists uniquely, and moreover prove that the dualcomponent of the path converges to the so-called analytic center of the Lagrange multiplier set. As another notable result, we clarify a regionaround the central path where Newton's equations relevant to primal-dual interior point methods are uniquely solvable.

2 - Developing Extended Memoryless Optimization Algorithms Based on the Ellipsoid Norms
Saman Babaie-Kafaki

Although the Euclidean norm has been widely used in the analysis of algorithmic and modeling aspects of optimization, the ellipsoid norm has also been employed in the analytical spectrum of literature. Effective applications of the ellipsoid norm can be observed in the convergence analysis of steepest descent algorithm, devising quasi-Newton updates, and particularly in developing scaled trust region methods. As an extension of the Euclidean norm, the ellipsoid norm is capable of being flexibly utilized in designing new memoryless algorithms to address high-dimensional real-world optimization problems that frequently arise in the contemporary world. Therefore, it deserves more attention. Hence, in an attempt to promote diversity and inclusion in the optimization tools in accordance with the current standards, the ellipsoid norm is applied in the least-squares framework, using quasi-Newton updating formulas as the index of the norm. As part of this endeavor, several extended formulas for the scaling parameter of memoryless quasi-Newton algorithms are developed, which can also be regarded as scalar approximations of the (inverse) Hessian. Furthermore, through a similar approach, several adaptive choices are suggested for the parameter of the Dai-Liao method, a well-known class of nonlinear conjugate gradient algorithms. The effect of such extensions on hybrid conjugate gradient algorithms is also studied. Finally, computational experiment results are reported.

3 - On a Unified Analysis of Kernel-based Interior Point Algorithms
Marianna E.-Nagy, Zsolt Darvay, Goran Lesaja, Petra Renáta Rigó, Anita Varga

Kernel-based interior point algorithms (IPA) were developed to improve the theoretical complexity of large step IPAs. In the literature, various kernel functions were defined and an upper bound was given on the iteration number of the corresponding IPAs. In 2010, Lesaja and Roos considered eligible kernel function based IPAs for sufficient linear complementarity problems and provided a general scheme for steps of the complexity analysis. However, this framework necessitates independently verifying its steps for different kernel functions.

To address this challenge, we propose additional, not very restrictive conditions on the class of eligible kernel functions and present a unified analysis for these kernel-based IPAs. Our subset of eligible kernel functions contains all kernel functions with polynomial and exponential barrier terms mentioned in the literature. We demonstrate consistent complexity bounds across all cases.

Friday, 11:25 - 12:40

■ FC-02

Friday, 11:25 - 12:40 - Room: M:O

Algorithms for Variational inequalities and equilibria

Stream: Advances in first-order optimization
Invited session
Chair: *Emanuel Laude*

1 - Recent advances in first-order methods for weak Minty variational inequalities
Thomas Pethick, Ioannis Mavrothalassitis, Volkan Cevher

First-order methods have classically been developed under the de facto assumption of monotonicity. However, recently there has been a surge of interest in developing methods for nonmonotone problems for which convergence guarantees of classical methods no longer hold. One such nonmonotone structure, which remains amenable to analysis, is the weak Minty variational inequality (MVI). The allowed degree of nonmonotonicity in the weak MVI has been pushed in recent years through a series of works. In this talk we will discuss the underlying mechanism behind these advances and provide a unifying framework for several existing methods.

2 - Progressive decoupling of linkage problems beyond elicitable monotonicity
Brecht Evens, Puya Latafat, Panagiotis Patrinos

In this talk, we study the progressive decoupling algorithm and Spingarn's method of partial inverses for finding a point in the graph of an operator such that the first, primal variable belongs to a closed subspace and the second, dual variable belongs to its orthogonal complement. This so-called linkage problem encompasses many problems emerging in optimization and variational inequalities, owing to its close connection with Lagrangian duality. In particular, we establish convergence of both methods in the absense of any (elicitable) monotonicity assumptions, by leveraging their connection with the preconditioned proximal point method. Additionally, we showcase the broad range of problems our theory is able to cover through several illustrative examples.

3 - Projected solutions for quasi-equilibria
Giancarlo Bigi

The concept of a projected solution for quasi-variational inequalities and generalized Nash equilibrium problems was recently introduced, due to the modeling of deregulated electricity markets where the constraint map is not a self-map. In this talk we show that the projected solutions of a quasi-equilibrium problem correspond to the classical solutions of an auxiliary quasi-equilibrium problem. This can be achieved by doubling the number of variables and adding an appropriate term. In this way algorithms for quasi-equilibria can be exploited for computing projected equilibria through the auxiliary problem. However, its structure does not guarantee the fulfilment of the assumptions required for the convergence of algorithms. In particular, descent and extragradient algorithms need that each feasible point is a fixed point for the constraint map and this not true in the case at hand. Thus, we show that under suitable assumption an ad-hoc choice of parameters allows convergence of the extragradient algorithm without the above requirement. the results of preliminary numerical tests are reported as well.

The talk is based on joint papers with Marco Castellani and Sara Latini.

■ FC-03

Friday, 11:25 - 12:40 - Room: M:J

In memory of Georg Still - part 4

Stream: In memory of Georg Still
Invited session
Chair: *Mirjam Duer*

1 - How to avoid the normal cone in the subdifferential calculus?
Marco A. López-Cerdá

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

2 - New results on copositive optimization obtained via linear semi-infinite optimization theory
Miguel Goberna, Andrea Ridolfi, Virginia N. Vera de Serio

In this talk we present existence theorems, results on the geometry of the feasible set, duality theorems, and optimality conditions for copositive optimization problems. Following the same approach used by Mirjam Dür and Georg Still in their pioneering papers "Copositive programming via semi-infinite optimization" [J. Optim. Theory Appl. 159 (2013) 322-340, with Faizan Ahmed] and "Genericity results in linear conic programming - A tour d'horizon" [Math. Oper. Res. 42 (2016) 77-94, with Bolor Jargalsaikhan], our new results have been derived from classical and updated ones of linear semi-infinite optimization theory.

3 - Some sincere thoughts of thanks about Professor Georg Still, our friend
Gerhard-Wilhelm Weber

This talk adds some professional and, not least, personal elements to the appreciation of Professor Georg Still. They cover a period since 1988, address some of his very important scientific achievements, some joint conferences and conference organizations, especially in EUROPT, his EUROPT fellowship, and especially his warm-hearted and loyal personality, which has manifested itself in the life of many people and forever in EUROPT.

■ FC-04

Friday, 11:25 - 12:40 - Room: M:M

Large-scale optimization III

Stream: Large-scale optimization
Invited session
Chair: *Anton Åkerman*

1 - A unified Euler–Lagrange system for analyzing continuous-time accelerated gradient methods
Mitsuru Toyoda, Akatsuki Nishioka, Mirai Tanaka

This talk presents a unified Euler–Lagrange system of continuous-time accelerated gradient methods and its convergence analysis based on a unified Lyapunov function. To understand the mysterious dynamics of discrete-time accelerated gradient methods, continuous-time ordinary differential equation (ODE) models have been recently developed; however, dumping terms appearing in the ODE models still have technical time-varying coefficients and have difficulties on the interpretability. Besides, in various problem settings (e.g., strong convexity of an objective function), an associated model and its Lyapunov function have been separately developed. Based on the background, this study proposes a Lagrangian, deriving an Euler–Lagrange system, and Lyapunov function for the unified analysis of existing ODE models. The introduced Lyapunov function naturally shows the conditions that the design parameters in the Lagrangian and Lyapunov function must satisfy. The obtained conditions suggests that the dumping terms in the ODE models are chosen to improve the resulting convergence rate. Furthermore, the extension to non-smooth objective functions with associated smooth approximation is available. The preprint of this study is available at “arXiv:2404.03383”.

2 - An optimal lower bound for smooth convex functions
Mihai I. Florea

First order methods endowed with global convergence guarantees operate using global lower bounds on the objective. The tightening of the bounds has been shown to lead to an increase in both the theoretical guarantees and practical performance. In this work, we define a global lower bound for smooth convex objectives that is optimal with respect to the collected oracle information. Using the machinery underlying the optimal bounds, we construct an Optimized Gradient Method with Memory possessing the best known convergence guarantees for its class of algorithms, even in terms of the proportionality constant. We additionally equip the method with an adaptive convergence guarantee adjustment procedure that is an effective replacement for line-search. Preliminary simulation results validate the theoretical properties of the proposed method.

3 - Coordinate Descent Algorithm for Nonlinear Matrix Decompositions with the ReLU function
Atharva Awari

Nonlinear Matrix Decompositions (NMD) solve the following problem: Given a matrix X, find low-rank factors W and H such that X is approximated by f(WH), where "f" is an element-wise nonlinear function. In this paper, we focus on the case when "f" is the rectified linear unit (ReLU) activation, that is, the function which maps all negative entries to zero and keeps positive entries the same. This is referred to as ReLU-NMD. All state-of-the-art algorithms for ReLU-NMD have been designed to solve a reformulation of ReLU-NMD. It turns out that this reformulation leads to a non-equivalent problem, and hence to suboptimal solutions. In this paper, we propose a coordinate-descent (CD) algorithm designed to solve ReLU-NMD directly. This allows us to compute more accurate solutions, with smaller error. This is illustrated on synthetic and real-world datasets.

■ FC-05

Friday, 11:25 - 12:40 - Room: M:N

Recent advances in bilevel optimization III

Stream: Bilevel optimization: strategies for complex decision-making
Invited session
Chair: *Alp Yurtsever*

1 - General single-loop splitting methods for bilevel optimization
Ensio Suonperä, Tuomo Valkonen

Bilevel problems have been traditionally solved through either treating the inner problem as a constraint, and solving the resulting Karush–Kuhn–Tucker conditions using a Newton-type solver; or by trivialising the inner problem to its solution mapping. These approaches are difficult to scale to large problems. The latter approach in principle requires near-exact solution of the inner problem for each outer iterate. Recently, intermediate approaches have surfaced that solve the inner problem to a low precision and still obtain some form of convergence. In this talk, we discuss the linear convergence of methods based on taking interleaved steps of proximal-type methods on both the inner and outer problem. We demonstrate numerical performance on imaging applications. Such bilevel learning problems have a large scale, as the dimension depends on the size of the images, which can be in the order of millions of pixels, the size of the training set, and the number of parameters to be learned.

2 - An adaptively inexact first-order method for bilevel learning
Mohammad Sadegh Salehi, Matthias J. Ehrhardt, Lindon Roberts

In various imaging and data science domains, tasks are modeled using variational regularization, which poses challenges in manually selecting regularization parameters, especially when employing regularizers involving a large number of hyperparameters. To tackle this, gradient-based bilevel learning, as a large-scale approach, can be used to learn parameters from data. However, the unattainability of exact function values and gradients with respect to hyperparameters necessitates reliance on inexact evaluations. State-of-the-art inexact gradient-based methods face difficulties in selecting accuracy sequences and determining appropriate step sizes due to unknown Lipschitz constants of hypergradients.

In this talk, we present our algorithm, the "Method of Adaptive Inexact Descent (MAID)," featuring a provably convergent backtracking line search that incorporates inexact function evaluations and hypergradients. This ensures convergence to a stationary point and adaptively determines the required accuracy. Our numerical experiments demonstrate MAID's practical superiority over state-of-the-art methods on an image denoising problem. Importantly, we showcase MAID's robustness across different initial accuracy and step size choices.

3 - Adaptive bilevel optimisation
Kimon Antonakopoulos, Shoham Sabach, Luca Viano, Mingyi Hong, Volkan Cevher

We propose a new textitadaptive optimization algorithm based on mirror descent for a class of possibly non-convex smooth bilevel optimization problems. The optimization template is broadly applicable in machine learning as it features two coupled problems where the optimal solution set of an inner problem serves as a constraint set for the outer problem. As such, existing algorithms require knowledge of gradient Lipschitz constants of both inner and outer levels and are often challenging to tune in practice. Our adaptive algorithm, to our knowledge the first in this setting, circumvents this difficulty by using an AdaGrad-type accumulation strategy on gradient norms and obtains a convergence rate of in terms of the outer objective function, when it is convex, where is the number of iterations. When the outer objective is non-convex, our algorithm obtains an best-iterate guarantee for the squared norm of the gradient of the outer objective function. We also provide numerical evidence to support the theory in a reinforcement learning setting where all problem parameters are accessible.

■ FC-07

Friday, 11:25 - 12:40 - Room: M:I

Optimization mosaics

Stream: Optimization in regression, classification and learning
Invited session
Chair: *Paula Amaral*
Chair: *Sebastian Lämmel*

1 - Relaxations for the elementary shortest path problem
Regina Schmidt, Mirjam Duer

Given a directed graph with negative arc costs and possibly negative cycles, the elementary shortst path problem consists in finding a shortest elementary path, i.e., a path from a source node to a target node that visits each node on the path exactly once. If negative arc costs are allowed, then the problem is NP-hard. In this talk, we discuss various formulations and relaxations for this problem.

2 - Extended convergence analysis of the Scholtes-type regularization for cardinality-constrained optimization problems
Sebastian Lämmel, Vladimir Shikhman

We extend the convergence analysis of the Scholtes-type regularization method for cardinality-constrained optimization problems. Its behavior is clarified in the vicinity of saddle points, and not just of minimizers as it has been done in the literature before. This becomes possible by using as an intermediate step the recently introduced regularized continuous reformulation of a cardinality-constrained optimization problem. We show that the Scholtes-type regularization method is well-defined locally around a nondegenerate T-stationary point of this regularized continuous reformulation. Moreover, the nondegenerate Karush-Kuhn-Tucker points of the corresponding Scholtes-type regularization converge to a T-stationary point having the same index, i.e. its topological type persists. As consequence, we conclude that the global structure of the Scholtes-type regularization essentially coincides with that of CCOP.

Friday, 14:10 - 15:50

■ FD-02

Friday, 14:10 - 15:50 - Room: M:O

Deterministic and stochastic optimization beyond Euclidean geometry

Stream: Advances in first-order optimization
Invited session
Chair: *Adrien Taylor*
Chair: *Hadrien Hendrikx*

1 - Horospherically Convex Optimization on Hadamard Manifolds
Christopher Criscitiello, Jungbin Kim

Many Euclidean notions, like affine functions, do not generalize well within the framework of geodesic convexity. Using Busemann functions as a building block, we introduce an alternative generalization of convexity to Hadamard manifolds called horospherical convexity (h-convexity). We provide algorithms for h-convex optimization which have rates *exactly* matching those from Euclidean space (including full acceleration). As a consequence, we obtain the best known rates for the minimal enclosing ball problem. We also establish necessary and sufficient conditions for h-convex interpolation.

2 - Implicit Regularisation of Mirror Flow on Separable Classification Problems
Radu-Alexandru Dragomir

We study continuous-time counterpart of mirror descent, namely mirror flow, on classification problems which are linearly separable. Such problems are minimised ‘at infinity’ and have many possible solutions; we study which solution is preferred by the algorithm depending on the mirror potential. We show that the iterates converge in direction towards the solution of a certain max-margin problem. This problem is determined by the horizon function of the potential, which can be seen as the norm induced by its shape ‘at infinity’. When the potential is separable, a simple formula allows to compute this function. We also prove the general existence of the horizon shape for subanalytic potentials.

Joint work with Scott Pesme and Nicolas Flammarion.

3 - Accelerated bregman divergence optimization with smart: an information geometric point of view
Stefania Petra

We consider the minimization of the Kullback-Leibler divergence between a linear model Ax and a positive vector b across various convex domains, such as the positive orthant, n-dimensional box, and probability simplex. Our focus is on the SMART method, which employs efficient multiplicative updates and is an exponentiated gradient method. This method offers dual interpretations as a Bregman proximal gradient method and as a Riemannian gradient descent on the parameter manifold of a corresponding distribution within the exponential family. This duality allows us to establish connections, facilitating accelerated SMART iterates while seamlessly incorporating constraints. Furthermore, it enables the development of a multilevel method utilizing first-order Riemannian optimization. To validate the efficacy of the acceleration schemes, we present results from extensive numerical experiments on large-scale datasets

4 - Investigating Variance Definitions for Stochastic Mirror Descent with Relative Smoothness
Hadrien Hendrikx

Mirror Descent is a popular algorithm, that extends Gradients Descent (GD) beyond the Euclidean geometry. One of its benefits is to enable strong convergence guarantees through smooth-like analyses, even for objectives with exploding or vanishing curvature. This is achieved through the introduction of the notion of relative smoothness, which holds in many of the common use-cases of Mirror descent. While basic deterministic results extend well to the relative setting, most existing stochastic analyses require additional assumptions on the mirror, such as strong convexity (in the usual sense), to ensure bounded variance. In this talk, we will revisit Stochastic Mirror Descent (SMD) proofs in the (relatively-strongly-) convex and relatively-smooth setting, and introduce a new (less restrictive) definition of variance which can generally be bounded (globally) under mild regularity assumptions. We will then investigate this notion in more details, and show that it naturally leads to strong convergence guarantees for stochastic mirror descent. Finally, we will leverage this new analysis to obtain convergence guarantees for the Maximum Likelihood Estimator of a Gaussian with unknown mean and variance.

■ FD-03

Friday, 14:10 - 15:50 - Room: M:J

Variational techniques in optimization

Stream: Variational analysis: theory and algorithms
Invited session
Chair: *Juan Enrique Martínez-Legaz*

1 - The Geometry of Sparsity-Inducing Balls
Michel De Lara

Sparse optimization seeks an optimal solution among vectors with at most k nonzero coordinates. This constraint is hard to handle, and a strategy to overcome that difficulty amounts to adding a norm penalty term to the objective function. The most widely used penalty is based on the l1-norm which is recognized as the archetype of sparsity-inducing norms. In this talk, we present generalized k-support norms, generated from a given source norm, and show how they contribute to induce sparsity via support identification. In case the source norms are the l1- and the l2-norms, we analyze the faces and normal cones of the unit balls for the associated k-support norms and their dual top-k norms.

2 - Parameter identification in PDEs by the solution of monotone inclusion problems
Pankaj Gautam, Markus Grasmair

In this paper, we consider a parameter identification problem for a semilinear parabolic PDE. For the regularized solution of this problem, we introduce a total variation based regularization method requiring the solution of a monotone inclusion problem. We show well-posedness in the sense of inverse problems of the resulting regularization scheme. In addition, we introduce and analyze a numerical algorithm for the solution of this inclusion problem using a nested inertial primal dual method. We demonstrate by means of numerical examples the convergence of both the numerical algorithm and the regularization method.

3 - Closed convex sets that are both Motzkin decomposable and generalized Minkowski sets
Juan Enrique Martínez-Legaz, Cornel Pintea

We consider and characterize closed convex subsets of the Euclidean space that are simultaneously Motzkin decomposable and generalized Minkowski sets or, shortly, MdgM sets. We also prove the existence of suitably defined fixed points for (possibly multivalued) functions defined on MdgM sets along with existence of classical fixed points for some single valued self functions of MdgM sets. The first mentioned type of existence results are based on the Kakutani fixed point theorem, and the second type are obtained by combining the Brouwer fixed point theorem with the Banach contraction principle.

■ FD-04

Friday, 14:10 - 15:50 - Room: M:M

Optimal control and stochastic optimal control - theory, methods and applications 1

Stream: Optimal control and stochastic optimal control - theory, methods and applications
Invited session
Chair: *Gerhard-Wilhelm Weber*

1 - Analysis of Finance-Human Factor Interactions Using Various Indicators
Betül Kalaycı, Vilda Purutcuoglu, Gerhard-Wilhelm Weber

There are several empirical phenomena regarding the behaviour of individual investors, such as how their emotions and opinions influence on their decisions. The Sentiment term describes all of these emotions and opinions. In finance, stochastic changes may occur based on sentiment levels of investors. Machine Learning methods are well-known and useful tools for prediction problems, and they have already been used successfully to solve a variety of financial problems. In this study, apart from pure financial related challlenges, we focus on behavior of financial problems which is based on the investors' sentiment levels. Accordingly, the purpose of this study is to assess sentiment index predicting performance by utilizing two-stage hybrid models which are the composition of multivariate adaptive regression splines, random forest and neural networks. We additionally intend to define people's attitudes on the economy based on their levels of confidence. Hereby, we perform hidden markov model (HMM) to estimate the underlying state of changes in the Consumer Confidence Index and to examine its link with some macroeconomic indicators (CPI, GDP, and currency rate) at monthly intervals. Our goal is to observe and understand the transition between these phases, as well as to define a path through them. Furthermore, we apply volatility models to each subgroup generated via HMM in order to check whether forecasting outcomes can be improved.

2 - Regime-switching models via stochastic optimal control & robust control theory, with applications in finance and insurance
Gerhard-Wilhelm Weber, Emel Savku, Ioannis Baltas, Athanasios Yannacopoulos

This presentation consists in newest advancements in both (i) stochastic optimal control and games with jumps in finance under delay and regime switching, and (ii) stochastic optimal control and games in pension fund systems with new elements of regime switching and longevity. (iii) Time is reserved to enjoy the beauty of the underlying and employed mathematics, to discussions about pros and cons of different approaches, e.g., maximum principle vs dynamic program and Isaacs-Hamilton-Jacobi-Bellman equation, as well as to outlooks at future studies and applications, such as in physics, neuroscience and cosmology.

The talks is based on joint works of G.W. Weber with E. Savku, I. Baltas, L. Dopierala, K. Kolodziejczyk, M. Szczepański and A.N. Yannacopoulos.

3 - Calibration and Higher Lower Partial Moments of Skew Elliptical Distributions
Kerem Ugurlu

Higher lower partial moments for the skew elliptical family of probability distributions are retrieved. Explicit characterizations for skew normal and skew-t distributions are derived. Thorough numerical study on financial data is also illustrated. In particular, it is validated that daily log differences of stock prices can be characterized by the skew elliptical family, and the normal distribution assumption can not be statistically validated by the data. The work paves explicit methods for portfolio optimization with respect to lower partial moments of the financial data.

■ FD-05

Friday, 14:10 - 15:50 - Room: M:N

Structured nonsmooth optimization

Stream: Nonsmooth and nonconvex optimization algorithms
Invited session
Chair: *Alireza Kabgani*

1 - The Boosted Difference of Convex Functions Algorithm for Value-at-Risk Constrained Portfolio Optimization
Marah-Lisanne Thormann, Vuong Phan, Alain Zemkoho

A highly relevant problem of modern finance is the design of Value-at-Risk (VaR) optimal portfolios. Due to contemporary financial regulations, banks and other financial institutions are tied to use the risk measure to control their credit, market, and operational risks. For a portfolio with a discrete return distribution and finitely many scenarios, a Difference of Convex (DC) functions representation of the VaR can be derived. Wozabal (2012) showed that this yields a solution to a VaR constrained Markowitz style portfolio selection problem using the Difference of Convex Functions Algorithm (DCA). A recent algorithmic extension is the so-called Boosted Difference of Convex Functions Algorithm (BDCA) which accelerates the convergence due to an additional line search step. It has been shown that the BDCA converges linearly for solving non-smooth quadratic problems with linear inequality constraints. In this paper, we prove that the linear rate of convergence is also guaranteed for a piecewise linear objective function with linear equality and inequality constraints using the Kurdyka- Łojasiewicz property. An extended case study under consideration of best practices for comparing optimization algorithms demonstrates the superiority of the BDCA over the DCA for real-world financial market data. We are able to show that the results of the BDCA are significantly closer to the efficient frontier compared to the DCA.

2 - Weak subgradients and radial epiderivatives: calculus and optimization
Refail Kasimbeyli

Nonsmooth analysis has its origins in the early 1970s when nonlinear programmers attempted to deal with optimality conditions for problems with nonsmooth functions. The pointwise maximum or minimum of smooth functions are simple examples that can be considered as an illustration of nonsmoothness. In this paper, we study two kinds of generalized derivatives which allow us to investigate global optimal solutions of nonconvex and nondifferentiable optimization problems. These are the weak subgradient and the radial epiderivative concepts. Unlike the classical derivative and many generalized derivative concepts, the weak subgardient and the radial epiderivative concepts can be used to characterize global behaviour of objective functions in nonconvex and nonsmooth optimization. We present calculus rules for some classes of nondifferentiable functions and some applications in nonsmooth optimization. We establish a necessary and sufficient condition for global descent direction via the radial epiderivative. We believe that this condition can be used to develop a solution method for escaping from local minima. The paper also presents necessary and sufficient conditions for global optimum in nonconvex nonsmooth optimization by using the weak subgradients and the radial epiderivatives.

3 - Trust-region methods for relatively smooth weakly convex optimization
Mohammad Hamed, Masoud Ahookhosh

Objective functions with relative smoothness (a generalization of the Lipschitz smoothness) and weak convexity (a generalization of convexity) encompass many applications in the domains such as signal and image processing, machine learning and inverse problems. In this talk, we first generate a smooth approximation of the objective function via Bregman forward-backward envelope (BFBE) and design a trust-region method to find a critical point of the original function. This includes, the convergence analysis and the numerical experiments demonstrating the theoretical and practical efficiency of the proposed method.

4 - Subgradient Methods for Minimizing Paraconvex Functions
Morteza Rahimi, Masoud Ahookhosh

The primary objective of this talk is to study the convergence analysis subgradient methods for paraconvex functions which satisfy an appropriate error bound property and grow sharply away from its solution set. We establish that a linear convergence of these subgradient methods in the paraconvex setting. In these methods, the parameters regarding paraconvex and error bound properties only appear to find a valid initial point. As such, the linear rate of convergence is independence of these constant values. Some preliminary numerical result validate or theoretical foundations.

■ FD-06

Friday, 14:10 - 15:50 - Room: M:H

Difference and decomposition methods

Stream: Methods for non-/monotone inclusions and their applications
Invited session
Chair: *Dânâ Davar*

1 - Douglas-Rachford DC methods for generalized DC programming
Avinash Dixit

Nonconvex functions are a trending framework among researchers in optimization. In this paper, we consider the difference of convex functions (DC) programming problems which is the backbone of nonconvex programming and global optimization. The classical problem contains the difference between two proper convex and lower semicontinuous functions. This paper deals with the generalized DC programming problem which deals with the minimization of three convex functions. We propose two novel methods, an inertial Douglas Rachford DC algorithm and a parametrized Douglas Rachford DC algorithm to solve the problem. We study their convergence behavior in the Hilbert space framework. Lastly, we conducted numerical experiments to study the application of proposed algorithms to solve real-world problems. The results show that the proposed algorithms outperform other already proposed algorithms.

2 - Revisiting Frank-Wolfe for Nonconvex Problems
Hoomaan Maskan, Suvrit Sra, Alp Yurtsever

Difference of convex (DC) programs are a class of nonconvex problems that address a wide range of applications. Therefore, there has been a growing effort to solve these problems more efficiently. In this talk, we introduce a Frank-Wolfe algorithm for minimizing difference of convex functions over a convex and compact set. Our analysis indicates that the proposed method converges to a stationary point of the problem at an O(1/sqrt(k)) rate. Notably, while this rate resembles that of nonconvex Frank-Wolfe methods, numerical simulations demonstrate that our proposed approach often identifies a better local optimal point with a smaller objective value compared to the standard nonconvex Frank-Wolfe algorithm.

3 - Penalty Decomposition methods for convex and nonconvex market equilibrium models
Giulio Scarponi, Marco Sciandrone

Market equilibrium problems are usually formulated as optimization problems involving the following issues: existence of an equilibrium and possible unicity. Both in convex and nonconvex case, existence of market equilibria is related to the optima of the welfare optimization problem. Penalty Decomposition methods take into account the peculiarities of the problem in order to efficiently compute market equilibria in large-scale settings. In this talk we analyze Penalty Decomposition methods both for convex and nonconvex problems and we present the results of a computational study in market equilibrium problems.

4 - A derivative-free trust-region method based on finite differences for composite nonsmooth optimization
Dânâ Davar, Geovani Grapiglia

We present a derivative-free trust-region method based on finite-difference gradient approximations for minimizing composite functions. The outer function is supposed to be convex and Lipschitz, possibly nonsmooth, while its argument is a vector blackbox function with Lipschitz continuous Jacobian. The proposed method approximates the Jacobian of the blackbox function by forward finite differences with stepsizes depending on an estimate of its Lipschitz constant. Such estimate is also used in the definition of the trust-region radius, allowing natural update rules to enforce sufficient decrease of the objective function. We establish a worst-case complexity bound for the number of function evaluations that the method needs to find an approximate stationary point. Notably, the obtained bound depends linearly on the problem dimension, and quadratically on the inverse of the desired accuracy. In addition, if the components of the blackbox function are convex and the outer function is monotone, the previous worst-case evaluation complexity is improved to a linear dependence on the inverse of the desired accuracy. Numerical results are also reported, showing the relative efficiency of the new method with respect to state-of-the-art derivative-free solvers for composite nonsmooth optimization.

■ FD-07

Friday, 14:10 - 15:50 - Room: M:I

Optimal Transport for Machine Learning and Inverse Problems

Stream: Optimization for Inverse Problems and Machine Learning
Invited session
Chair: *Cheik Traoré*

1 - Gradient flows and kernelization in the Hellinger-Kantorovich (a.k.a. Wasserstein-Fisher-Rao) space
Jia-Jie Zhu

Motivated by applications of the optimal transport theory in optimization and machine learning, we present a principled investigation of gradient flow dissipation geometry, emphasizing the Hellinger (Fisher-Rao) type gradient flows and the connections with the Wasserstein space. The talk will introduce new advances in two directions: 1) the kernelization of Hellinger type distance and gradient flows, revealing precise connections with Stein flows, kernel discrepancies, and nonparametric regression; 2) new convergence results of the Hellinger-Kantorovich, a.k.a. Wasserstein-Fisher-Rao, gradient flows.

Joint work with Alexander Mielke.

2 - Learning Total-Variation Regularization Parameters via Weak Optimal Transport
Enis Chenchene, Kristian Bredies

We introduce a novel method for data-driven tuning of regularization parameters in total-variation denoising. The proposed approach leverages the semi-dual Brenier formulation of weak optimal transport between the distributions of clean and noisy images to devise a new loss function for total variation parameter learning. Our loss has a close connection to the traditional bilevel quadratic setting, but it leads to fully explicit monolevel problems, which are, in fact, convex under certain conditions. For training, we introduce a new conditional-gradient-type method, which can handle a complex and unbounded constraint set with computations up to numerical precision. Numerical experiments demonstrate the effectiveness of our approach and suggest promising avenues for future extensions.

3 - An Optimal Transport-based approach to Total-Variation regularization for the Diffusion MRI problem
Rodolfo Assereto

Diffusion Magnetic Resonance Imaging (dMRI) is a non-invasive imaging technique that draws structural information from the interaction between water molecules and biological tissues. Common ways of tackling the derived inverse problem include, among others, Diffusion Tensor Imaging (DTI), High Angular Resolution Diffusion Imaging (HARDI) and Diffusion Spectrum Imaging (DSI). However, these methods are structurally unable to recover the full diffusion distribution, only providing partial information about particle displacement. In our work, we introduce a Total-Variation (TV) regularization defined from an optimal transport perspective using 1-Wasserstein distances. Such a formulation produces a variational problem that can be handled by well-known algorithms enjoying good convergence properties, such as the primal-dual proximal method by Chambolle and Pock. It allows for the reconstruction of the complete diffusion spectrum from measured undersampled k/q space data.

4 - Application of the Opial property in Wasserstein spaces to inexact JKO schemes
Emanuele Naldi

The Opial property is a metric characterization of weak convergence for a suitable class of Banach spaces. It plays an important role in the study of weak convergence of iterates of mappings and of the asymptotic behavior of nets satisfying some metric properties. Since it involves only metric quantities, it is possible to define this property also in metric spaces provided with a suitable notion of weak convergence. This is the case for the space of probability measures endowed with the Kantorovich-Rubinstein-Wasserstein metric deriving by optimal transport. In this talk, we present the Opial property in the Wasserstein space of Borel probability measures with finite quadratic moment on a separable Hilbert space. We present applications of this property to convergence of sequences generated by the JKO scheme (the proximal point algorithm in Wasserstein spaces) when the functional is lower semicontinuous and convex along generalized geodesics. In practice, the computation of a proximal step in Wasserstein spaces can be challenging and it is often carried out only approximately. For this reason, we discuss various type of inexactness that can be introduced in the method, showing convergence and providing rates.

Advances in first-order optimization

Puya Latafat
KU LEuven
puya.latafat@kuleuven.be

Emanuel Laude
KU Leuven
emanuel.laude@esat.kuleuven.be

Panagiotis Patrinos
KU Leuven
panos.patrinos@esat.kuleuven.be

Track(s): 2

Bilevel optimization: strategies for complex decision-making

Alp Yurtsever
Umeå University
alp.yurtsever@umu.se

Track(s): 5

Challenges in nonlinear programming

Pavel Dvurechensky
Weierstrass Institute for Applied Analysis and Stochastics
pavel.dvurechensky@gmail.com

Jakub Marecek
Czech Technical University
jakub.marecek@fel.cvut.cz

Mathias Staudigl
Universität Mannheim
mathias.staudigl@uni-mannheim.de

Track(s): 6

Conic optimization: theory, algorithms and applications

Miguel Anjos
University of Edinburgh
anjos@stanfordalumni.org

Etienne De Klerk
Tilburg University
e.deklerk@uvt.nl

Track(s): 2

Developments in interior point methods

Utkarsh Detha
Danmarks Tekniske Universitet
utde@dtu.dk

Track(s): 2

Global optimization

Sonia Cafieri
ENAC - Ecole Nationale d'Aviation Civile
sonia.cafieri@enac.fr

Track(s): 7

In memory of Georg Still

Mirjam Duer
Augsburg University
mirjam.duer@math.uni-augsburg.de

Oliver Stein
Karlsruhe Institute of Technology
stein@kit.edu

Track(s): 3

Large-scale optimization

Pontus Giselsson
Lund University
pontusg@control.lth.se

Track(s): 4

Methods for non-/monotone inclusions and their applications

Dimitri Papadimitriou
Belgium Research Center (Leuven, Belgium), MCO Institute & Université libre de Bruxelles (ULB)
papadimitriou.dimitri.be@gmail.com

Track(s): 6

Multiobjective optimization

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

Track(s): 4

Nonsmooth and nonconvex optimization algorithms

Masoud Ahookhosh
University of Antwerp
masoud.ahookhosh@uantwerp.be

Sorin-Mihai Grad
ENSTA Paris
sorin-mihai.grad@ensta-paris.fr

Alireza Kabgani
University of Antwerp
alireza.kabgani@uantwerpen.be

Track(s): 5

Opening session

Track(s): 1

Optimal control and stochastic optimal control - theory, methods and applications

Gerhard-Wilhelm Weber
Poznan University of Technology
gerhard-wilhelm.weber@put.poznan.pl

Track(s): 4

Optimization applications

Track(s): 7

Optimization for Inverse Problems and Machine Learning

Enis Chenchene
enis.chenchene@gmail.com

Emanuele Naldi
Università di Genova
naldi@dima.unige.it

Cheik Traoré
University of Genoa
temci76@gmail.com

Track(s): 7

Optimization for learning

Pontus Giselsson
Lund University
pontusg@control.lth.se

Manu Upadhyaya
Lund University
manu.upadhyaya@control.lth.se

Track(s): 5

Optimization in neural architectures: convergence and solution characterization

Manish Krishan Lal
TU Munich
manish.krishanlal@tum.de

Maria-Luiza Vladarean
Ecole Polytechnique Federale de Lausanne
maria-luiza.vladarean@epfl.ch

Track(s): 3

Optimization in regression, classification and learning

Paula Amaral
Universidade Nova de Lisboa
paca@fct.unl.pt

Track(s): 4 7

Plenaries

Giancarlo Bigi
Universita’ di Pisa
giancarlo.bigi@unipi.it

Pontus Giselsson
Lund University
pontusg@control.lth.se

Track(s): 1

Randomized optimization algorithms

Laurent Condat
KAUST
laurent.condat@kaust.edu.sa

Track(s): 5

Tutorials

Track(s): 4

Variational analysis: theory and algorithms

Francisco Javier Aragón Artacho
University of Alicante
francisco.aragon@ua.es

Abderrahim Hantoute
Universidad de Alicante
hantoute@ua.es

Track(s): 3

Session Chair Index

Åkerman, Anton FB-04, FC-04
anton.akerman@control.lth.se
Department of Automatic Control, Lund University, Lund, Sweden

Alacaoglu, Ahmet FB-05
ahmetalacaoglu@gmail.com
University of Wisconsin-Madison, United States

Amaral, Paula TB-04, WC-04, FC-07
paca@fct.unl.pt
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Lisbon, Portugal

Andersen, Martin Skovgaard TC-02
mskan@dtu.dk
Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark

Anjos, Miguel WC-02
anjos@stanfordalumni.org
School of Mathematics, University of Edinburgh, United Kingdom

Aragón Artacho, Francisco Javier WF-03
francisco.aragon@ua.es
Mathematics, University of Alicante, Alicante, Spain

Bigi, Giancarlo TA-01, TE-01
giancarlo.bigi@unipi.it
Dipartimento di Informatica, Universita' di Pisa, Pisa, Italy

Bomze, Immanuel WD-02
immanuel.bomze@univie.ac.at
Dept. of Statistics and OR, University of Vienna, Vienna, Austria

Cafieri, Sonia TB-07, TC-07
sonia.cafieri@enac.fr
ENAC - Ecole Nationale d'Aviation Civile, Toulouse, France

Cesaroni, Edoardo WD-07
cesaroni@diag.uniroma1.it
Department of Computer, Control, and Management Engineering Antonio Ruberti, Università di Roma "La Sapienza", Roma, Italy, Italy

Chenchene, Enis TD-07
enis.chenchene@gmail.com
Austria

Condat, Laurent WE-05, WF-05
laurent.condat@kaust.edu.sa
AI Initiative, KAUST, Thuwal, Saudi Arabia

Davar, Dânâ FD-06
dana.davar@uclouvain.be
Applied Mathematics, Université catholique de Louvain, Belgium

Detha, Utkarsh TC-02
utde@dtu.dk
DTU Compute, Danmarks Tekniske Universitet, Copenhagen, Denmark

Duer, Mirjam FC-03, TC-03
mirjam.duer@math.uni-augsburg.de
Augsburg University, Augsburg, Germany

Dvurechensky, Pavel TB-06
pavel.dvurechensky@wias-berlin.de
Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Berlin, Germany

E.-Nagy, Marianna FB-06
marianna.eisenberg-nagy@uni-corvinus.hu
Corvinus University of Budapest, Hungary

Evens, Brecht FB-02
brecht.evens@kuleuven.be
Department of Electrical Engineering ESAT-STADIUS, KU Leuven, Leuven, Vlaams-Brabant, Belgium

Ghaderi, Susan TD-05
susan.ghaderi@kuleuven.be
ESAT, KU Leuven, Leuven, Flanders, Belgium

Giselsson, Pontus TE-01, WA-01, WB-01
pontusg@control.lth.se
Dept. of Automatic Control, Lund University, Sweden

Hantoute, Abderrahim TD-03
hantoute@ua.es
Mathematics, Universidad de Alicante, Alicante, Spain, Spain

Hendrikx, Hadrien FD-02
hadrien.hendrikx@inria.fr
Inria Grenoble, France

Hu, Feifei TC-06
fh17302@bristol.ac.uk
Department of Mathematics, University of Bristol, United Kingdom

Kabgani, Alireza FD-05
alireza.kabgani@uantwerpen.be
Department of Mathematics, University of Antwerp, Antwerp, BELGIUM, Belgium

Kamri, Yassine TB-02
yassinekamri1996@gmail.com
INMA, UCLouvain, Louvain-la-Neuve, Belgium

Kovacova, Gabriela TD-04
kovacova@ucla.edu
Department of Mathematics, University of California, Los Angeles, Los Angeles, California, United States

Krishan Lal, Manish WC-03, WD-03
manish.krishanlal@tum.de
TU Munich, Germany

Laude, Emanuel FC-02
emanuel.laude@esat.kuleuven.be
KU Leuven, Belgium

Lämmel, Sebastian FC-07
sebastian.laemmel@mathematik.tu-chemnitz.de
Chemnitz University of Technology, Chemnitz, Germany

Maji, Atanu WE-07
maji@diag.uniroma1.it
Department of Computer, Automatic and Management Engineering, Sapienza University of Rome, Rome, Lazio, Italy

Martínez-Legaz, Juan Enrique FD-03
juanenrique.martinez.legaz@uab.cat
Departament d'Economia, Universitat Autònoma de Barcelona, Bellaterra, Spain

Naldi, Emanuele WF-07
naldi@dima.unige.it
Mathematics, Università di Genova, Genova, Italy

Nilsson, Max WE-04, TB-05
max.nilsson@control.lth.se
Department of Automatic Control, Lund University, Sweden

Papadimitriou, Dimitri WC-06, WD-06
papadimitriou.dimitri.be@gmail.com
Belgium Research Center (Leuven, Belgium), MCO Institute & Université libre de Bruxelles (ULB), Leuven, Belgium

Ricci, Jacopo Maria WC-07
jacopomaria.ricci@uniroma3.it
Economia Aziendale, Università degli Studi Roma Tre, Rome, Italy, Italy

Shulgin, Egor WF-05
egor.shulgin@kaust.edu.sa
KAUST AI Initiative, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

Staudigl, Mathias WE-06
mathias.staudigl@uni-mannheim.de
Department of Mathematics, Universität Mannheim, Mannheim, Germany

Stein, Oliver TE-01, FB-03, TB-03
stein@kit.edu
Institute of Operations Research, Karlsruhe Institute of Technology, Karlsruhe, Germany

Stillfjord, Tony TD-06
tony.stillfjord@math.lth.se
Centre for Mathematical Sciences, Lund University, Lund, Sweden

Strugariu, Radu WF-04
rstrugariu@tuiasi.ro
Mathematics, Gheorghe Asachi Technical University & Octav Mayer Institute of Mathematics, Iasi, Iasi, Romania

Taylor, Adrien FD-02, TD-02, WE-02, WF-02
adrien.taylor@inria.fr
Inria/ENS, France

Traoré, Cheik FD-07
temci76@gmail.com
University of Genoa, Genova, Liguria, Italy

Ulus, Firdevs TC-04
firdevs@bilkent.edu.tr
Industrial Engineering, Bilkent University, Ankara, Turkey

Upadhyaya, Manu . . . TD-02, WE-02, WF-02, WC-05, WD-05
manu.upadhyaya@control.lth.se
Lund University, Sweden

Vladarean, Maria-Luiza WC-03, WD-03
maria-luiza.vladarean@epfl.ch
IC, Ecole Polytechnique Federale de Lausanne, Lausanne, Vaud, Switzerland

Weber, Gerhard-Wilhelm FD-04
gerhard-wilhelm.weber@put.poznan.pl
Faculty of Engineering Management, Poznan University of Technology, Poznan, Poland

Weissmann, Simon WF-06
simon.weissmann@uni-mannheim.de
Department of Mathematics, Universität Mannheim, Germany

Yang, Xiaoqi WE-03
xiao.qi.yang@polyu.edu.hk
The Hong Kong Polytechnic University, Hong Kong, Hong Kong

Yurtsever, Alp FA-01, FB-05, FC-05, TC-05
alp.yurtsever@umu.se
Umeå University, Sweden

Zhang, Yi WD-04
yzhang@minlp.com
The Optimization Firm, Hangzhou, Zhejiang, China

Author Index

Abbaszadehpeivasti, Hadi WF-02
h.abbaszadehpeivasti@tilburguniversity.edu
Tilburg University, Tilburg, North Brabant, Netherlands

Absil, P.-a. WD-05
pa.absil@uclouvain.be
UCLouvain, Louvain-la-Neuve, Belgium

Åkerman, Anton WF-02
anton.akerman@control.lth.se
Department of Automatic Control, Lund University, Lund, Sweden

Ahmed, Faizan TC-03
faizan.ahmed@utwente.nl
Computer Science, University of Twente, Enschede, Netherlands

Ahookhosh, Masoud FD-05, TD-05
masoud.ahookhosh@uantwerp.be
Department of Mathematics, University of Antwerp, Antwerp, Belgium

Alacaoglu, Ahmet WC-06
ahmetalacaoglu@gmail.com
University of Wisconsin-Madison, United States

Allerbo, Oskar TB-04
oallerbo@kth.se
Mathematical Statistics, KTH, Stockholm, Sweden

Amaral, Paula WC-04
paca@fct.unl.pt
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Lisbon, Portugal

Andersen, Martin Skovgaard FB-05
mskan@dtu.dk
Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark

Anjos, Miguel WC-02
anjos@stanfordalumni.org
School of Mathematics, University of Edinburgh, United Kingdom

Antonakopoulos, Kimon FC-05
kimon.antonakopoulos@epfl.ch
EPFL, Switzerland

Aragón Artacho, Francisco Javier WF-03
francisco.aragon@ua.es
Mathematics, University of Alicante, Alicante, Spain

Ararat, Cagin WC-07
cararat@bilkent.edu.tr
Industrial Engineering, Bilkent University, Ankara, Turkey

Aspman, Johannes WF-06
johannes.aspman@cvut.cz
Computer Science, Czech Technical University, Kärna, Sweden

Assereto, Rodolfo FD-07
assereto.rodolfo@gmail.com
Department of Mathematics and Scientific Computing, University of Graz, Graz, Austria

Atenas, Felipe WF-03
felipe.atenas@unimelb.edu.au
University of Melbourne, Australia

Awari, Atharva FC-04
atharvaabhijit.awari@umons.ac.be
Mathematics and Operations Research, University of Mons, Belgium, Mons, Hainaut, Belgium

Babaie-Kafaki, Saman FB-06
saman.babaiekafaki@unibz.it
Faculty of Engineering, Free University of Bozen-Bolzano, Italy

Baldacci, Roberto WE-07
rbaldacci@hbku.edu.qa
College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar

Baltas, Ioannis FD-04
jmpaltas@aegean.gr
Financial and Management Engineering, University of the Aegean, Chios isl., Greece

Banert, Sebastian TD-02, WF-02
banert@uni-bremen.de
Uni Bremen, Germany

Banert, Sebastian WE-02
sebastian.banert@control.lth.se
Department of Automatic Control, Lund University, Lund, Sweden

Banholzer, Dirk FB-03
dirk.banholzer@soton.ac.uk
School of Mathematics, University of Southampton, United Kingdom

Beck, Amir TA-01, TD-05
becka@tauex.tau.ac.il
School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel

Behrens, Freya WD-03
freya.behrens@epfl.ch
EPFL, Switzerland

Bigi, Giancarlo FC-02
giancarlo.bigi@unipi.it
Dipartimento di Informatica, Universita' di Pisa, Pisa, Italy

Birdal, Tolga TB-06
tbirdal@imperial.ac.uk
Computing, Imperial College London, London, London, United Kingdom

Bock, Andreas TC-02
aasbo@dtu.dk
Copenhagen, Denmark

Bomze, Immanuel WD-02
immanuel.bomze@univie.ac.at
Dept. of Statistics and OR, University of Vienna, Vienna, Austria

Bortz, Michael WE-07
michael.bortz@itwm.fraunhofer.de
Optimization, Fraunhofer ITWM, Kaiserslautern, Deutschland, Germany

Bot, Radu Ioan FB-04
radu.bot@univie.ac.at
Faculty of Mathematics, University of Vienna, Vienna, Austria

Bousselmi, Nizar WE-02
niz.bousselmi@gmail.com
ICTEAM, UCLouvain, Wavre, Belgium

Bouza Allende, Gemayqzel TC-03
gema@matcom.uh.cu
Cuba

Bredies, Kristian FD-07
kristian.bredies@uni-graz.at
Institute for Mathematics and Scientific Computing, University of Graz, Graz, Austria

Cafieri, Sonia TC-07
sonia.cafieri@enac.fr
ENAC - Ecole Nationale d'Aviation Civile, Toulouse, France

Camacho, Jesús TD-03
j.camacho@umh.es
Center of Operations Research, Miguel Hernández University of Elche, Elche, Spain

Campoy, Rubén WF-03
ruben.campoy@ua.es
Department of Mathematics, Universidad de Alicante, San Vicente del Raspeig, Alicante, Spain

Cascarano, Pasquale WE-05
pasquale.cascarano2@unibo.it
University of Bologna, Italy

Cassimiro de Oliveira, Nicole Cristina WD-06
nicole@ime.unicamp.br
Computational & Applied Mathematics, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil

Castera, Camille TD-07
camille.castera@protonmail.com
Department of Mathematics, University of Tübingen, Tübingen, Germany

Cesarone, Francesco WC-07
francesco.cesarone@uniroma3.it
Department of Business Studies, Roma Tre University, Rome, Italy

Cesaroni, Edoardo WD-07
cesaroni@diag.uniroma1.it
Department of Computer, Control, and Management Engineering Antonio Ruberti, Università di Roma "La Sapienza", Roma, Italy, Italy

Cevher, Volkan FC-02, FC-05
volkan.cevher@epfl.ch
School of Engineering, EPFL, Lausanne, Switzerland

Chen, Chien-Hsin TD-04
chienhsinch@gmail.com
Computer Science, National Tsing Hua University, Taiwan

Chen, Lujing FB-05
lujc@dtu.dk
Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Region Hovedstaden, Denmark

Chen, Po-An TD-04
poanch@gmail.com
Institute of Information Management, National Chiao Tung University, Taiwan

Chen, Tianshi FB-05
tschen@cuhk.edu.cn
The Chinese University of Hong Kong, Shenzhen, China

Chen, Yuwen TC-02
yuwen.chen@eng.ox.ac.uk
Engineering Science, University of Oxford, Oxford, Oxfordshire, United Kingdom

Chenchene, Enis WF-02, FD-07
enis.chenchene@gmail.com
Austria

Chirinos Rodriguez, Jonathan WF-07
jonathanchirinosrodriguez@gmail.com
Mathematics, University of Genoa, Genova, Liguria, Italy

Chouzenoux, Emilie WE-04
emilie.chouzenoux@univ-mlv.fr
Université Paris-Est Marne-La-Vallée, Marne La Vallee, France

Cialenco, Igor TD-04
cialenco@iit.edu
Illinois Institute of Technology, Chicago, United States

Ciccarelli, Fabio WE-07
f.ciccarelli@uniroma1.it
Department of computer, control and management engineering (DIAG), Sapienza University of Rome, Roma, Italy, Italy

Cohen, Jeremy TB-04
jeremy.cohen@cnrs.fr
CNRS, France

Condat, Laurent WF-05
laurent.condat@kaust.edu.sa
AI Initiative, KAUST, Thuwal, Saudi Arabia

Coniglio, Stefano WE-07
stefano.coniglio@gmail.com
Mathematical Sciences, University of Southampton, Southampton, United Kingdom

Cortild, Daniel FB-04
d.cortild@rug.nl
Faculty of Science and Engineering, University of Groningen, Groningen, Netherlands

Cosson, Romain WD-05
romain.cosson@outlook.fr
Inria, France

Criscitiello, Christopher FD-02
christopher.criscitiello@epfl.ch
Mathematics, EPFL, Saint Sulpice, Vaud, Switzerland

Croella, Anna Livia WE-07
annalivia.croella@uniroma1.it
Sapienza University of Rome, Rome, Lazio, Italy, Italy

Cui, Hugo WD-03
hugo.cui@epfl.ch
EPFL, Switzerland

da Silva, Joakim WF-04
Joakim.daSilva@elekta.com
Elekta, Sweden

Dao ngoc, Minh WF-03
daonminh@gmail.com
Federation University Australia, sydny, Australia

Darvay, Zsolt FB-06
zsolt.darvay@ubbcluj.ro
Department of Mathematics and Computer Science of the Hungarian Line, Babes-Bolyai University, Cluj-Napoca, Romania

Davar, Dânâ FD-06
dana.davar@uclouvain.be
Applied Mathematics, Université catholique de Louvain, Belgium

De Klerk, Etienne TB-03
e.deklerk@uvt.nl
Econometrics and OR, Tilburg University, Tilburg, Netherlands

De Lara, Michel FD-03
michel.delara@enpc.fr
École des Ponts ParisTech, Marne la Vallée Cedex, France

De Marchi, Alberto FB-04
alberto.demarchi@unibw.de
University of the Bundeswehr Munich, Neubiberg, Germany

de Oliveira, Welington WF-03
welington.oliveira@minesparis.psl.eu
CMA, Mines Paris PSL, Sophia Antipolis, France

de Vicente, Daniel WD-02
a12032143@unet.univie.ac.at
University of Vienna, Austria

Dias, Tiago WC-04
tme.dias@campus.fct.unl.pt
Department of Mathematics, NOVA School of Science and Technology | FCT NOVA, Caparica, Setúbal, Portugal

Dieuleveut, Aymeric TD-02, TB-05
aymeric.dieuleveut@gmail.com
CMAP (Applied Maths), Ecole Polytechnique, Institut Polytechnique de Paris, France

Dieuleveut, Aymeric TD-02
aymeric.dieuleveut@polytechnique.edu
CMAP, Ecole Polytechnique, Paris, France

Difonzo, Fabio V. WF-06
fabiovito.difonzo@cnr.it
IAC, CNR, Italy

Dixit, Avinash FD-06
avinash@kmc.du.ac.in
Mathematics, Kirori Mal College, University of Delhi, Delhi, Delhi, India

Dobrovoczki, Peter WD-06
pdobrovoczki@sztaki.hu
Engineering and Management Intelligence Research Laboratory, Hun-ren Sztaki, Hungary

Doikov, Nikita WE-06
nikita.doikov@epfl.ch
EPFL, Lausanne, Switzerland

Dragomir, Radu-Alexandru FD-02
dragomir@telecom-paris.fr
Telecom Paris, France

Duer, Mirjam TB-03, TC-03, FC-07
mirjam.duer@math.uni-augsburg.de
Augsburg University, Augsburg, Germany

Durea, Marius WF-04
durea@uaic.ro
Mathematics, University Al. I. Cuza and Octav Mayer Institute of Mathematics, Iasi, Iasi, Romania

Dvurechensky, Pavel WE-06
pavel.dvurechensky@gmail.com
Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany

Dvurechensky, Pavel TB-06
pavel.dvurechensky@wias-berlin.de
Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Berlin, Germany

E.-Nagy, Marianna FB-06
marianna.eisenberg-nagy@uni-corvinus.hu
Corvinus University of Budapest, Hungary

Ehrhardt, Matthias J. FC-05
matthias.ehrhardt@bath.ac.uk
University of Bath, United Kingdom

Eichfelder, Gabriele TE-01, TC-04
gabriele.eichfelder@tu-ilmenau.de
Institute of Mathematics, Technische Universität Ilmenau, Ilmenau, Germany

Eisenmann, Monika TD-06
monika.eisenmann@math.lth.se
Centre for Mathematical Sciences, Lund University, Sweden

Elvira, Víctor WE-04
victor.elvira@ed.ac.uk
School of Mathematics, University of Edinburgh, United Kingdom

Evens, Brecht FC-02
brecht.evens@kuleuven.be
Department of Electrical Engineering ESAT-STADIUS, KU Leuven, Leuven, Vlaams-Brabant, Belgium

Fajardo, Maria Dolores TD-03
md.fajardo@ua.es
Mathematics, University Of Alicante, Alicante, Spain

Fatkhullin, Ilyas WF-06
ilyas.fn979@gmail.com
Computer Science, ETH Zürich, Zurich, Switzerland

Fincato, Ettore TD-06
zc21901@bristol.ac.uk
Mathematics, University of Bristol, United Kingdom

Fliege, Joerg FB-03
J.Fliege@soton.ac.uk
University of Southampton, United Kingdom

Flinth, Axel WC-03
axel.flinth@umu.se
Mathematics and Mathematical Statistics, Umeå University, Sweden

Florea, Mihai I. FC-04
mihai.florea@uclouvain.be
Mathematical Engineering, Université catholique de Louvain, Louvain-la-Neuve, Belgium

Forsgren, Anders TC-02
andersf@kth.se
Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden

Fortunato Filho, Fabio WC-07
f235849@dac.unicamp.br
Department of Applied MathematicsIMECC, The State University of Campinas (UNICAMP), Campinas, SP, Brazil

Franchini, Giorgia WE-05
giorgia.franchini@unimore.it
University of Modena and Reggio Emilia, Modena, Italy

Frank, Martin WC-04
martin.frank@kit.edu
Karlsruhe Institute of Technology (KIT), Germany

Furini, Fabio WE-07
fabio.furini@uniroma1.it
DIAG, La Sapienza, Roma, Please Select, Italy

G.-Tóth, Boglárka TB-07
boglarka@inf.szte.hu
Department of Computational Optimization, University of Szeged, Szeged, Hungary

Gabl, Markus WD-02
markus@gabl.com
Austria

Gally, Tristan TC-07
tristangally@fico.com
Fair Isaac Services Limited, Birmingham, West Midlands, United Kingdom

Gao, Bin WD-05
gaobin@lsec.cc.ac.cn
Chinese Academy of Sciences, Beijing, Beijing, China

Gautam, Pankaj FD-03
pgautam908@gmail.com
Applied Mathematics and Scientific Computing, Indian Institute of Technology, Roorkee, Saharanpur, Uttar Pradesh, India

Gencsi, Mihály TB-07
gencsi@inf.u-szeged.hu
Computational Optimization, University of Szeged, Szeged, Hungary

Ghaderi, Susan TD-05
susan.ghaderi@kuleuven.be
ESAT, KU Leuven, Leuven, Flanders, Belgium

Ghitri, Mohammed Ibrahim Abdelkayoum TD-03
moghitri@gmail.com
Mathematics, Universidad de Alicante, Alicante, Alicante, Spain

Gidel, Gauthier WE-02
gauthier.gidel@gmail.com
Computer Science, Mila, Université de Montréal, Canada

Gillis, Nicolas TB-04, WC-05
nicolas.gillis@umons.ac.be
Mathematics and Operational Research, Université de Mons, Mons, Belgium

Giselsson, Pontus WE-02, WF-02, WC-05
pontusg@control.lth.se
Dept. of Automatic Control, Lund University, Sweden

Glineur, François TB-02, TD-02, WE-02, WF-02, WC-07
francois.glineur@uclouvain.be
Icteam/inma & Core, Université catholique de Louvain (UCLouvain), Louvain-la-Neuve, Belgium

Goberna, Miguel FC-03
mgoberna@ua.es
Departamento de Matemáticas, Universidad de Alicante, San Vicente del Raspeig, Alicante, Spain

Gorbunov, Eduard WE-02
eduard.gorbunov@mbzuai.ac.ae
Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dabi, United Arab Emirates

Goujaud, Baptiste TD-02
baptiste.goujaud@gmail.com

CMAP, Ecole Polytechnique, Palaiseau, IDF, France

Goujaud, Baptiste TD-02
baptiste.goujaud@polytechnique.edu
CMAP, Ecole Polytechnique, Palaiseau, France

Grapiglia, Geovani FD-06
geovani.grapiglia@uclouvain.be
Applied Mathematics, Université catholique de Louvain, LOUVAIN-LA-NEUVE, Belgium

Grasmair, Markus FD-03
markus.grasmair@ntnu.no
Mathematical Sciences, Norwegian University of Science and Technology, Norway, Trondheim, Norway

Gruhlke, Robert WF-06
r.gruhlke@fu-berlin.de
FU Berlin, Germany

Guenzel, Harald FB-03
guenzel@mathc.rwth-aachen.de
Dept. of Mathematics (C), RWTH Aachen University, Aachen, Germany

Guerra Pires, Arthur Leandro WE-07
a.guerrapires@hotmail.com
Post-Graduate Program in Production Engineering and Systems - Federal University of Paraiba, Federal University of Paraiba, João Pessoa, Paraíba, Brazil

Guilmeau, Thomas WE-04
thomas.guilmeau@inria.fr
Université Paris-Saclay, CentraleSupélec, INRIA, France

Hallak, Nadav TD-05
ndvhllk@technion.ac.il
The Technion, Haifa, Israel

Hamed, Mohammad FD-05
mohammad.hamed@uantwerpen.be
Mathematics, University of Antwerpen, Antwerpen, Belgium

Hantoute, Abderrahim TD-03
hantoute@ua.es
Mathematics, Universidad de Alicante, Alicante, Spain, Spain

Hanzely, Slavomír WF-05
shanzely@gmail.com
Machine Learning, MBZUAI, Abu Dhabi, United Arab Emirates

He, Niao WF-06
niaohe@illinois.edu
Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States

Heeman, Pim TC-02
pimh@kth.se
Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden

Hendrickx, Julien TB-02, WE-02
julien.hendrickx@uclouvain.be
ICTEAM, Louvain-la-Neuve, Belgium

Hendrikx, Hadrien FD-02
hadrien.hendrikx@inria.fr
Inria Grenoble, France

Hernandez Escobar, Daniel WF-04
daniel.hernandez.escobar@it.uu.se
Department of Information Technology, Uppsala University, Uppala, Sweden

Hernandez Escobar, Daniel FB-03
Daniel.Hernandez@uib.no
Department of Informatics, University of Bergen, Bergen, Norway

Hladik, Milan WD-06
hladik@kam.mff.cuni.cz
Department of Applied Mathematics, Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic

Hoff, Daniel TC-04
daniel.hoff@tu-ilmenau.de
Institiute of Mathematics, Technische Universität Ilmenau, Ilmenau, Germany

Hon, Wing-Kai TD-04
wkhon@cs.nthu.edu.tw
Computer Science, National Tsing Hua University, Hsinchu, Hsinchu, Taiwan

Hong, Mingyi FC-05
mhong@umn.edu
University of Minnesota, United States

Horvath, Samuel WE-02
samuel.horvath@mbzuai.ac.ae
Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dabi, United Arab Emirates

Houta, Zakaria WD-07
houta@laplace.univ-tlse.fr
ENSEEIHT, Laplace laboratory, Toulouse, France

Hu, Feifei TC-06
fh17302@bristol.ac.uk
Department of Mathematics, University of Bristol, United Kingdom

Hu, Yaohua WE-03
hyh19840428@163.com
School of Mathematical Sciences, Shenzhen University, Shenzhen, Guangdong Province, China

Hu, Yifan WF-06
yifan.hu@epfl.ch
EPFL, Switzerland

Huguet, Thomas WD-07
huguet@laplace.univ-tlse.fr
LAPLACE Laboratory, Toulouse, France

Janda, Karel WE-07
karel-janda@seznam.cz
Charles University, Praha, Czech Republic

Jaramillo Villalba, Catalina TC-06
c230186@dac.unicamp.br
Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas, Campinas, SP, Brazil

Jargalsaikhan, Bolor TC-03
bolor.jargalsaikhan@gmail.com
Ulaanbaatar, Mongolia

Jäkel, Oliver WC-04
o.jaekel@dkfz-heidelberg.de
German Cancer Research Center (DKFZ), Germany

Kabgani, Alireza TD-05
alireza.kabgani@uantwerpen.be
Department of Mathematics, University of Antwerp, Antwerp, BELGIUM, Belgium

Kalaycı, Betül FD-04
betulkalayci@gmail.com
Financial Mathematics, Institute of Applied Mathematics, Ankara, Turkey

Kamri, Yassine WC-07
yassine.kamri@uclouvain.be
INMA, UCLouvain, Louvain-La-Neuve, Belgium

Kamri, Yassine TB-02
yassinekamri1996@gmail.com
INMA, UCLouvain, Louvain-la-Neuve, Belgium

Kamzolov, Dmitry WE-06
kamzolov.dmitry@mbzuai.ac.ae
Machine Learning, Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dabi, United Arab Emirates

Karagulyan, Avetik WD-05
avetik.karagulyan@kaust.edu.sa
Artificial Intelligence Initiative, King Abdullah University of Science and Technology, Thuwal, Makkah, Saudi Arabia

Karsu, Ozlem TC-04
ozlemkarsu@bilkent.edu.tr
Industrial Engineering, Bilkent University, Ankara, Turkey

Kasimbeyli, Refail FD-05
refail.kasimbeyli@gmail.com
Industrial Engineering, Eskisehir Technical University, Eskisehir, Turkey

Ketkov, Sergei FB-05
sergei.ketkov@business.uzh.ch
Department of Business Administration, University of Zurich, Zurich, Switzerland

Kim, Chesoong TC-06
dowoo@sangji.ac.kr
Department of Business Administration, Sangji University, Wonju, Gangwon-do, Korea, Republic of

Kim, Donghwan WC-06
donghwankim@kaist.ac.kr
KAIST, Korea, Republic of

Kim, Jaeyeon WC-05
kjy011102@snu.ac.kr
Seoul National University, Korea, Republic of

Kim, Jungbin FD-02
kjb2952@snu.ac.kr
Seoul National University, Korea, Republic of

Kis, Tamas WD-06
kis.tamas@sztaki.mta.hu
Institute for Computer Science and Control, Budapest, Hungary

Klein, Sara WF-06
sara.klein@uni-mannheim.de
Mathematics Institute, University of Mannheim, Mannheim, Germany

Kobler, Erich WE-05
kobler@uni-bonn.de
University of Bonn, Germany

Kocuk, Burak WC-02
burakkocuk@sabanciuniv.edu
Industrial Engineering, Sabanci University, Istanbul, Turkey

Kovacova, Gabriela TD-04
kovacova@ucla.edu
Department of Mathematics, University of California, Los Angeles, Los Angeles, California, United States

Krishan Lal, Manish WC-03
manish.krishanlal@tum.de
TU Munich, Germany

Krzakala, Florent WD-03
florent.krzakala@epfl.ch
EPFL, Switzerland

Küfer, Karl-Heinz WE-07
kuefer@itwm.fhg.de
Optimization, Fraunhofer ITWM, Kaiserslautern, Germany

Kungurtsev, Vyacheslav WF-06
vyacheslav.kungurtsev@fel.cvut.cz
Czech Technical University, Czech Republic

Labarrière, Hippolyte TD-07
hippolyte.labarriere@edu.unige.it
Università di Genova, Italy

Lampariello, Lorenzo FB-02
lorenzo.lampariello@uniroma3.it
Roma Tre University, Italy

Latafat, Puya FB-02, FC-02
puya.latafat@kuleuven.be
Electrical Engineering (ESAT) STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU LEuven, Leuven, Belgium

Lämmel, Sebastian FB-03, FC-07
sebastian.laemmel@mathematik.tu-chemnitz.de
Chemnitz University of Technology, Chemnitz, Germany

Le Digabel, Sébastien TC-04
sebastien.le-digabel@polymtl.ca
Polytechnique Montréal, Montréal, QC, Canada

Lebbe, Nicolas WD-07
lebbe@laplace.univ-tlse.fr
LAPLACE Laboratory, Toulouse, France

Lee, Jon TB-07
jonxlee@umich.edu
Industrial and Operations Engineering Department, University of Michigan, Ann Arbor, Michigan, United States

Lesaja, Goran FB-06
goran@georgiasouthern.edu
Mathematical Sciences, Georgia Southern University, Statesboro, Georgia, United States

Li, Hanmin WD-05
hanmin.li@kaust.edu.sa
Artificial Intelligence Initiative, King Abdullah University of Science and Technology, Thuwal, Makkah, Saudi Arabia

Li, Xiang WF-05
xiang.li@inf.ethz.ch
Department of Computer Science, ETH Zurich, Switzerland

Li, Zhening TD-05
zheningli@gmail.com
Department of Mathematics, University of Portsmouth, Portsmouth, United Kingdom

Loizou, Nicolas WE-02
nicolasloizou1@gmail.com
Department of Computer Science and Operations Research, Mila - Quebec Artificial Intelligence Institute & University of Montreal, Montreal, Quebec, Canada

López-Cerdá, Marco A. FC-03
marco.antonio@ua.es
Statistics and Operations Research, Alicante University, Alicante, Spain

Ludovic, Stephan TB-05
ludovic.stephan@epfl.ch
IDEPHICS, EPFL, Lausanne, Vaud, Switzerland

Macedo, Eloisa WC-02
macedo@ua.pt
Department of Mechanical Engineering, University of Aveiro, Aveiro, Aveiro, Portugal

Madani, Mehdi WC-07
mma@n-side.com
N-SIDE, Louvainn-la-Neuve, Belgium

Maier, Severin TD-07
sevemaier98@googlemail.com
University of Tübingen, Germany

Maji, Atanu WE-07
maji@diag.uniroma1.it
Department of Computer, Automatic and Management Engineering, Sapienza University of Rome, Rome, Lazio, Italy

Malha, Rui WC-04
rj.malha@campus.fct.unl.pt
Mathematic, Sst-fct Nova, CAPARICA, Portugal

Malo, Pekka TC-05
pekka.malo@aalto.fi
Information and Service Economy, Aalto University School of Business, Helsinki, Finland

Mansueto, Pierluigi TB-04
pierluigi.mansueto@unifi.it
Department of Information Engineering, University of Florence, Florence, Italy

Marecek, Jakub TB-06, WF-06
jakub.marecek@fel.cvut.cz
Dept. of Computer Science, Czech Technical University, Prague, Czech Republic

Mariosa, Raffaele WC-04
mariosa@diag.uniroma1.it
DIAG, Sapienza, University of Rome, Rome, Italy, Italy

Martínez, José Mario WC-07
martinez@ime.unicamp.br
Dept. Applied Mathematics, University of Campinas, Campinas, SP, Brazil

Martínez-Legaz, Juan Enrique FD-03
juanenrique.martinez.legaz@uab.cat
Departament d'Economia, Universitat Autònoma de Barcelona, Bellaterra, Spain

Maskan, Hoomaan FD-06
hoomaanhezaveh@yahoo.com
Umeå University, Umeå, Sweden

Mavrothalassitis, Ioannis FC-02
ioannis.mavrothalassitis@epfl.ch
EPFL, Switzerland

McRae, Andrew WE-04
andrew.mcrae@epfl.ch
Institute of Mathematics, EPFL, Switzerland

Messine, Frédéric WD-07
Frederic.Messine@laplace.univ-tlse.fr
LAPLACE-ENSEEIHT-INPT University of Toulouse, Toulouse, France

Mimouni, Daniel WF-03
daniel.mimouni@minesparis.psl.eu
IFP Energies Nouvelles, France

Mößner, Nikola WD-07
nikola.moessner@ier.uni-stuttgart.de
Institute of Energy Economics and the Rational Use of Energy, Universität Stuttgart, Stuttgart, Germany

Mogalle, David WE-07
david.mogalle@itwm.fraunhofer.de
Optimization, Fraunhofer ITWM, Kaiserslautern, Germany

Molinari, Cesare WF-07

cecio.molinari@gmail.com
Università di Genova, Genova, Italy

Molinari, Cesare WE-05
molinari@dima.unige.it
Università di Genova, Genova, GE, Italy

Mongeau, Marcel TC-07
marcel.mongeau@enac.fr
ENAC, Toulouse, France

Moreau, Yves TD-05
Yves.Moreau@esat.kuleuven.be
KU Leuven, Leuven, Belgium

Morini, Benedetta WE-05
benedetta.morini@unifi.it
Dipartimento di Ingegneria Industriale, Universita di Firenze, Firenze, Italy

Nadisic, Nicolas TB-04
nicolas.nadisic@ugent.be
Ghent University and Royal Institute for Cultural Heritage (KIK-IRPA), Gent, Belgium

Naldi, Emanuele WF-02, FD-07
naldi@dima.unige.it
Mathematics, Università di Genova, Genova, Italy

Narushima, Yasushi FB-05
narushima@keio.jp
Department of Industrial and Systems Engineering, Keio University, Yokohama, Kanagawa, Japan

Nemecek, Jiri WF-06
nemecj38@fel.cvut.cz
Czech Technical University in Prague, Czech Republic

Nguyen, Dang-Khoa FB-04
dang-khoa.nguyen@univie.ac.at
Ho Chi Minh City University of Science, Viet Nam

Nishioka, Akatsuki FC-04
akatsuki_nishioka@mist.i.u-tokyo.ac.jp
Department of Mathematical Informatics, The University of Tokyo, Japan

Nouy, Anthony WF-06
anthony.nouy@ec-nantes.fr
Centrale Nantes & Nantes Université, France

Nowak, Ivo TB-07
ivo.nowak@haw-hamburg.de
HAW Hamburg, Hamburg, Germany

Ochs, Peter TD-07
ochs@cs.uni-saarland.de
Saarland University, Germany

Okuno, Takayuki FB-06
takayuki-okuno@st.seikei.ac.jp
Seikei University, Tokyo, Japan

Oliveira, Aurelio TC-06, WD-06
aurelio@ime.unicamp.br
Computational & Applied Mathematics, UNICAMP, Campinas, SP, Brazil

Ortkamp, Tim WC-04
tim.ortkamp@kit.edu
Karlsruhe Institute of Technology (KIT), Germany

Özen, Buket WC-02
bozen@sabanciuniv.edu
Industrial Engineering, Sabancı University, Istanbul, Istanbul, Turkey

Palagi, Laura WC-04, WE-07
palagi@diag.uniroma1.it
Department of Computer, Control, and Management Engineering A. Ruberti, Sapienza University of Rome, Rome, Italy

Pan, Shaohua WE-03
shhpan@scut.edu.cn
School of Mathematics, South China University of Technology, Guangzhou, Tianhe district, China

Panahi, Ashkan WF-02
ashkan.panahi@chalmers.se
Chalmers University of Technology, Sweden

Papadimitriou, Dimitri WC-06
dpapadimitriou@3nlab.org
3nLab, Leuven, Belgium

Patrinos, Panagiotis FB-02, FC-02, WF-02
panos.patrinos@esat.kuleuven.be
Electrical Engineering, KU Leuven, Leuven, Belgium

Paul, Malisani WF-03
paulmalisani@ifpen.fr
IFP Energies Nouvelles, France

Peng, Bo WD-02
bo.peng@univie.ac.at
University of Vienna, Vienna, Vienna, Austria

Pethick, Thomas FC-02
thomas.pethick@epfl.ch
EPFL, Switzerland

Petra, Stefania FD-02
stefania.petra@uni-a.de
Heidelberg University, Germany

Peypouquet, Juan FB-04
j.g.peypouquet@rug.nl
Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, Netherlands

Peyré, Gabriel WB-01
gabriel.peyre@ens.fr
CNRS and École Normale Supérieure, Paris, Paris, France

Phan, Vuong FD-05
t.v.phan@soton.ac.uk
University of Southampton, Southampton, United Kingdom

Philippenko, Constantin ... TB-05
cphilippenko@gmail.com
Inria Paris, Paris, France

Pinar, Mustafa ... WC-07
mustafap@bilkent.edu.tr
Department of Industrial Engineering, Bilkent University, Ankara, Turkey

Pinta, Titus ... WE-05
t.pinta@math.uni-goettingen.de
Mathematics, Goettingen University, Germany

Pintea, Cornel ... FD-03
cpintea@math.ubbcluj.ro
Babes Bolyai University, Romania

Poon, Clarice ... TC-05
clarice.poon@warwick.ac.uk
Mathematics Institute, University of Warwick, United Kingdom

Porcelli, Margherita ... WC-05
margherita.porcelli@unibo.it
Dipartimento di Matematica, Università di Bologna, Italy

Porta, Federica ... WE-05
federica.porta@unimore.it
Universita' di Modena e Reggio Emilia, Italy

Prakhya, Karthik ... TB-06
karthik.prakhya@umu.se
Umeå University, Umeå, Sweden

Purutcuoglu, Vilda ... FD-04
vpurutcu@metu.edu.tr
Statistics, Middle East Technical University, Ankara, Turkey

Qiu, Yuzhou ... WD-02
y.qiu-16@sms.ed.ac.uk
School of Mathematics, The University of Edinburgh, Edinburgh, Scotland, United Kingdom

Rahimi, Morteza ... FD-05
morteza.rahimi@uantwerpen.be
Mathematics, University of Antwerp, Antwerp, Belgium

Rando, Marco ... WF-07
marco.rando0396@gmail.com
Malga - DIBRIS, Universita degli Studi di Genova, Genova, Italy

Razin, Noam ... WD-03
noamrazin@mail.tau.ac.il
Computer Science, Tel Aviv University, Israel

Rebegoldi, Simone ... WE-05
simone.rebegoldi@unimore.it
Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Italy

Ricci, Jacopo Maria ... WC-07
jacopomaria.ricci@uniroma3.it
Economia Aziendale, Università degli Studi Roma Tre, Rome, Italy, Italy

Richtarik, Peter ... WD-05
peter.richtarik@kaust.edu.sa
KAUST, Thuwal, Saudi Arabia

Richtarik, Peter ... TB-05
richtarik@gmail.com
Computer Science, KAUST, Thuwal, Saudi Arabia

Ridolfi, Andrea ... FC-03
abridolfi@gmail.com
Universidad Nacional de Cuyo, San Rafael, Mendoza, Argentina

Rigó, Petra Renáta ... FB-06
petrarenata92@gmail.com
Corvinus University of Budapest, Budapest, Hungary

Roberts, Lindon ... FC-05
lindon.roberts@anu.edu.au
Australian National University, Australia

Rocha, Luiz Celio Souza ... WE-07
luizrochamg@hotmail.com
Management, Federal Institute of Northern Minas Gerais, Almenara, Minas Gerais, Brazil

Rosasco, Lorenzo ... WF-07
lrosasco@mit.edu
MaLGa, DIBRIS, Università di Genova, Genova, Italy

Rotaru, Teodor ... WF-02
teodor.rotaru@kuleuven.be
Electrical Engineering, KU Leuven, Leuven, Belgium

Rotella Junior, Paulo ... WE-07
paulo.rotela@gmail.com
Faculty of Social Sciences, Charles University, Czech Republic, and Federal University of Paraiba - UFPB, Brazil

Rubbens, Anne ... WE-02
anne.rubbens@uclouvain.be
ICTEAM, UCLouvain, Bruxelles, Bruxelles, Belgium

Ruckmann, Jan-J ... FB-03
Jan-Joachim.Ruckmann@ii.uib.no
Department of Informatics, University of Bergen, Bergen, Norway

Ryu, Ernest ... WC-05
ernestryu@snu.ac.kr
National University Ryu, Seul, Korea, Republic of

Sabach, Shoham ... FC-05
ssabach@technion.ac.il
Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa, Israel

Sagratella, Simone ... FB-02
sagratella@dis.uniroma1.it
Ingegneria informatica automatica e gestionale A. Ruberti, La Sapienza Università di Roma, Roma, Italy

Sahinidis, Nikolaos ... TC-07
nikos@gatech.edu

Georgia Institute of Technology, Atlanta, GA, United States

Salehi, Mohammad Sadegh FC-05
mss226@bath.ac.uk
Mathematical Sciences, University of Bath, Bath, United Kingdom

Salomon, Ludovic TC-04
ludovic.salomon@polymtl.ca
MAGI, Polytechnique Montréal, Montreal, Quebec, Canada

Sasso, Valerio Giuseppe FB-02
sasso@diag.uniroma1.it
Dipartimento di ingegneria Informatica, Automatica e Gestionale, Sapienza Università di Roma, Roma, Italy

Savku, Emel FD-04
esavku@gmail.com
Department of Mathematics, University of Oslo, Oslo, –Select state–, Norway

Scarponi, Giulio FD-06
giulio.scarponi@uniroma1.it
Sapienza, Italy

Schmidt, Regina FC-07
regina.schmidt@math.uni-augsburg.de
University of Augsburg, Augsburg, Germany

Schneider, Christopher TD-04
christopher.schneider@eah-jena.de
Fachbereich Grundlagenwissenschaften, Ernst-Abbe-Hochschule Jena, Jena, Thuringia, Germany

Schoen, Fabio TB-04
fabio.schoen@unifi.it
Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Firenze, Firenze, Italy

Sciandrone, Marco FD-06
sciandrone@diag.uniroma1.it
DIAG, Sapienza Università di Roma, Italy

Sebastiani, Andrea WE-05
asebastiani@unimore.it
Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, Modena, Italy

Sedlmayer, Michael FB-04
michael.sedlmayer@univie.ac.at
Faculty of Mathematics, University of Vienna, Vienna, Vienna, Austria

Seidel, Tobias WE-07
tobias.seidel@itwm.fhg.de
Optimization, Fraunhofer Institute ITWM, Germany

Seraghiti, Giovanni WC-05
giovanni.seraghiti@umons.ac.be
Umons, Mons, Belgium

Shafiei, Allahkaram WC-06
shafie.allahkaram@gmail.com
Computer Science, Czech Technical University, Prague, Czech Republic

Shikhman, Vladimir FB-02, 03, FC-07
vladimir.shikhman@mathematik.tu-chemnitz.de
Chemnitz University of Technology, Germany

Shtern, Shimrit TB-06
shimrits@technion.ac.il
Industrial Engineering and Management, Technion-Israel Institute of Technology, Haifa, Israel

Shulgin, Egor TB-05
egor.shulgin@kaust.edu.sa
KAUST AI Initiative, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

Singh, Shambhavi WF-07
sambha@mail.ubc.ca
Combinatorics and Optimization, University of Waterloo, Kelowna, British Columbia, Canada

Sinjorgo, Lennart WC-02
l.m.sinjorgo@tilburguniversity.edu
Tilburg University, Tilburg, Netherlands

Sjölund, Jens WF-04
jens.sjolund@it.uu.se
Department of Information Technology, Uppsala University, Sweden

Sonja, Steffensen WE-06
steffensen@igpm.rwth-aachen.de
RWTH Aachen University, Aachen, Germany

Sotirov, Renata WC-02
r.sotirov@uvt.nl
Tilburg University, Tilburg, Outside of Canada & USA, Netherlands

Sra, Suvrit FD-06
suvrit@mit.edu
Massachusetts Institute of Technology, Cambridge, United States

Staudigl, Mathias TB-06, WE-06
mathias.staudigl@uni-mannheim.de
Department of Mathematics, Universität Mannheim, Mannheim, Germany

Stein, Oliver TB-03
stein@kit.edu
Institute of Operations Research, Karlsruhe Institute of Technology, Karlsruhe, Germany

Stich, Sebastian FA-01
stich@cispa.de
CISPA Helmholtz Center for Information Security, Germany

Still, Georg TC-03
g.still@math.utwente.nl
Mathematics, University of Twente, Enschede, Netherlands

Stillfjord, Tony TD-06
tony.stillfjord@math.lth.se
Centre for Mathematical Sciences, Lund University, Lund, Sweden

Strugariu, Radu WF-04
rstrugariu@tuiasi.ro
Mathematics, Gheorghe Asachi Technical University & Octav Mayer Institute of Mathematics, Iasi, Iasi, Romania

Suh, Jaewook WC-05
jacksuhkr@snu.ac.kr
Seoul National University, Korea, Republic of

Suonperä, Ensio FC-05
ensio.suonpera@helsinki.fi
Mathematics and statistics, University of Helsinki, Helsinki, Other (Non U.S.), Finland

Tam, Matthew WF-03
matthew.tam@unimelb.edu.au
School of Mathematics and Statistics, University of Melbourne, Australia

Tanaka, Mirai FC-04
mirai@ism.ac.jp
Department of Statistical Inference and Mathematics, The Institute of Statistical Mathematics, Tachikawa, Tokyo, Japan

Tanji, Sofiane WC-07
sofiane.tanji@uclouvain.be
UCLouvain, Belgium

Taylor, Adrien TD-02, WE-02
adrien.taylor@inria.fr
Inria/ENS, France

Themelis, Andreas FB-02, FB-04
andreas.themelis@ees.kyushu-u.ac.jp
Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

Thormann, Marah-Lisanne FD-05
m.-l.thormann@soton.ac.uk
Mathematical Sciences, University of Southampton, Southampton, Hampshire, United Kingdom

Tofani, Alberto WE-07
alberto.tofani@enea.it
ENEA, Italy

Tondji, Lionel WF-07
tngoupeyou@aimsammi.org
Braunschweig, Lower Saxony, Germany

Toyoda, Mitsuru FC-04
toyoda@tmu.ac.jp
Department of Mechanical Systems Engineering, Tokyo Metropolitan University, Hino-shi, Tokyo, Japan

Traoré, Cheik TD-07
temci76@gmail.com
University of Genoa, Genova, Liguria, Italy

Tribes, Christophe TC-04
christophe.tribes@polymtl.ca
Mathematics and Industrial Engineering, Polytechnique Montréal, Canada

Trunschke, Philipp WF-06
philipp.trunschke@univ-nantes.fr
Centrale Nantes & Nantes Université, France

Ugurlu, Kerem FD-04
keremugurlu@gmail.com
Mathematics, Nazarbayev University, Kazakhstan

Ulus, Firdevs TC-04
firdevs@bilkent.edu.tr
Industrial Engineering, Bilkent University, Ankara, Turkey

Upadhyaya, Manu WE-02
manu.upadhyaya@control.lth.se
Lund University, Sweden

Valkonen, Tuomo FC-05
tuomov@iki.fi
Escuela Politécnica Nacional, Ecuador

Van Dessel, Guillaume TB-02
guiomvds3@gmail.com
Mathematical Engineering, UCLouvain, Overijse, Belgium

Vandaele, Arnaud TB-04, WC-05
arnaud.vandaele@gmail.com
Mathematics and Operations Research, University of Mons, Mons, Belgium

Varga, Anita FB-06
anita.varga@uni-corvinus.hu
Corvinus Centre for Operations Research, Corvinus University of Budapest, Budapest, Hungary

Varre, Aditya WD-03
aditya.varre@epfl.ch
EDIC, EPFL, Lausanne, Vaud, Switzerland

Vera de Serio, Virginia N. FC-03
vvera@uncu.edu.ar
Facultad de Cs. Económicas - Instituto de Cs. Básicas, Universidad Nacional de Cuyo, MENDOZA, MENDOZA, Argentina

Viano, Luca FC-05
luca.viano@epfl.ch
École Polytechnique Fédérale de Lausanne, Switzerland

Vidal-Nunez, Jose TD-03
j.vidal@uah.es
University of Alcalá, Spain

Villa, Silvia FB-02, WF-07
silvia.villa@unige.it
Department of Mathematics, MaLGa, università di Genova, Genova, Italy

Volk, Maximilian TB-03
maximilian.volk@student.kit.edu
Institute of Operations Research, Karlsruhe Institute of Technology, Germany

Vu, Cong Bang WC-06
bangcvvn@gmail.com
3NLab, Huawei Belgium Research Center (BeRC), Leuven,

Belgium, Leuven, Belgium

Wahl, Niklas WC-04
n.wahl@dkfz-heidelberg.de
German Cancer Research Center (DKFZ), Germany

Wang, Guan-Ren WE-04
guanren.wang.tw@gmail.com
National Taiwan University, Taiwan

Wang, Hao WE-03
haw309@gmail.com
Information Science and Technology, ShanghaiTech University, Shanghai, China

Weber, Gerhard-Wilhelm FC-03, FD-04
gerhard-wilhelm.weber@put.poznan.pl
Faculty of Engineering Management, Poznan University of Technology, Poznan, Poland

Werner, Ralf FB-03
werner_ralf@gmx.net
Institut für Mathematik, Universität Augsburg, Augsburg, Germany

Williamson, Måns TD-06
mans.williamson@math.lth.se
Centre for Mathematical Sciences, Lund University, Sweden

Wodecki, Ales TB-06
wodecale@fel.cvut.cz
Computer science, Czech Technical University in Prague, Prague, Czech Republic

Wright, Stephen WC-06
swright@cs.wisc.edu
Computer Science, University of Wisconsin - Madison, Madison, WI, United States

Wu, Yuqia WE-03
yuqia.wu@connect.polyu.hk
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong

Yamamoto, Seima FB-05
seima19991111@gmail.com
Keio University, Japan

Yang, Xiaoqi WE-03
mayangxq@polyu.edu.hk
Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Yannacopoulos, Athanasios FD-04
ayannaco@aueb.gr
Athens University of Economics and Business, Athens, Greece

Yildirim, E. Alper WD-02
eayildirim@gmail.com
School of Mathematics, The University of Edinburgh, Edinburgh, United Kingdom

Yoon, TaeHo WC-05
tetrzim@gmail.com
Mathematical Sciences, Seoul National University, Seoul, Korea, Republic of

Yurtsever, Alp FD-06, TB-06
alp.yurtsever@umu.se
Umeå University, Sweden

Zalinescu, Constantin WF-04
zalinesc@uaic.ro
Institute of Mathematics, Iasi Branch of Romanian Academy, Iasi, Romania

Zamani, Moslem TD-02
moslem.zamani@uclouvain.be
Mathematical Engineering, UCLouvain, Louvain-la-Neuve, Belgium

Zdeborová, Lenka WD-03
lenka.zdeborova@epfl.ch
EPFL, Switzerland

Zemkoho, Alain FD-05
a.b.zemkoho@soton.ac.uk
Mathematics, University of Southampton, Southampton, Hampshire, United Kingdom

Zhang, Yi WD-04, TC-07
yzhang@minlp.com
The Optimization Firm, Hangzhou, Zhejiang, China

Zhu, Jia-Jie FD-07
zplusj@gmail.com
Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Berlin Germany, Berlin, Germany

Zhu, Jiamin WF-03
jiamin.zhu@ifpen.fr
IFP Energies Nouvelles, France

Wednesday, 8:30 - 8:45

WA-01: Opening session (M:A) 1

Wednesday, 8:45 - 9:35

WB-01: Plenary I - Gabriel Peyré (M:A) 1

Wednesday, 10:05 - 11:20

WC-02: Conic and Semidefinite Optimization (M:O) 2
WC-03: Optimization in neural architectures I (M:J) 2
WC-04: Optimization in regression, classification and learning I (M:M) 3
WC-05: Optimization for learning I (M:N) 3
WC-06: Advances in monotone inclusions and related methods (M:H) 4
WC-07: Optimization applications I (M:I) 4

Wednesday, 11:25 - 12:40

WD-02: Conic and polynomial optimization (M:O) 6
WD-03: Optimization in neural architectures II (M:J) 6
WD-04: BARON Tutorial (M:M) 7
WD-05: Optimization for learning II (M:N) 7
WD-06: Linear Programming (M:H) 8
WD-07: Optimization applications II (M:I) 8

Wednesday, 14:10 - 15:50

WE-02: Recent advances in computer-aided analyses of optimization algorithms I (M:O) 10
WE-03: Structured sparse optimization (M:J) 10
WE-04: Large-scale optimization I (M:M) 11
WE-05: Randomized optimization algorithms part 1/2 (M:N) 12
WE-06: Higher-order Methods in Mathematical Programming I (M:H) 13
WE-07: Optimization applications III (M:I) 13

Wednesday, 16:20 - 18:00

WF-02: Recent advances in computer-aided analyses of optimization algorithms II (M:O) 15
WF-03: Splitting algorithms (M:J) 15
WF-04: Multiobjective Optimization I (M:M) 16
WF-05: Randomized optimization algorithms part 2/2 (M:N) 17
WF-06: Stochastic Gradient Methods: Bridging Theory and Practice (M:H) 17
WF-07: Regularization methods for Machine Learning and Inverse Problems (M:I) 18

Thursday, 8:45 - 9:35

TA-01: Plenary II - Amir Beck (M:A) ... 20

Thursday, 10:05 - 11:20

TB-02: Solver-based optimization algorithms (M:O) ... 21
TB-03: In memory of Georg Still - part 1 (M:J) ... 21
TB-04: Optimization in regression, classification and learning II (M:M) ... 22
TB-05: Optimization for learning III (M:N) ... 22
TB-06: Advances in Semi-Definite Programming (M:H) ... 23
TB-07: Global optimization I (M:I) ... 23

Thursday, 11:25 - 12:40

TC-02: Developments in interior point methods (M:O) ... 25
TC-03: In memory of Georg Still - part 2 (M:J) ... 25
TC-04: Multiobjective Optimization II (M:M) ... 26
TC-05: Recent advances in bilevel optimization I (M:N) ... 26
TC-06: Stochastic methods and applications (M:H) ... 27
TC-07: Global optimization II / MINLP (M:I) ... 27

Thursday, 14:10 - 15:50

TD-02: Recent advances in computer-aided analyses of optimization algorithms III (M:O) ... 29
TD-03: Variational techniques and subdifferentials (M:J) ... 29
TD-04: Multiobjective Optimization III (M:M) ... 30
TD-05: Nonsmooth optimization algorithms (M:N) ... 31
TD-06: Stochastic methods (M:H) ... 31
TD-07: Accelerated methods in modern optimization (M:I) ... 32

Thursday, 16:20 - 17:20

TE-01: Plenary III - Gabriele Eichfelder EUROPT Fellow (M:A) ... 33

Friday, 8:45 - 9:35

FA-01: Plenary IV - Sebastian Stich (M:A) ... 34

Friday, 10:05 - 11:20

FB-02: First-order methods for multi-level and multi-objective optimization (M:O) ... 35
FB-03: In memory of Georg Still - part 3 (M:J) ... 35
FB-04: Large-scale optimization II (M:M) ... 36
FB-05: Recent advances in bilevel optimization II (M:N) ... 36

FB-06: Higher-order Methods in Mathematical Programming II (M:H) 37

Friday, 11:25 - 12:40

FC-02: Algorithms for Variational inequalities and equilibria (M:O) 38
FC-03: In memory of Georg Still - part 4 (M:J) 38
FC-04: Large-scale optimization III (M:M) 39
FC-05: Recent advances in bilevel optimization III (M:N) 39
FC-07: Optimization mosaics (M:I) 40

Friday, 14:10 - 15:50

FD-02: Deterministic and stochastic optimization beyond Euclidean geometry (M:O) 41
FD-03: Variational techniques in optimization (M:J) 41
FD-04: Optimal control and stochastic optimal control - theory, methods and applications 1 (M:M) 42
FD-05: Structured nonsmooth optimization (M:N) 42
FD-06: Difference and decomposition methods (M:H) 43
FD-07: Optimal Transport for Machine Learning and Inverse Problems (M:I) 44