

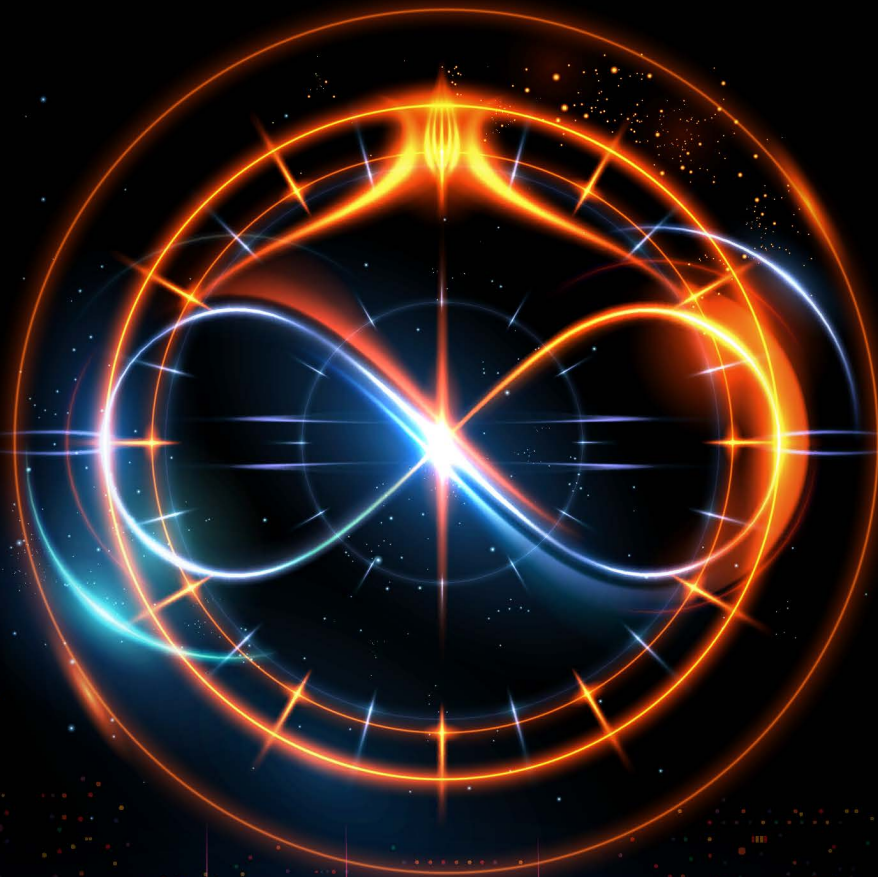
ICOML 2026

International Conference on
Optimization and Machine Learning

July 27-29, 2026 | **Taipei X Taiwan** |
National Taiwan Normal University

S102, Lecture Hall, Gongguan Campus, NTNU

Program Book



ORGANIZERS :



College of Science,
National Taiwan Normal University



$\Sigma\Delta\Gamma\text{H}$ 80
Department of Mathematics,
National Taiwan Normal University



Center for Optimal Intelligent
Data Analytics and Prediction,
National Taiwan Normal University

SPONSORS :



NSTC 國家科學及技術委員會
National Science and Technology Council



Office of Research and Development,
National Taiwan Normal University

Program

Day 1 (July 27, 2026, Monday)

Time	Activity
09:00–09:30	Registration and Morning Refreshments
09:30–10:00	Welcome and Opening Ceremony
Session 1 Chair: Shuzhong Zhang	
10:00–10:30	<i>Robust and Flexible Reward Modeling in Reinforcement Learning</i> Masashi Sugiyama
10:30–11:00	<i>Orthogonalized Moment Aberration for Hyperparameter Optimization under Heterogeneity</i> Ming-Chung Chang
11:00–11:30	<i>Data-Driven Analysis of First-Order Methods via Distributionally Robust Optimization</i> Bartolomeo Stellato
11:30–12:00	<i>Extended-Variable Relaxations for the Constrained Generalized Maximum-Entropy Sampling Problem</i> Jon Lee
12:00–13:30	Lunch Break
Session 2 Chair: Bartolomeo Stellato	
13:30–14:00	<i>New Results on the Majorization Schemes for Nonconvex Constrained Optimization</i> Shuzhong Zhang
14:00–14:30	<i>Learning-Enhanced Approximations of Mixed Integer Nonlinear Optimization Problems</i> Claudia D'Ambrosio
14:30–15:00	<i>Evolution of Network Intelligence: From Classical Optimization to AI Agents</i> Chih-Wei Huang
15:00–15:30	Coffee and Tea Break
Session 3 Chair: Claudia D'Ambrosio	
15:30–16:00	<i>Stochastic Block-Iterative Splitting of Systems of Monotone Inclusions</i> Patrick L. Combettes
16:00–16:30	<i>A Conditional-Gradient-Based Single-Loop Augmented Lagrangian Method for Inequality Constrained Problems</i> Ting Kei Pong
16:30–17:00	<i>Point Sketches versus Coresets in k-Means</i> Leo Liberti

Day 2 (July 28, 2026, Tuesday)

Time	Activity
09:30–10:00	Morning Refreshments
Session 4 Chair: Kim-Chuan Toh	
10:00–10:30	<i>Non-Stationary Douglas-Rachford Revisited</i> Matthew K. Tam
10:30–11:00	<i>The Duality between Sampling and Optimization: New Perspectives from Diffusions and Optimal Control</i> Lorenz Richter
11:00–11:30	<i>Robustness against Noise</i> LieJune Shiau
11:30–12:00	<i>An Anchored Random Subspace Newton Algorithm</i> Pierre-Louis Poirion
12:00–13:30	Lunch Break
13:30–15:00	Poster Session
15:00–15:30	Coffee and Tea Break
Session 5 Chair: Lorenz Richter	
15:30–16:00	<i>RiNNAL+: A Riemannian ALM Solver for SDP-RLT Relaxations of Mixed-Binary Quadratic Programs</i> Kim-Chuan Toh
16:00–16:30	<i>On Circumcentered Direct Methods for Monotone Variational Inequality Problems</i> Luiz-Rafael Santos
16:30–17:00	<i>A Comparison of Convex Relaxations of Optimization Problems with Hard Cardinality Constraints</i> Emre Alper Yildirim
17:00–17:30	Group Photo
17:30–19:00	Banquet (by invitation only)

Day 3 (July 29, 2026, Wednesday)

Time	Activity
09:30–10:00	Morning Refreshments
Session 6 Chair: David M. Rosen	
10:00–10:30	<i>State-Space Representation of Neural Networks via Implicit Models and Applications</i> Laurent El Ghaoui
10:30–11:00	<i>Sum of Squares Submodularity</i> Anna Deza
11:00–11:30	<i>Tackling Dimensional Collapse toward Comprehensive Universal Domain Adaptation</i> Hsuan-Tien Lin
11:30–12:00	<i>Linear Interior-Point Methods</i> Bento Natura
12:00–13:30	Lunch Break
Session 7 Chair: Patrick L. Combettes	
13:30–14:00	<i>Concave Certificates: Geometric Framework for Distributionally Robust Risk and Complexity Analysis</i> Chu Thi Mai Hong
14:00–14:30	<i>Approximation Theory of Laplacian-Based Neural Operators for Reaction-Diffusion Systems</i> Jenn-Nan Wang
14:30–15:00	<i>Online Learning of Quantum States with Logarithmic Loss</i> Yen-Huan Li
15:00–15:30	Coffee and Tea Break
Session 8 Chair: Laurent El Ghaoui	
15:30–16:00	<i>Efficient Global Polynomial Optimization for Machine Perception</i> David M. Rosen
16:00–16:30	<i>Adaptive Forward-Backward Splitting for Multi-Layer Convolutional Dictionary Learning</i> Guan-Ju Peng
16:30–17:00	<i>A Neurodynamical Neural Network Approach for Solving Support Vector Machine Problems with Generalized Pinball Loss via Hyperbolic Tangent Smoothing</i> Rabian Wangkeeree

Abstracts

Day 1

Monday, July 27, 2026

Robust and Flexible Reward Modeling in Reinforcement Learning

Masashi Sugiyama

10:00–10:30

RIKEN/The University of Tokyo

Abstract. Reinforcement learning (RL) has achieved remarkable success in robotics, games, and language model post-training, where reward signals are essential for effective agent training. In this talk, I will present our recent research on advanced reward modeling. This includes robustifying RL through transfer and weakly supervised learning, coping with interval-based rewards, and exploring diverse reward aggregation frameworks that move beyond the traditional discounted sum.

Orthogonalized Moment Aberration for Hyperparameter Optimization under Heterogeneity

Ming-Chung Chang

10:30–11:00

Institute of Statistical Science, Academia Sinica

Abstract. Hyperparameter optimization often takes place under heterogeneous conditions, such as different datasets, hardware settings, and random seeds. Classical experimental designs are usually developed for homogeneous settings and may therefore lose efficiency in practice. In this talk, I will introduce a framework that applies orthogonalized moment aberration (OMA) to hyperparameter tuning. OMA was originally proposed for mixed-level multi-stratum factorial designs and provides a model-free yet statistically justified way to compare designs through kernel-based similarities while accounting for heterogeneous experimental units. By treating hyperparameters as treatment factors and tuning environments as strata, the proposed approach constructs robust and diverse hyperparameter configurations. The resulting designs are closely connected to Bayesian D-optimality and Gaussian-process-based entropy criteria, providing an efficient and principled foundation for hyperparameter experimentation under heterogeneity.

Data-Driven Analysis of First-Order Methods via Distributionally Robust Optimization

Bartolomeo Stellato

11:00–11:30

Department of Operations Research and Financial Engineering, Princeton University

Abstract. We consider the problem of analyzing the probabilistic performance of first-order methods when solving convex optimization problems drawn from an unknown distribution only accessible through samples. By combining performance estimation (PEP) and Wasserstein distributionally robust optimization (DRO), we formulate the analysis as a tractable semidefinite program. Our approach unifies worst-case and average-case analyses by incorporating data-driven information from the observed convergence of first-order methods on a limited number of problem instances. This yields probabilistic, data-driven performance guarantees in terms of the expectation or conditional value-at-risk of the selected performance metric. Experiments on smooth convex minimization, logistics regression, and Lasso show that our method significantly reduces the conservatism of classical worst-case bounds and narrows the gap between theoretical and empirical performance.

Extended-Variable Relaxations for the Constrained Generalized Maximum-Entropy Sampling Problem

Jon Lee

11:30–12:00

Industrial and Operations Engineering Department, University of Michigan

Abstract. The generalized maximum-entropy sampling problem (GMESP) is to select an order- s principal submatrix from an order- n covariance matrix, to maximize the product of its t greatest eigenvalues, $0 < t \leq s < n$. Introduced more than 25 years ago, GMESP is a natural generalization of two fundamental problems in statistical design theory: (i) maximum-entropy sampling problem (MESP); (ii) binary D-optimality (D-Opt). In the general case, it can be motivated by a selection problem in the context of principal component analysis (PCA). We present (i) non-convex extended variable formulations, (ii) first non-convex and then convex continuous relaxations, (iii) results analyzing our new (upper) bounds and demonstrating some relations between different bounds, including bounds from the literature and our new bounds, (iv) theory related to how to carry out B&B, and (v) favorable numerical results.

New Results on the Majorization Schemes for Nonconvex Constrained Optimization

Shuzhong Zhang

13:30–14:00

Department of Industrial & Systems Engineering, University of Minnesota

Abstract. Constrained nonconvex optimization models are frequently encountered in machine learning research. In this talk, I shall present some new results (based on the joint work with Nuo Zhou Wang and Junyu Zhang) in the study of convex surrogation that majorizes the objective and the constraints. In the case where the objective and constraint functions are gradient Holderian continuous, the surrogate functions can be readily constructed, and the solution method can be efficiently implemented. The surrogate envelopes are extended to the setting where the second-order information is available, and the convex subproblems are further represented by the Dikin ellipsoids using the self-concordance of the convex surrogate constraints. Iteration complexities have been developed for both convex and nonconvex optimization models. Preliminary numerical results show the promises of the proposed approaches.

Learning-Enhanced Approximations of Mixed Integer Nonlinear Optimization Problems

Claudia D'Ambrosio

14:00–14:30

École Polytechnique

Abstract. Mixed Integer Non Linear Optimization (MINLP) problems are a challenging class of mathematical optimization problems. In this talk, we present and compare several approaches to approximate MINLP with the aim of finding good quality heuristic solutions. In particular, we focus on classic and learning-enhanced piecewise linear and non-linear approximations. The methods are computationally tested on a benchmark of diverse MINLP instances, some from real-world applications arising in engineering.

Evolution of Network Intelligence: From Classical Optimization to AI Agents

Chih-Wei Huang

14:30–15:00

Department of Communication Engineering, National Central University

Abstract. Traditional model-based optimization has long been the foundation for resource allocation in wireless networks. However, as future network systems become increasingly complex, we need new approaches to handle highly dynamic and non-stationary environments. This talk traces the evolution of network intelligence through three main stages. The first stage focuses on system perception and prediction using hybrid convolutional and recurrent neural networks. The second stage discusses dynamic decision-making and control. We show how deep reinforcement learning and multi-agent algorithms can be deployed to efficiently solve resource allocation problems that are often too complex for traditional solvers. The third stage explores the concept of generalization, introducing how foundation model based agentic architectures can be used to manage and execute complex networking tasks. Together, these developments show how network intelligence is shifting from classical optimization to cognitive agents, and we will conclude by discussing the open challenges for applied optimization in the AI era.

Stochastic Block-Iterative Splitting of Systems of Monotone Inclusions

Patrick L. Combettes

15:30–16:00

North Carolina State University, USA.

Abstract. We propose an abstract stochastic scheme for solving a broad range of monotone operator inclusion problems in Hilbert spaces. The application we highlight concerns randomized block-iterative projective splitting methods for solving systems of coupled inclusions involving a mix of set-valued, cocoercive, and Lipschitzian monotone operators combined via various monotonicity-preserving operations. Almost sure and L^2 convergence results are obtained for the iterates of the proposed scheme.

This is a joint work with Javier I. Madariaga.

A Conditional-Gradient-Based Single-Loop Augmented Lagrangian Method for Inequality Constrained Problems

Ting Kei Pong

16:00–16:30

Department of Applied Mathematics, The Hong Kong Polytechnic University

Abstract. We consider the problem of minimizing the sum of a Lipschitz differentiable convex function and a proper closed convex function that admits efficient linear optimization oracles, subject to multiple smooth convex inequality constraints. We adapt the classical augmented Lagrangian (AL) method for these problems: in each iteration, our algorithm consists of one step of conditional gradient (CG) method applied to the AL function, followed by an update of the dual variable as in classical AL methods with a diminishing dual stepsize. We study the convergence rate of our algorithm under two standard stepsize rules for the CG method, namely, an open-loop stepsize and the short stepsize, and obtain a rate that matches the best-known complexity for this class of problems. We also establish accelerated rates when the aforementioned proper closed convex function is the indicator function of a uniformly convex set. This is a joint work with Xiaozhou Wang and Zev Woodstock.

Point Sketches versus Coresets in k-Means

Leo Liberti

16:30–17:00

LIX, CNRS and Ecole Polytechnique, Institut Polytechnique de Paris

Abstract. Given a set of n points in d dimensions, the k -means problem consists in finding a clustering of given cardinality k such that the sum of squared Euclidean distances from each given point to the centroid of the cluster it belongs to is minimum. The problem inherits its name from a homonymous heuristic algorithm that is known to be both fast and effective. Random projections have been applied to the k -means problem to decrease the dimensions of each point, while coresets have been used to decrease the number of points. We show that the latter feat can also be achieved through a matrix sketch—a construction similar to random projections—applied to the rows of the data matrix, i.e. the given points in a k -means instance. Our computational tests indicate that, on single values of k , coresets give better quality results faster: but point sketches become much faster than coresets when one needs to rapidly compute approximate clusterings for many values of k .

Day 2

Tuesday, July 28, 2026

Non-Stationary Douglas-Rachford Revisited**Matthew K. Tam**

10:00–10:30

School of Mathematics and Statistics, The University of Melbourne

Abstract. The Douglas-Rachford algorithm is a method for minimisation the sum of two nonsmooth functions using their proximity operators. The standard analysis of the method interprets it as the fixed-point iteration of a nonexpansive operator, from which convergence follows. In contrast to many methods in optimisation, the fixed-points of this underlying operator are difficult to work with; they are not solutions to problem at hand and they have a parametric dependence on the stepsize. Both of these properties make analysing "non-stationary" versions of the Douglas-Rachford algorithm difficult. In this talk, we develop machinery for analysing Douglas-Rachford methods with non-constant stepsize. The approach relies on a simple but so-far unused observation about proximity operators of convex functions.

The Duality between Sampling and Optimization: New Perspectives from Diffusions and Optimal Control**Lorenz Richter**

10:30–11:00

Zuse Institute Berlin

Abstract. The duality between optimization and sampling provides a powerful mathematical framework for exploring complex probability measures and highly non-convex landscapes. In this talk, we examine this relationship in both directions and show how ideas from one domain can be used to overcome fundamental limitations in the other. We begin with the direction from optimization to sampling. By formulating stochastic optimal control as an optimization problem over path space measures, we derive a new trust-region-based approach for sampling from high-dimensional, multimodal target distributions. This perspective enables principled control of exploration and improves robustness in challenging sampling settings. We then reverse the viewpoint and use diffusion-based sampling techniques to address global optimization problems. We introduce a non-convex optimization method that combines controlled diffusion dynamics with temperature tempering. Through systematic benchmarking, we analyze the failure modes of classical first-order optimizers and traditional MCMC approaches such as Simulated Annealing. Our results show that leveraging diffusion dynamics leads to improved scaling behavior and more reliable escape from local minima in highly multimodal objectives.

Robustness against Noise

LieJune Shiau

11:00–11:30

Department of Mathematics, University of Houston, Clear Lake

Abstract. Noise, fluctuations, and variability in biological systems are often viewed as undesirable or even detrimental, as excessive noise can contribute to disease and dysfunction. Nevertheless, noise is ubiquitous and arises inevitably from both intrinsic and extrinsic sources. How, then, does the body cope with such pervasive variability? In this work, we demonstrate that, under certain conditions, biological noise can in fact play a beneficial role in supporting a variety of fundamental physiological functions.

An Anchored Random Subspace Newton Algorithm

Pierre-Louis Poirion

11:30–12:00

RIKEN Center for Advanced Intelligence Project

Abstract. Random subspace methods help mitigate the curse of dimensionality by enabling the computation of inexpensive descent directions for high-dimensional optimization problems. However, this gain in efficiency often comes at the cost of slower convergence. In particular, in the Newton setting, the use of random subspaces generally prevents superlinear convergence.

In this talk, we present a new method that augments the random subspace with an anchor direction designed to better approximate the true Newton direction. We show that this approach restores fast local behavior and prove that the method achieves local superlinear convergence.

Joint work with Haruki Yamakawa and Akiko Takeda

RiNNAL+: A Riemannian ALM Solver for SDP-RLT Relaxations of Mixed-Binary Quadratic Programs

Kim-Chuan Toh

15:30–16:00

National University of Singapore, Singapore

Abstract. Doubly nonnegative (DNN) relaxation usually provides a tight lower bound for a mixed-binary quadratic program (MBQP). However, solving DNN problems is challenging because: (1) the problem size is $\Omega((n+l)^2)$ for an MBQP with n variables and l inequality constraints, and (2) the rank of optimal solutions cannot be estimated a priori due to the absence of theoretical bounds. In this work, we propose RiNNAL+, a Riemannian augmented Lagrangian method (ALM) for solving DNN problems. We prove that the DNN relaxation of an MBQP, with matrix dimension $n + l + 1$, is equivalent to the SDP-RLT relaxation (based on the reformulation-linearization technique) with a smaller matrix dimension $n + 1$. In addition, we develop a hybrid method that alternates between two phases to solve the ALM subproblems. In phase one, we apply low-rank matrix factorization and random perturbation to transform the feasible region into a lower-dimensional manifold so that we can use the Riemannian gradient descent method. In phase two, we apply a single projected gradient step to update the rank of the underlying variable and escape from spurious local minima arising in the first phase if necessary. To reduce the computation cost of the projected gradient step, we develop pre-processing and warm-start techniques for acceleration. Unlike traditional rank-adaptive methods that require extensive parameter tuning, our hybrid method requires minimal tuning. Extensive experiments confirm the efficiency and robustness of RiNNAL+ in solving various classes of large-scale DNN problems.

This talk is based mainly on joint work with Di Hou and Tianyun Tang.

On Circumcentered Direct Methods for Monotone Variational Inequality Problems

Luiz-Rafael Santos

16:00–16:30

Departamento de Matemática, Universidade Federal de Santa Catarina

Abstract. The variational inequality problem (VIP) plays a central role in the theory and applications in continuous optimization. In particular, minimization problems and KKT systems can be regarded as VIPs. In this work, we present the first methods using circumcenters for solving VIPs. The circumcentered-reflection method (CRM) is a projection-based tool developed with the aim of finding a point in the intersection of finitely many closed convex sets. CRM has gone through enhancements and adaptations over the last few years and was proven to be faster in many settings than competitors such as alternating projections and the Douglas-Rachford methods. One of CRM features is that it is able to deal with approximate projections, as we do in this paper. Here, we present circumcenter schemes that solve VIPs with two types of operators: paramonotone and monotone ones. For the former we employ exact projections, whereas for the latter we base our approach on approximate ones. We establish convergence results for both methods and we perform numerical experiments, which show a remarkable performance when compared to other well know methods, like the extragradient algorithm.

A Comparison of Convex Relaxations of Optimization Problems with Hard Cardinality Constraints

Emre Alper Yildirim

16:30–17:00

The University of Edinburgh

Abstract. Sparsity plays a vital role in a wide range of applications, such as machine learning, data science, image processing, portfolio management, and sparse regression, due to the interpretability, robustness, and ease of implementation of sparse solutions. As such, optimization problems with cardinality constraints arise in a plethora of applications such as interpretable machine learning, portfolio optimization, and principal component analysis.

In this talk, we consider alternative formulations of the NP-hard class of optimization problems with hard cardinality constraints. We introduce various polyhedral and convex conic relaxations of such formulations. We compare the resulting convex relaxations in terms of their strength and computational cost.

This is joint work with Merve Bodur and Buket Ozen.

Day 3

Wednesday, July 29, 2026

State-Space Representation of Neural Networks via Implicit Models and Applications**Laurent El Ghaoui**

10:00–10:30

VinUniversity

Abstract. Implicit models are generalizations of neural networks, which define hidden states and outputs through equilibrium (fixed-point) equations rather than fixed layer sequences, enabling adaptive depth and feedback loops in the network's computational graph. Such models can always be represented in a "state-space" form that resembles the classical state-space models in control theory, allowing for a rigorous matrix formalism and opening many connections with convex optimization. In this talk, I will review results in this area, including: a) representation of existing networks in the state-space format; hidden-state optimization (HSO), where neural network parameters are optimized while keeping hidden states fixed; neural network inversion; controlling dynamical systems with neural networks; robustness analysis; and so on.

Sum of Squares Submodularity**Anna Deza**

10:30–11:00

MERL

Abstract. We introduce the notion of t -sum of squares (sos) submodularity, which is a hierarchy, indexed by t , of sufficient algebraic conditions for certifying submodularity of set functions. We show that, for fixed t , each level of the hierarchy can be verified via a semidefinite program of size polynomial in n , the size of the ground set of the set function. This is particularly relevant given existing hardness results around testing whether a set function is submodular (Crama, 1989). We derive several equivalent algebraic characterizations of t -sos submodularity and identify submodularity-preserving operations that also preserve t -sos submodularity. We further present a complete classification of the cases for which submodularity and t -sos submodularity coincide, as well as examples of t -sos-submodular functions. We demonstrate the usefulness of t -sos submodularity through three applications: (i) a new convex approach to submodular regression, involving minimal manual tuning; (ii) a systematic procedure to derive lower bounds on the submodularity ratio in approximate submodular maximization, and (iii) improved difference-of-submodular decompositions for difference-of-submodular optimization. Based on joint work with Georgina Hall.

Tackling Dimensional Collapse toward Comprehensive Universal Domain Adaptation

Hsuan-Tien Lin

11:00–11:30

National Taiwan University

Abstract. Universal Domain Adaptation (UniDA) addresses unsupervised domain adaptation where target classes may differ arbitrarily from source ones, except for a shared subset. A widely used approach, partial domain matching (PDM), aligns only shared classes but struggles in extreme cases where many source classes are absent in the target domain, underperforming the most naive baseline that trains on only source data. In this work, we identify that the failure of PDM for extreme UniDA stems from dimensional collapse (DC) in target representations. To address target DC, we propose to jointly leverage the alignment and uniformity techniques in self-supervised learning on the unlabeled target data to preserve the intrinsic structure of the learned representations. Our experimental results confirm that SSL consistently advances PDM and delivers new state-of-the-art results across a broader benchmark of UniDA scenarios with different portions of shared classes, representing a crucial step toward truly comprehensive UniDA.

Linear Interior-Point Methods

Bento Natura

11:30–12:00

Industrial Engineering and Operations Research, Columbia University

Abstract. We study algorithms for linear programs of the form $\min c \cdot x$, $Ax = b$, $x \geq 0$, that access the right-hand side b and objective c only through linear comparisons. An algorithm that operates under this restriction is called a linear algorithm. Such an algorithm can query an oracle which, given vectors u and v and scalars z and y , returns the signs of $u \cdot b + z$ and $v \cdot c + y$. The simplex method, as well as several known strongly polynomial combinatorial algorithms, can be implemented in this model. We show that interior-point methods can likewise be adapted to run as linear algorithms, which yields faster algorithms for several classes of linear programs and, through the framework of Norton, Plotkin, and Tardos, extends the family of problems that admit strongly polynomial algorithms.

Concave Certificates: Geometric Framework for Distributionally Robust Risk and Complexity Analysis

Chu Thi Mai Hong

13:30–14:00

College Engineering and Computer Science, VinUniversity

Abstract. Distributionally Robust (DR) optimization aims to certify worst-case risk within a Wasserstein uncertainty set. Current certifications typically rely either on global Lipschitz bounds, which are often conservative, or on local gradient information, which provides only a first-order approximation. In this talk, we introduce a novel geometric framework based on the least concave majorants of the growth rate function. Our proposed concave certificate establishes a tight bound of DR risk that remains applicable to non-Lipschitz and non-differentiable losses. We extend this framework to complexity analysis, introducing a deterministic bound that complements standard statistical generalization bound. For practical application in deep learning, we introduce the adversarial score as a tractable relaxation that enables efficient and layer-wise analysis of neural networks.

Approximation Theory of Laplacian-Based Neural Operators for Reaction-Diffusion Systems

Jenn-Nan Wang

14:00–14:30

National Taiwan University

Abstract. In this talk, I would like to discuss neural operators applied to the solution mapping from initial conditions to time-dependent solutions of some reaction-diffusion equations, in particular, generalized Gierer-Meinhardt reaction-diffusion equations, a prototypical model of nonlinear pattern formation. Our main results establish explicit approximation error bounds in terms of network depth, width, and spectral rank by exploiting the Laplacian spectral representation of the Green's function underlying the PDE. We show that the required parameter complexity grows at most polynomially with respect to the target accuracy, demonstrating that Laplacian eigenfunction-based neural operator architectures alleviate the curse of parametric complexity encountered in generic operator learning. Numerical experiments on the Gierer-Meinhardt system support the theoretical findings. This talk is based on the joint work with Furuya and Ryo.

Online Learning of Quantum States with Logarithmic Loss

Yen-Huan Li

14:30–15:00

National Taiwan University

Abstract. The problem of online learning of quantum states with logarithmic loss arises naturally in quantum state tomography, a task that is essential for building reliable quantum computing units. In terms of mathematical structure, it is a quantum generalization of, and hence more challenging than, online portfolio selection, a famous open problem in online learning for more than three decades. In this talk, I will introduce this problem, characterize the optimal regret rate for it, and provide an algorithm that is nearly regret-optimal. I will then discuss some applications of these results beyond quantum state tomography.

Efficient Global Polynomial Optimization for Machine Perception

David M. Rosen

15:30–16:00

Northeastern University

Abstract. Many fundamental perception and state estimation tasks in robotics and computer vision are naturally formalized as high-dimensional polynomial optimization problems (POPs) that are known to be computationally hard (NP-hard) to solve in general. Nevertheless, in this talk we present a novel class of certifiably correct estimation algorithms that are provably capable of efficiently recovering verifiably globally optimal solutions in many practical settings. In brief, our approach is based upon convex relaxation: first, we develop semidefinite relaxations whose minimizers we prove provide exact, globally optimal solutions to the original POP for sufficiently small measurement noise; next, we describe specialized, structure-exploiting semidefinite optimization algorithms that enable even large-scale instances of these relaxations to be solved efficiently in practice. We illustrate the design of this class of methods using the fundamental problem of pose-graph optimization (a mathematical abstraction of robotic navigation) as a motivating example, demonstrating that our approach enables the recovery of certifiably globally optimal solutions of POPs involving tens to hundreds of thousands of variables in a matter of seconds.

Adaptive Forward-Backward Splitting for Multi-Layer Convolutional Dictionary Learning

Guan-Ju Peng

16:00–16:30

Graduate Institute of Data Science and Information Computing, National Chung Hsing University

Abstract. Multi-layer convolutional sparse coding (ML-CSC) is a powerful technique for obtaining parsimonious representations of signals in machine learning applications. However, achieving maximum parsimony without sacrificing data fidelity requires learning a dictionary from data, resulting in the non-convex non-smooth optimization problem of multi-layer convolutional dictionary learning (ML-CDL). Existing ML-CDL solvers derived from traditional dictionary learning employ the convex ℓ_1 function as the sparsity-inducing function to promote convergence; nevertheless, analyzing the convergence of candidate alternate optimization schemes is still challenging. In this paper, we reformulate the ML-CDL problem by replacing the convolution with the Hadamard product to reduce computational complexity and, to achieve this without destroying the consistency between the variables in the time and frequency domains, additional constraints are introduced to the reformulated problem. We also modify a forward-backward splitting algorithm to solve the reformulated ML-CDL problem using simultaneous coefficients and dictionary elements updates across all layers. In contrast, these variables are updated alternately in the previous approaches. The proposed algorithm includes a parameter adaptation scheme to improve optimization performance and ensure convergence. In numerical simulations, the proposed method outperforms existing methods by achieving convergence in fewer iterations and generating a smaller final functional value. The high quality of recovered images demonstrates the applicability of multi-layer sparse representation to image reconstruction.

A Neurodynamical Neural Network Approach for Solving Support Vector Machine Problems with Generalized Pinball Loss via Hyperbolic Tangent Smoothing

Rabian Wangkeeree

16:30–17:00

Naresuan University

Abstract. In this talk, we present a neurodynamical neural network approach for solving support vector machine (SVM) problems with generalized pinball loss functions by using a hyperbolic tangent smoothing technique. Since the generalized pinball loss function is nonsmooth, we first construct a smooth approximation through hyperbolic tangent smoothing. This allows us to reformulate the original nonsmooth SVM problem as a smooth optimization problem that can be handled within a continuous-time dynamical framework.

Based on the smoothed formulation, we propose a continuous-time neurodynamical neural network model whose trajectories provide continuous real-time solutions of the SVM problem. We prove that the set of optimal solutions of the hyperbolic tangent smoothed optimization problem coincides with the set of equilibrium points of the proposed dynamical system. In addition, we establish several important theoretical properties of the model, including the existence and uniqueness of solutions, global asymptotic stability, global convergence, and exponential stability.

To demonstrate the effectiveness of the proposed approach, we perform numerical experiments on both synthetic datasets and benchmark real-world datasets. The performance of the model is evaluated using accuracy, F1-score, and Matthews correlation coefficient (MCC). We also compare the proposed method with three existing benchmark models for SVM problems with pinball loss functions. Furthermore, Friedman test analysis followed by the Nemenyi post-hoc test is conducted to examine the statistical significance of the observed differences among the competing methods. The results show that the proposed model achieves strong classification performance and compares favorably with the existing approaches. In particular, beyond competitive empirical performance, the proposed framework provides theoretical guarantees of global stability and convergence, making it an effective and reliable method for solving generalized pinball loss SVM problems.

Organizing Committees

Organizing Committee

- Jein-Shan Chen, National Taiwan Normal University
- Jann-Long Chern, National Taiwan Normal University
- Antoine Deza, McMaster University
- Tsung-Ming Huang, National Taiwan Normal University
- Sebastian Pokutta, Technische Universität Berlin and Zuse Institute Berlin
- Mei-Heng Yueh, National Taiwan Normal University

Local Committee

- Shyan-Shiou Chen, National Taiwan Normal University
- Tsung-Ming Huang, National Taiwan Normal University
- Tsui-Shan Lu, National Taiwan Normal University
- Shih-Feng Shieh, National Taiwan Normal University
- Ching-Yu Yang, National Taiwan Normal University