MOPTA 2021

Modeling and Optimization: Theory and Applications

2-4 August '21

http://coral.ie.lehigh.edu/~mopta

Lehigh University
Bethlehem, PA, USA

Program and Abstracts

■ Welcome to MOPTA 2021
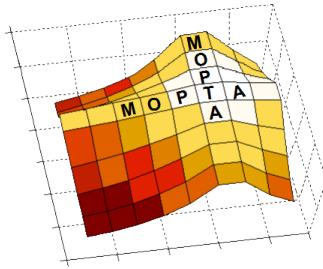
Modeling and Optimization: Theory and Applications (MOPTA) aims at bringing together a diverse group of people from both discrete and continuous optimization, working on both theoretical and applied aspects. There will be a small number of invited talks from distinguished speakers and contributed talks, spread over three days. Our target is to present a diverse set of exciting new developments from different optimization areas while at the same time providing a setting which will allow increased interaction among the participants. We aim to bring together researchers from both the theoretical and applied communities who do not usually have the chance to interact in the framework of a medium-scale event.

Tamás Terlaky
Lehigh University

■ Conference Details

Conference Location

The **conference** will take place at:
Lehigh University
Mohler Labs Department of Industrial & Systems Engineering
200 W. Packer Ave
Bethlehem, PA 18015, US

Conference in Numbers

- 3 day conference;
- 8 plenary speakers;
- 4 parallel sessions;

- 29 contributed sessions, 87 talks;
- optimization and modeling competition, 14 participants from around the world.

Conference Format

We are excited to announce that with the current University, Commonwealth, and National guidelines in place the **MOPTA** Conference will offer an in-person component this year with an option for remote attendance. This format accommodates the wishes of participants and speakers who would like to attend remotely, while still offering the in-person experience for those who would like to attend in-person. So, **MOPTA 2021** will be offered in an innovative dual **VIRT/IPVS** mode. Participants will select between two attendance modes (all guests will experience the talks in the same format):

1. **Virtual (VIRT)**, as meetings have been held during the pandemic (via Zoom). **VIRT** participants and speakers will actually not be able to tell that there is another mode.

2. **In-Person Virtual-Sharing (IPVS)**, as on-campus meetings were held before the pandemic with the single exception that all speakers will present remotely. **IPVS** participants will be able to attend talks together in rooms equipped with large screens and room sound system. Questions are asked via personal laptops and/or cell phones in the Zoom chat. **IPVS** participants will share an in-person social program (coffee breaks, lunches, dinners, etc.). **IPVS** speakers will be assigned a room/office to deliver their talks remotely using their laptops, and can then interact in-person with IPVS participants after their virtual talks.
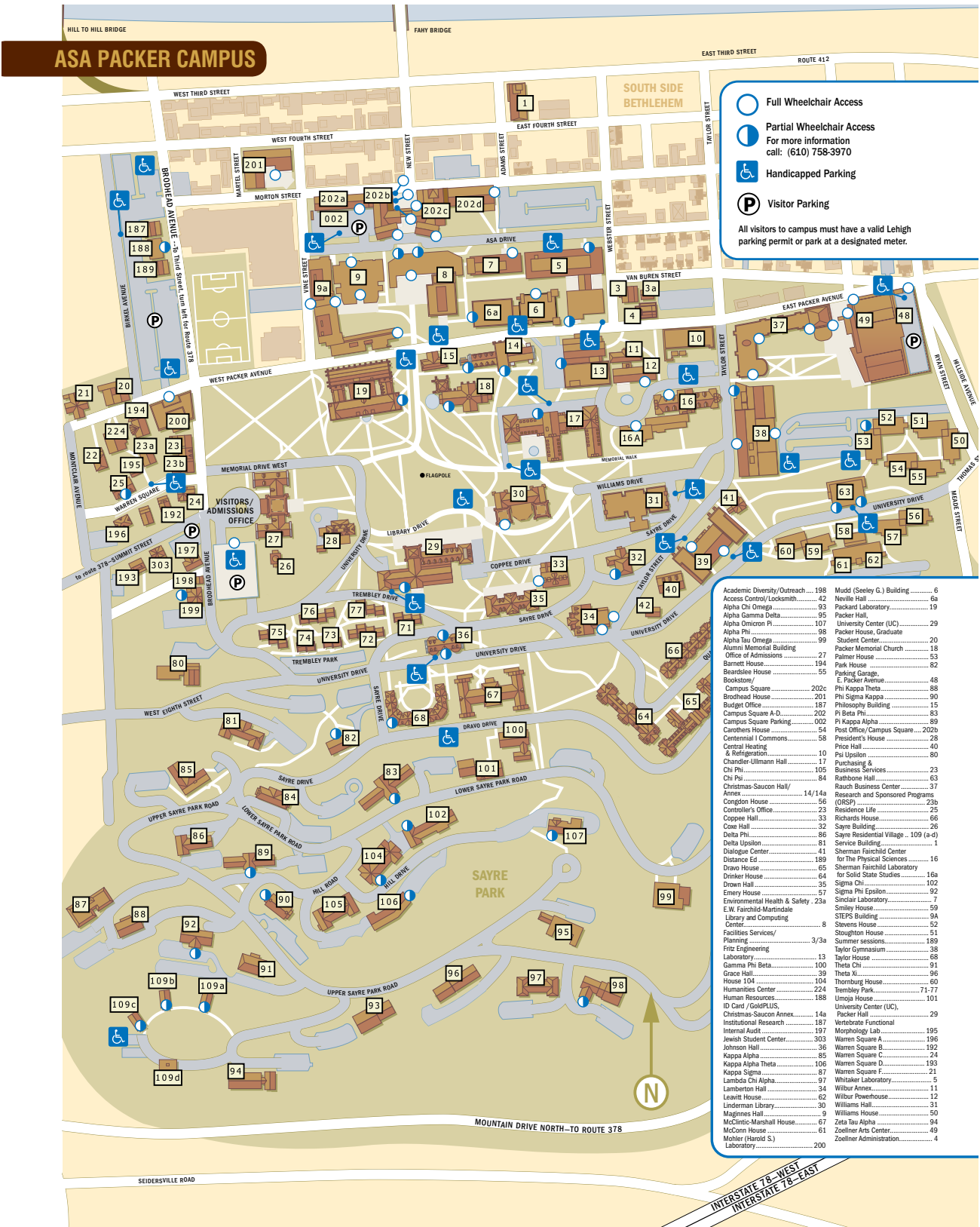
Other details

Information about presentation rooms and zoom links can be found in the table below.

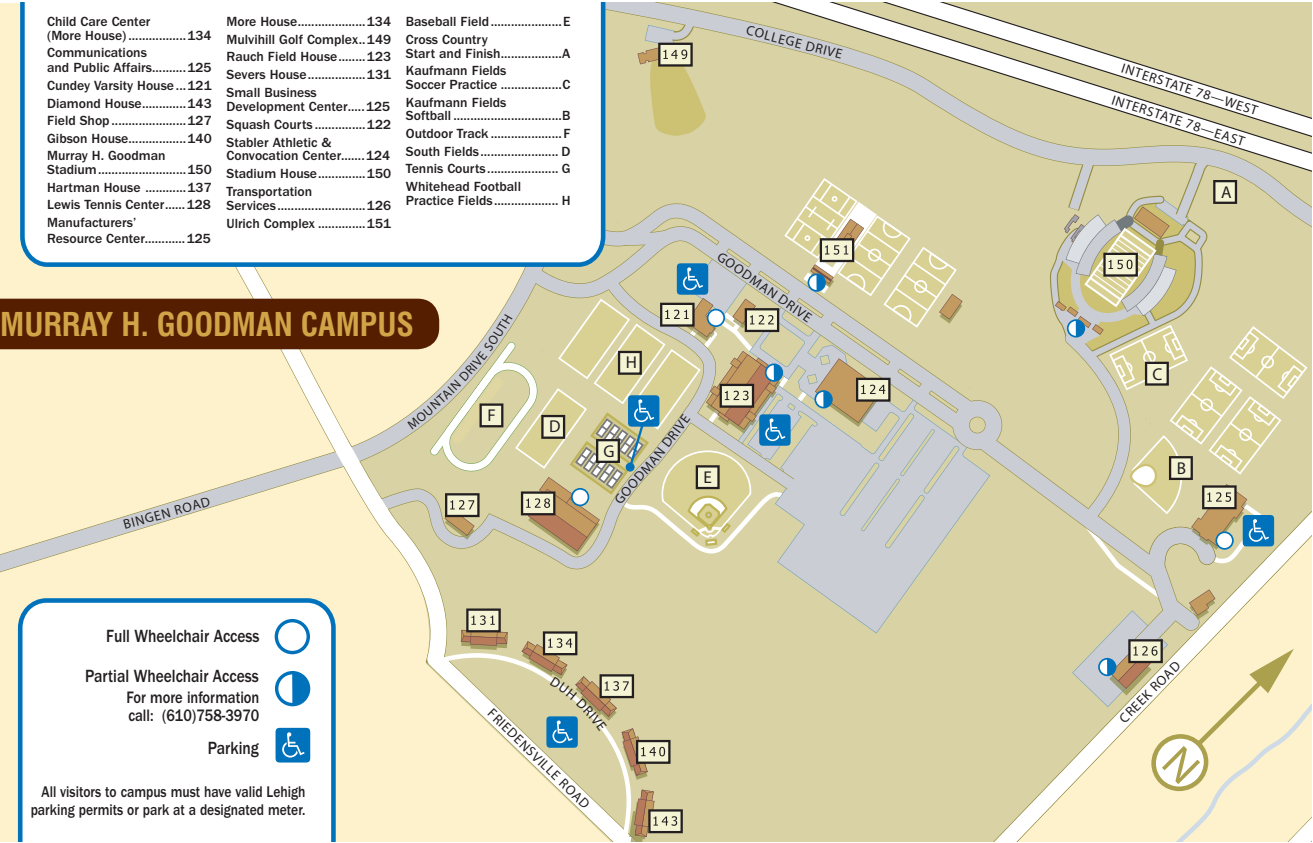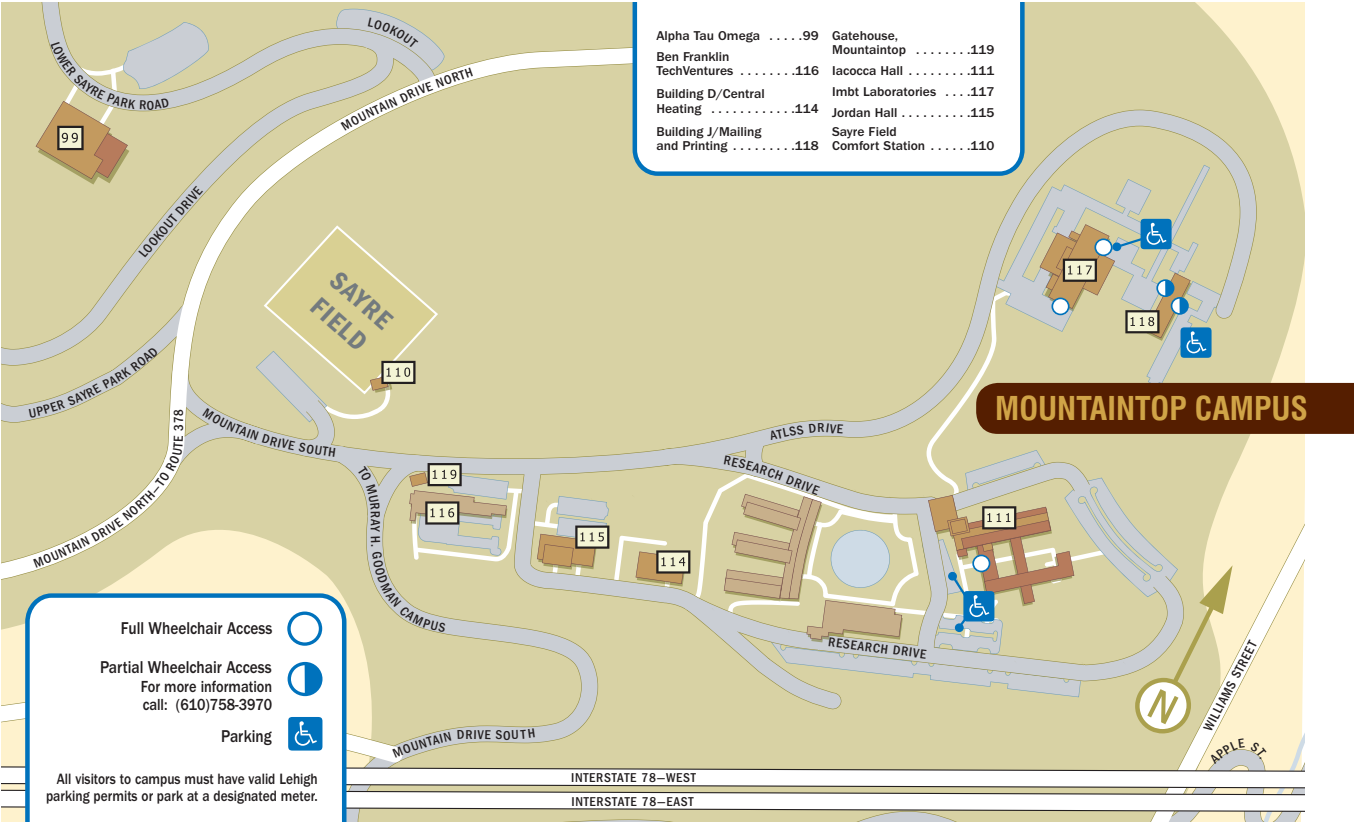| Room Name/Zoom Link	MONRO	LEHIGH	IACOCCA	WHITEHOUSE
Mohler Labs	Mohler 451	Mohler 375	Mohler 453	Mohler 304

- Passwords will be sent to all **registered** participants via email a few days before the conference.

- Plenary talks will be broadcast live in rooms Mohler 451 and Mohler 453, and will be streamed on Zoom Room MONRO.

- The **AIMMS-MOPTA Optimization Modeling Competition** will be held in room Mohler 453, and will be streamed on Zoom Room IACOCCA.

■ Maps

ASA PACKER CAMPUS

Full Wheelchair Access

Partial Wheelchair Access
For more information
call: (610) 758-3970

Handicapped Parking

Visitor Parking

All visitors to campus must have a valid Lehigh
parking permit or park at a designated meter.

SOUTH SIDE
BETHLEHEM

SAYRE PARK

revised August 2013

LEHIGH UNIVERSITY

LEHIGH UNIVERSITY® CAMPUS MAPS

MOUNTAINTOP CAMPUS

Full Wheelchair Access

Partial Wheelchair Access
For more information
call: (610)758-3970

Parking

All visitors to campus must have valid Lehigh
parking permits or park at a designated meter.

MURRAY H. GOODMAN CAMPUS

Full Wheelchair Access

Partial Wheelchair Access
For more information
call: (610)758-3970

Parking

All visitors to campus must have valid Lehigh
parking permits or park at a designated meter.

revised August 2012

■ Plenary Talks

Monday 09:00–10:00. Amy Ellen Mainville Cohn

Addressing Challenges in Multi-Criteria Optimization Problems for Healthcare Personnel Scheduling Problems

University of Michigan

amycohn@umich.edu

Healthcare personnel scheduling problems look on the surface to be clear candidates for combinatorial optimization techniques such as integer programming. A key challenge, however, is the lack of a single, well-defined objective function. Instead, there are many competing metrics of importance including service coverage and patient care, provider well-being and satisfaction, and (in the case of medical trainees) educational considerations. In this talk, I will consider a number of personnel scheduling problems from the University of Michigan health system, focusing not only on the models and algorithms needed to solve these problems, but also on the techniques needed to identify solutions that are desirable and acceptable for practical use.

Speaker Biography. Amy Ellen Mainville Cohn is an Arthur F. Thurnau Professor in the Department of Industrial and Operations Engineering at the University of Michigan, where she also holds an appointment in the Department of Health Management and Policy in the School of Public Health. Dr. Cohn is the Faculty Director of the Center for Healthcare Engineering and Patient Safety (CHEPS). She holds an A.B. in applied mathematics, magna cum laude, from Harvard University and a PhD in operations research from the Massachusetts Institute of Technology. Her primary research interests are in applications of combinatorial optimization, particularly to healthcare and aviation, and to the challenges of optimization problems with multiple objective criteria. She values teaching, mentoring, having a positive impact on society through her work, and helping to foster a vibrant, diverse, nurturing community. She and her husband Jonathan are the proud parents of two sons, Tommy and Peter.

Monday 12:00–13:00. Yurii Nesterov   *Distinguished Spencer C. Schantz Speaker*

Inexact high-order proximal-point methods with auxiliary search procedure

UC Louvain

yurii.nesterov@uclouvain.be

In this paper, we present the further development of Bi-Level Unconstrained Minimization by a new pth-order proximal-point method with the convergence rate $O(1/k^{(1+3p)/2})$, where k is the iteration counter. In this method, we replace the auxiliary line-search procedure by a segment search. This allows bounding its complexity by a logarithm of the desired accuracy. Each step in this search needs an approximate computation of a high-order proximal-point operator. Under assumption on the boundedness of the $(p+1)$th derivative of the objective function, this can be done by one step of the pth order augmented tensor method. In this way, for $p = 2$, we get a new second-order method with the rate of convergence $O(1/k^{7/2})$ and logarithmic complexity of the auxiliary search at each iteration. Another possibility is to compute the proximal-point operator by a lower-order minimization method. As an example, for $p = 3$, we consider the upper-level process convergent as $O(1/k^5)$. Assuming boundedness of the fourth derivative, an appropriate approximation of the proximal-point operator can be computed by a second-order method in a logarithmic number of iterations. This combination gives a second-order scheme with much better complexity than the existing theoretical limits.

Speaker Biography. Born: 1956, Moscow. Master degree 1977, Moscow State University. Doctor degree 1984. Professor at Center for Operations Research and Econometrics, UCLouvain, Belgium. Author of 5 monographs and more than 120 refereed papers in the leading optimization journals. International recognition: Dantzig Prize 2000, John von Neumann Theory Prize 2009, Charles Broyden prize 2010, Francqui Chair (Liege University 2011-2012), SIAM Outstanding paper award (2014), EURO Gold Medal 2016. Main research direction is the development of efficient numerical methods for convex and nonconvex optimization problems supported by the global complexity analysis: general interior-point methods (theory of self-concordant functions), fast gradient methods (smoothing technique), global complexity analysis of second-order and tensor schemes (cubic regularization of the Newton's method), accelerated proximal-point methods.

Spencer C. Schantz Distinguished Lecture Series
This lecture series is endowed in the name of the late **Spencer C. Schantz**, who graduated from Lehigh in 1955 with a B.S. in Industrial Engineering. Following progressive responsibilities with several electrical manufacturing companies, in 1969 he founded U.S. Controls Corporation and became its first CEO and President.
The **Spencer C. Schantz Distinguished Lecture Series** was established by his wife, Jerelyn, as a valuable educational experience for faculty, students and friends of Lehigh?s Industrial and Systems Engineering department.

Monday 15:45–16:45. Catherine McGeoch

Introduction to Quantum Annealing

D-Wave

cmcgeoch@dwavesys.com

Quantum annealing is a novel computational paradigm that runs on purpose-built computing platforms that exploit quantum effects to heuristically solve (NP-hard) problems in combinatorial optimization. The talk will present an overview of how quantum annealing works, and how to use it. It will survey features of D-Wave's Advantage processor, which contains 5000+ qubits and was launched in late 2020, and of the software stack above it.

Speaker Biography.   Catherine McGeoch is on the benchmarking team at D-Wave, charged with evaluating performance of its quantum systems and related software. Before joining D-Wave she held a chaired position in the computer science department at Amherst College, and for thirty years was engaged in research on experimental methods for evaluating algorithm and heuristic performance. She co-founded the (SIAM) ALENEX meetings on algorithms and experiments; co-founded the DIMACS Challenge series; and is past Editor-in-Chief of the ACM Journal on Experimental Algorithmics. She is the author of two books, one on experimental algorithmics, and one on quantum annealing and adiabatic quantum computing.

Tuesday 09:00–10:00. Anna Nagurney

Labor and Supply Chain Networks: Insights from Models Inspired by the COVID-19 Pandemic

UMASS-Amherst

nagurney@isenberg.umass.edu

The COVID-19 pandemic has dramatically illustrated the importance of labor in supply chain networks in numerous economic sectors, from agriculture to healthcare. In this talk, I will discuss our recent research on the inclusion on labor in supply chains both in optimization and game theory frameworks to elucidate the impacts of disruptions of labor in terms of availability as well as productivity on product flows, prices, as well as the profits of firms. I will also highlight what can be done to ameliorate negative impacts and will also discuss the power of setting appropriate wages on supply chain links from production and transportation to storage and ultimate distribution to points of demand.

Speaker Biography.   Anna Nagurney is the Eugene M. Isenberg Chair in Integrative Studies at the Isenberg School of Management at the University of Massachusetts Amherst and the Director of the Virtual Center for Supernetworks, which she founded in 2001. She holds ScB, AB, ScM and PhD degrees from Brown University in Providence, RI. She is the author/editor of 15 books, more than 200 refereed journal articles, and over 50 book chapters. Professor Nagurney has been a Fulbrighter twice (in Austria and Italy), was a Visiting Professor at the School of Business, Economics and Law at the University of Gothenburg in Sweden, and was a Distinguished Guest Visiting Professor at the Royal Institute of Technology (KTH) in Stockholm. She was a Visiting Fellow at All Souls College at Oxford University during the 2016 Trinity Term and a Summer Fellow at the Radcliffe Institute for Advanced Study at Harvard in 2017 and 2018. Anna has held visiting appointments at MIT and at Brown University and was a Science Fellow at the Radcliffe Institute for Advanced Study at Harvard University in 2005-2006. She has been recognized for her research on networks with the Kempe Prize from the University of Umea, the Faculty Award for Women from the US National Science Foundation, the University Medal from the University of Catania in Italy, the 2019 Constantin Caratheodory Prize, and the 2020 Harold Larnder Prize, and was elected a Fellow of the RSAI (Regional Science Association International) as well as INFORMS (Institute for Operations Research and the Management Sciences) and the Network Science Society, among other awards. She has also been recognized with several awards for her mentorship of students and her female leadership with the WORMS Award, for example. Her research has garnered support from the AT&T Foundation, the Rockefeller Foundation through its Bellagio Center programs, the Institute for International Education, and the National Science Foundation. She has given plenary/keynote talks and tutorials on 5 continents. She is an active member of professional societies, including INFORMS, POMS, and RSAI.

Tuesday 12:00–13:00. Jeffrey T. Linderoth

High-Rank Matrix Completion by Integer Programming

UW-Madison

linderoth@wisc.edu

In the High-Rank Matrix Completion (HRMC) problem, we are given a collection of n data points, arranged into columns of a matrix $X \in \mathbb{R}^{d \times n}$, and each of the data points is observed only on a subset of its coordinates. The data points are assumed to be concentrated near a union of low-dimensional subspaces. The goal of HRMC is to recover the missing elements of the data matrix X. State-of-the-art algorithms for HRMC can fail on instances with a large amount of missing data or if the data matrix X is nearly full-rank. We propose a novel integer programming based approach for HRMC. The approach is based on dynamically determining a set of candidate subspaces and optimally assigning points to selected subspaces. The problem structure is identical to the classical facility-location problem, with subspaces playing the role of facilities and data points that of customers. We propose a column-generation approach for identifying candidate subspaces combined with a Benders decomposition approach for solving the linear programming relaxation of the formulation. An empirical study demonstrates that the proposed approach can achieve better clustering accuracy than state-of-the-art methods when the data is high-rank, the percentage of missing data is high, or there are a small number of data points in each subspace. This is joint work with Jim Luedtke, Daniel Pimentel-Alarcón and Akhilesh Soni, all of UW-Madison.

Speaker Biography.   Jeff Linderoth is the Harvey D. Spangler Professor of the department of Industrial and Systems Engineering at the University of Wisconsin-Madison. Prof. Linderoth holds a courtesy appointment in the Computer Sciences department and as a Discovery Fellow in the Optimization Theme at the Wisconsin Institutes of Discovery. Dr. Linderoth received his Ph.D. degree from the Georgia Institute of Technology in 1998. He was previous employed in the Mathematics and Computer Science Division at Argonne National Laboratory, with the optimization-based financial products firm of Axioma, and as an Assistant Professor at Lehigh University. His awards include an Early Career Award from the Department of Energy, the SIAM Activity Group on Optimization Prize, and the INFORMS Computing Society (ICS) Prize. He 2016, he was elected to membership as an INFORMS Fellow.

Tuesday 15:45–16:45. Simge Küçükyavuz

Strong formulations for Joint Chance-Constrained Programs

Northwestern University

simge@northwestern.edu

An important substructure in modeling joint linear chance-constrained programs with finite sample space is the intersection of mixing sets with common binary variables. In this talk, we first revisit basic mixing sets by establishing a strong and previously unrecognized connection to submodularity. We show that mixing inequalities with binary variables are nothing but the polymatroid inequalities associated with a specific submodular function. This submodularity viewpoint enables us to unify and extend existing results on valid inequalities and convex hulls of the intersection of multiple mixing sets with common binary variables. Next, we consider exact deterministic reformulations of distributionally robust chance-constrained programs (DR-CCP) over Wasserstein ambiguity sets. The existing formulations are known to have weak continuous relaxation bounds, and, consequently, for hard instances with small radius, or with large problem sizes, the branch-and-bound based solution processes suffer from large optimality gaps even after hours of computation time. Motivated by these challenges, we conduct a polyhedral study to strengthen these formulations. We reveal several hidden connections between DR-CCP and its nominal counterpart (the sample average approximation), mixing sets, and robust 0-1 programming. By exploiting these connections in combination, we provide an improved formulation and new valid inequalities for DR-CCP. We test the impact of our results on a stochastic transportation problem numerically. Our experiments demonstrate that our proposed approach reduces the overall solution times from hours to seconds. This is joint work with Nam Ho-Nguyen, Fatma Kilinc-Karzan, and Dabeen Lee.

Speaker Biography. Simge Küçükyavuz is a professor in the Industrial Engineering and Management Sciences Department at Northwestern University. She is an expert in mixed-integer programming, large-scale optimization, and optimization under uncertainty. Her research has been supported by multiple grants from the National Science Foundation (NSF) and Office of Naval Research. She is the recipient of the NSF CAREER Award and a co-winner of the INFORMS Computing Society (ICS) Prize. She is the Chair of ICS and serves on the editorial boards of Mathematics of Operations Research, Mathematical Programming (B,C), SIAM Journal on Optimization, and MOS-SIAM Optimization Book Series.

Wednesday 09:00–10:00. Hamsa Balakrishnan

Towards a Fair and Efficient Air Transportation System

MIT

hamsa@mit.edu

Efficiency âĂŞ for example, minimizing delays âĂŞ has traditionally been the primary objective of air traffic management algorithms. With the emergence of new vehicle types (e.g., drones and urban air mobility) and aircraft operators, there is increased competition for constrained airspace resources. In such competitive environments, fairness and privacy, in addition to efficiency, become important considerations in traffic management and scheduling. In this talk, I will present some of our recent work on balancing the tradeoffs between efficiency and fairness in transportation systems. We will focus on two contexts: First, we consider the perspective of vehicle operators while equitably allocating delays during traffic congestion management. Next, we propose a scalable approach to balancing efficiency and fairness in vehicle routing, which has applications to rideshare platforms and multiplexing tasks on vehicles.

Speaker Biography. Hamsa Balakrishnan is the William E. Leonhard (1940) Professor of Aeronautics and Astronautics at the Massachusetts Institute of Technology (MIT). She received her PhD from Stanford University, and a B.Tech. from the Indian Institute of Technology Madras. Her research is in the design, analysis, and implementation of control and optimization algorithms for cyber-physical infrastructures, with an emphasis on air transportation. She is the co-founder and chief scientist of Lumo, a Boston-based travel startup. Prof. Balakrishnan is the recipient of an NSF CAREER Award in 2008, the CNA Award for Operational Analysis in 2012, the AIAA Lawrence Sperry Award in 2012, the American Automatic Control Council's Donald P. Eckman Award in 2014, the MIT AIAA Undergrad Advising (2014) and Undergraduate Teaching (2019) Awards, and multiple best paper awards. She is an Associate Fellow of the American Institute of Aeronautics and Astronautics.

Wednesday 12:00–13:00. Michael Mahoney

Incorporating second order ideas into first class machine learning methods

UC Berkeley

mmahoney@stat.berkeley.edu

At one point, second order optimization methods were obviously good to use to solve machine learning problems; then, second order methods were obviously bad to use to solve machine learning problems; and now, we have an increasingly fine understanding of how to incorporate second order information and ideas into machine learning methods. While existing computational infrastructure provides a strong bias toward first order methods, first order methods actually don't perform so well out of the box, and so one must introduce lots of hyperparameterized variants to make them work well. These hyperparameters are often aimed to address problems second order ideas are designed to address, but their use introduces a whole host of other problems, not the least of which is the need for extremely expensive hyperparameter search and the associated model instability. In this talk, we'll cover recent results on incorporating second order ideas directly into first class state-of-the-art machine learning methods: some theory; some implementations; and some applications.

Speaker Biography. Michael W. Mahoney is at the University of California at Berkeley in the Department of Statistics and at the International Computer Science Institute (ICSI). He is also a faculty scientist at the Lawrence Berkeley National Laboratory. He works on algorithmic and statistical aspects of modern large-scale data analysis. Much of his recent research has focused on large-scale machine learning, including randomized matrix algorithms and randomized numerical linear algebra, geometric network analysis tools for structure extraction in large informatics graphs, scalable implicit regularization methods, computational methods for neural network analysis, physics informed machine learning, and applications in genetics, astronomy, medical imaging, social network analysis, and internet data analysis. He received his PhD from Yale University with a dissertation in computational statistical mechanics, and he has worked and taught at Yale University in the mathematics department, at Yahoo Research, and at Stanford University in the mathematics department. Among other things, he is on the national advisory committee of the Statistical and Applied Mathematical Sciences Institute (SAMSI), he was on the National Research Council's Committee on the Analysis of Massive Data, he co-organized the Simons Institute's fall 2013 and 2018 programs on the foundations of data science, he ran the Park City Mathematics Institute's 2016 PCMI Summer Session on The Mathematics of Data, and he runs the biennial MMDS Workshops on Algorithms for Modern Massive Data Sets. He is the Director of the NSF/TRIPODS-funded FODA (Foundations of Data Analysis) Institute at UC Berkeley. More information is available at https://www.stat.berkeley.edu/~mmahoney/.

■ AIMMS/MOPTA Optimization Modeling Competition 2021

The **13th AIMMS-MOPTA Optimization Modeling Competition** is a result of cooperation between **AIMMS** and the organizers of the **MOPTA** conference. Teams of three graduate students participated and solved a challenging home service assignment, routing, and appointment scheduling (H-SARA) problem under uncertainty. The teams had to form a mathematical model of the problem, implement it in **AIMMS**, solve it, create a graphical user interface, and write a 15-page report on the project.

For more information about the competition and the full problem description, please see https://coral.ise.lehigh.edu/~mopta/competition.

Problem Description

Home Service Assignment, Routing, and Scheduling with Stochastic Service Time, Travel time, and Cancellation

Home services are the provision of providing services to people at their homes, e.g., home health care, home beauty care, banking service, appliance repair service, and more. The home service industry, especially home healthcare, has been rapidly growing worldwide due to emergent changes in family structures, work obligations, aging populations, and outspread of chronic and infectious diseases. In 2018, the global home service market was valued at \$282 Billion and is expected to grow by 18.91% annually from 2019-2026, reaching \$1,133.4 Billion by 2026 (Verified Market Research, 2019[1]). Accordingly, the development of computationally efficient and implementable tools is essential to support decision-making in all areas of the home service industry.

Home services require professional service teams (or one operator) to travel for delivering the services to geographically distributed customers. Service providers often quote an appointment time (planned service start time) to each customer in advance to avoid delivery failure. Therefore, when home service providers plan for service, they need to decide: how many service teams to hire (i.e., sizing problem), how to assign service teams to the customers (i.e., an assignment problem), how to route service teams (i.e., a vehicle routing problem), and how to assign appointment times for the customers (i.e., appointment scheduling problem). This competition aims to address these specific critical operational decision problems in home service. Stochasticity is an intrinsic property of the home service assignment, routing, and appointment scheduling (H-SARA) problem. Here, we will focus on solving H-SARA under three key random factors: service duration, travel time, and customer cancellation. Your team's goal is to develop an efficient and implementable method to solve H-SARA, which will be used by home service providers.

Finalists

We are happy that 14-teams from around the world registered for the competition.

The panel of judges (Karmel Shehadeh from Lehigh University and Mohan Chiriki from AIMMS) selected the following three teams as finalists:

<div align="center">

Team "Flamerunners", Brown University
Enrique Areyan Viqueira, Shamay Samuel
advised by Serdar Kadioglu.

Team "SolvED", University of Edinburgh
Shunee Johnn, Yiran Zhu, Andreǎs Miniguano-Trujillo
advised by Akshay Gupte.

Team "The Optimistics", Universidad de los Andes
Daniel Yamin, Daniel Barahona
advised by Alfaima Solano.

</div>

The finalist teams will each give 30-minute presentations (25 minutes for the talk + 5 minutes for questions) on their work on **Tuesday, August 3rd, starting at 10:15am**. The winning team will be announced during the **AIMMS-MOPTA Modeling Competition & Plenary Award Session** on Tuesday (Prizes: 1st place \$1200; 2nd place \$600; 3rd place \$300). The panel of judges will consist of: Karmel S. Shehadeh, Ana-Iulia Alexandrescu and Luis Zuluaga (Lehigh University), Albert S. Berahas (University of Michigan), Mohan Chiriki (AIMMS).

We thank all the teams for their participation. We believe that it has been a very positive experience for all parties involved in the process.

---

[1]Verified Market Research. Global home services market by deployment, by geographic scope and forecast to 2026., Published Date: April 2019.

■ Instructions

MOPTA 2021 will be offered in an innovative dual **VIRT/IPVS** mode. We will have both **VIRT** (virtual) and **IPVS** (in-person) participants, speakers and session chairs. The instructions below apply to all attendees of **MOPTA 2021**.

For Speakers (VIRT/IPVS)

- We ask all speakers to be familiar with the time and the location (virtual and/or in person) of their stream and talk, as specified in the conference booklet.

- Speakers should arrive at the location of their stream and talk 10 minutes prior to the scheduled start time of the session.

- Upon arrival you will be met (virtually or in person) by the chair of the session. Please introduce yourself and, if applicable, test your presentation (share screen and presentation on zoom).

- Talks are strictly 25 minutes long plus 5 minutes for questions and answers. Anyone going over this time will be asked to stop by the chair.

- To aid you with the timing of your presentation, the session chairs are asked to send 1 minute warning messages to presenters.

For Chairs (VIRT/IPVS)

- Please arrive at the appropriate location (virtual and/or in person) 10 minutes before the start of the stream you will be chairing.

- In the event of a problem you should immediately seek the help of one of the conference organizers.

- Delegates presenting in the stream should also be present in the seminar room 10 minutes before the start of the stream. You should introduce yourself to the speakers. They will test out their presentation on the virtual platform prior to the start of the session.

- Your main role will be to ensure that the stream runs to time. The speaker has 25 minutes for presentation followed by 5 minutes of questions and answers. Each talk is followed by a 5 minute break for the comfort of the audience and to allow for movement between streams.

- If a speaker fails to show for their talk, advise the audience to attend a talk in an alternative seminar room. Please, do not move the next talk forward.

- Before each speaker presents, you should introduce them and remind the audience that all interruptions and questions are to be reserved until the scheduled 5 minute question and answer session following each presentation.

- Should a speaker overrun, you must politely but firmly stop their presentation and move on to the question and answer section of the time slot.

- After each talk, thank the speaker, encourage applause, and open the floor to questions.

■ Conference Sponsors

■ MOPTA 2021 Committee

Albert S. Berahas
MOPTA Conference Chair

Karmel S. Shehadeh
MOPTA Competition Chair
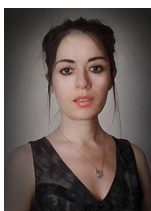
Tamás Terlaky

Frank E. Curtis
SIAM Representative

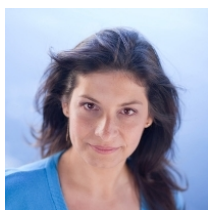**Lawrence V. Snyder**
Co-director of IDISC

Luis Nunes Vicente
Department Chair

Martin Takáč

Aida Khajavirad
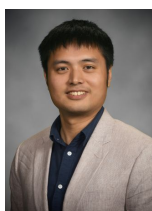
Daniel P. Robinson

Monica Gabriela Cojocaru

Luis Zuluaga

Katya Scheinberg

Xiu Yang
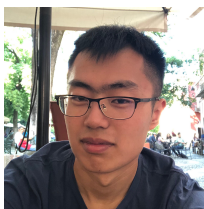
Boris Defourny

Alberto Lamadrid

Ted K. Ralphs

Baoyu Zhou

PROGRAM

■ MOPTA 2021 Program

MOPTA 2021

Conference information (passwords will be sent to all registered participants)

Room Name		MONRO	LEHIGH	IACOCCA	WHITEHOUSE
Room		Mohler 451	Mohler 375	Mohler 453	Mohler 304
Zoom Room		https://lehigh.zoom.us/j/99177535128	https://lehigh.zoom.us/j/97019673134	https://lehigh.zoom.us/j/99156475897	https://lehigh.zoom.us/j/96392454462

EST	**Monday August 2, 2021**				
08:30-08:45	Registration				
08:45-09:00	Opening remarks				
09:00-10:00	Plenary Talk #1	**Amy Cohn**			
10:00-10:15	Coffee break				
10:15-11:45	Parallel Technical Sessions M1	Nonlinear and Stochastic Optimization Algorithms	Constrained Optimization	Nonconvex optimization - Part (i)	Optimization and Healthcare
11:45-12:00	Coffee break				
12:00-13:00	Plenary Talk #2	**Yurii Nesterov**			
13:00-14:00	Lunch break				
14:00-15:30	Parallel Technical Sessions M2	Methods for Large-Scale, Nonlinear and Stochastic Opt	Quantum Optimization I	Nonconvex optimization, Part (ii)	OR in Healthcare
15:30-15:45	Coffee break				
15:45-16:45	Plenary Talk #3	**Catherine McGeoch**			
16:45-18:15	Parallel Technical Sessions M3	Advances in Large-Scale Nonlinear Optimization	Quantum Optimization II	Advances in Nonconvex Optimization	Optimization with Applications in Healthcare
18:30-20:30	Student Social				

	Tuesday August 3, 2021				
08:30-09:00	Registration				
09:00-10:00	Plenary Talk #4	**Anna Nagurney**			
10:00-10:15	Coffee break				
10:15-11:45	Parallel Technical Sessions T1	Advances in Nonlinear ADMM and Related Comp. Methods	Algorithms for Derivative-Free Optimization	**AIMMS-MOPTA Optimization Modeling Competition**	-
11:45-12:00	Coffee break				
12:00-13:00	Plenary Talk #5	**Jeffrey T. Linderoth**			
13:00-14:00	Lunch break				
14:00-15:30	Parallel Technical Sessions T2	Advances in Stochastic Optimization	New Deterministic and Stochastic Methods for DFO I	Distributed Optimization	Robust and Stochastic Optimization under Uncertainty
15:30-15:45	Coffee break				
15:45-16:45	Plenary Talk #6	**Simge Kucukyavuz**			
16:45-18:15	Parallel Technical Sessions T3	First Order Methods	New Deterministic and Stochastic Methods for DFO II	Robust Learning	Optimization & Machine Learning
18:15-18:30	AIMMS-MOPTA Modeling Competition & Plenary Award Session				
19:00	Dinner (for in person participants)				

	Wednesday August 4, 2021				
08:30-09:00	Registration				
09:00-10:00	Plenary Talk #7	**Hamsa Balakrishnan**			
10:00-10:15	Coffee break				
10:15-11:45	Parallel Technical Sessions W1	Large Scale Optimization	Optimization and Energy	Fairness in Machine Learning	Applications of Optimization
11:45-12:00	Coffee break				
12:00-13:00	Plenary Talk #8	**Michael Mahoney**			
13:00-14:00	Lunch break				
14:00-15:30	Parallel Technical Sessions W2	Higher Order Optimization Methods	Energy and Optimization	Optimization & Applications 1	-
15:30-15:45	Closing remarks				

08:00–08:45   **Registration**

08:45–09:00   **Opening Remarks**

09:00–10:00   **Plenary talk** (Mohler 451, Mohler 453, MONRO)   *Chair: Karmel S. Shehadeh*

- **Addressing Challenges in Multi-Criteria Optimization Problems for Healthcare Personnel Scheduling Problems** *Amy Ellen Mainville Cohn*

10:00–10:15   **Break**

10:15–11:45   **Nonlinear and Stochastic Optimization Algorithms** (Mohler 451, MONRO)   *Chair: Baoyu Zhou*

1. **A Theoretical and Empirical Comparison of Gradient Approximations Methods in Derivative-Free Optimization** *Liyuan Cao*
2. **Efficient algorithms for some variants of the extended trust-region subproblems** *Maziar Salahi*
3. **Generalized Cyclic Stochastic Approximation and its Application in Multi-agent Systems** *Jiahao Shi*

10:15–11:45   **Constrained Optimization** (Mohler 375, LEHIGH)   *Chair: Negar Soheili Azad*

1. **New notions of simultaneous diagonalizability of quadratic forms with applications to QCQPs** *Alex Wang*
2. **Second-Order Conic and Polyhedral Approximations of the Exponential Cone: Application to Mixed-Integer Exponential Conic Programs** *Qing Ye*
3. **On the rescaling and projection algorithm** *Negar Soheili Azad*

10:15–11:45   **Nonconvex optimization - Part (i)** (Mohler 453, IACOCCA)   *Chair: Aida Khajavirad*

1. **Proximity in Concave Integer Quadratic Programming** *Alberto Del Pia*
2. **Two-halfspace closure** *Amitabh Basu*
3. **Convexification of the Lennard-Jones Potential** *Anatoliy Kuznetsov*

10:15–11:45   **Optimization and Healthcare** (Mohler 304, WHITEHOUSE)   *Chair: Karmel S. Shehadeh*

1. **Distributionally Robust Optimization Approaches for a Mobile Facility Routing and Scheduling Problem** *Karmel Shehadeh*
2. **Distributionally Robust Home Service Routing and Appointment Scheduling with Random Travel and Service Times** *Man Yiu Tsang*
3. **A Voice base Nonsmooth Nonconvex optimization model for Parkinson Disease Severity Estimation.** *Habib Ghaffari Hadigheh*

11:45–12:00   **Break**

12:00–13:00   **Plenary talk,** *Distinguished Spencer C. Schantz Speaker* (Mohler 451, Mohler 453, MONRO)   *Chair: Frank E. Curtis*

- **Inexact high-order proximal-point methods with auxiliary search procedure** *Yurii Nesterov*

13:00–14:00   **Lunch**

14:00–15:30   **Methods for Large-Scale, Nonlinear and Stochastic Optimization** (Mohler 451, MONRO)   *Chair: Jiahao Shi*

1. **SQP for Nonlinear Equality Constrained Stochastic Optimization** *Baoyu Zhou*
2. **Diagonalized Hessian Estimates for Convex and Nonconvex Optimization and Comparison with Natural Gradient Method** *Shiqing Sun*
3. **Complexity of Projected Newton Methods for Bound-constrained Optimization** *Yue Xie*

14:00–15:30   **Quantum Optimization I** (Mohler 375, LEHIGH)   *Chair: Brandon Augustino*

1. **An Inexact-Feasible Interior Point Method for Linear Optimization with High Adaptability to Quantum Computers** *Mohammadhossein Mohammadisiahroudi*
2. **Quantum Interior Point Methods for Semidefinite Optimization** *Brandon Augustino*
3. **Characterization and Mitigation of Errors in Quantum Computing via Consistent Bayesian** *Muqing Zheng*

14:00–15:30   **Nonconvex optimization, Part (ii)** (Mohler 453, IACOCCA)   *Chair: Aida Khajavirad*

1. **Sum-of-squares lower bounds for random combinatorial problems** *Dmitriy Kunisky*
2. **Sparse regression: decompositions, convexifications and algorithms** *Andres Gomez*
3. **Rank Pump: A Feasibility Heuristic For Polynomial Optimization** *Chen Chen*

14:00–15:30   **OR in Healthcare** (Mohler 304, WHITEHOUSE)   *Chair: Tugce Isik*

1. **Solution Methods for Integrated Surgery Scheduling and Inventory Problem** *Amogh Bhosekar*
2. **Dynamic Tuberculosis Screening for Healthcare Employees** *Mahsa Kiani*

3. **Optimal Nurse Allocation for the Surgical System: A Tandem Network with Flexible Servers** *Tong Zhang*

15:30–15:45   **Break**

15:45–16:45   **Plenary talk** (Mohler 451, Mohler 453, MONRO)   *Chair: Tamas Terlaky*

- **Introduction to Quantum Annealing** *Catherine McGeoch*

16:45–18:15   **Advances in Large-Scale Nonlinear Optimization** (Mohler 451, MONRO)   *Chair: Raghu Bollapragada*

1. **Average Curvature FISTA for Nonconvex Smooth Composite Optimization Problems** *Jiaming Liang*
2. **Implicit Regularization of Sub-Gradient Method in Robust Matrix Recovery: Don't be Afraid of Outliers** *Jianhao Ma*
3. **Inexact Proximal Gradient Methods** *Daniel Robinson*

16:45–18:15   **Quantum Optimization II** (Mohler 375, LEHIGH)   *Chair: Brandon Augustino*

1. **Quantum-inspired formulations for the max k-cut problem** *Ramin Fakhimi*
2. **Improving QAOA with Warm-Start Initializations and Custom Mixers** *Reuben Tate*
3. **Qubo Reformulations of Combinatorial Optimization Problems** *Rodolfo Alexander Quintero Ospina*

16:45–18:15   **Advances in Nonconvex Optimization** (Mohler 453, IACOCCA)   *Chair: Meisam Razaviyayn*

1. **A Riemannian Block Coordinate Descent Method for Computing the Projection Robust Wasserstein Distance** *Minhui Huang*
2. **A Smoothing Scheme for Nonconvex-Concave Min-Max Problems** *Weiwei Kong*
3. **Smooth nonconvex-nonconcave min-max optimization problems with small inner maximization constraint set** *Meisam Razaviyayn*

16:45–18:15   **Optimization with Applications in Healthcare** (Mohler 304, WHITEHOUSE)   *Chair: Secil Sozuer*

1. **Data-Driven Distributionally Robust Surgery Planning in Flexible Operating Rooms Over a Wasserstein Ambiguity** *Karmel S. Shehadeh*
2. **Simulation-Based Optimization of Dynamic Appointment Scheduling Problem with Patient Unpunctuality and Provider Lateness** *Secil Sozuer*
3. **Supply Side Flexibility in Revenue Management: An Application in Medical Wire and Device Manufacturing** *Cigdem Gurgur*

18:30–20:30   **Student Social**

08:30–09:00 **Registration**

09:00–10:00 **Plenary talk** (Mohler 451, Mohler 453, MONRO)   *Chair: Karmel S. Shehadeh*

- **Labor and Supply Chain Networks: Insights from Models Inspired by the COVID-19 Pandemic** *Anna Nagurney*

10:00–10:15 **Break**

10:15–11:45 **Advances in Nonlinear ADMM and Related Computational Methods** (Mohler 451, MONRO) *Chair: Dimitri Papadimitriou*

1. **Escaping from Zero Gradient: Revisiting Action-Constrained Reinforcement Learning via Frank-Wolfe Policy Optimization** *Ping-Chun Hsieh*
2. **Relaxed alternating minimization algorithm for convex optimization problems in image denoising** *Yuchao Tang*
3. **Power of Alternating Direction Method of Multipliers (ADMM) in Deep Learning** *Junxiang Wang*

10:15–11:45 **Algorithms for Derivative-Free Optimization** (Mohler 375, LEHIGH)   *Chair: Baoyu Zhou*

1. **Zeroth-Order Riemannian Optimization** *Jiaxiang Li*
2. **New Hybrid Algorithms for Global and Local Derivative-Free Optimization** *Ahmad Almomani*
3. **Manifold Sampling for Optimizing Nonsmooth Nonconvex Compositions** *Baoyu Zhou*

10:15–11:45 **AIMMS-MOPTA Optimization Modeling Competition** (Mohler 453, IACOCCA)   *Chair: Karmel Shehadeh*

11:45–12:00 **Break**

12:00–13:00 **Plenary talk** (Mohler 451, Mohler 453, MONRO)   *Chair: Daniel P. Robinson*

- **High-Rank Matrix Completion by Integer Programming** *Jeffrey T. Linderoth*

13:00–14:00 **Lunch**

14:00–15:30 **Advances in Stochastic Optimization** (Mohler 451, MONRO)   *Chair: Dimitri Papadimitriou*

1. **A Stochastic Decoupling Method for Minimizing the Sum of Smooth and Non-Smooth Functions** *Konstantin Mishchenko*
2. **An Online Algorithm for Maximum-Likelihood Quantum State Tomography** *Yen-Huan Li*
3. **Perturbed Fenchel duality and first-order methods** *Javier Pena*

14:00–15:30 **New Deterministic and Stochastic Methods for Derivative-Free Optimization I** (Mohler 375, LEHIGH)

Chair: Tommaso Giovannelli

1. **Full-Low Evaluation Methods for Derivative-Free Optimization** *Oumaima Sohab*
2. **Randomized DFO Methods for Fitting Numerical Physics Models** *Matt Menickelly*
3. **A derivative-free method for structured optimization problems** *Andrea Cristofari*

14:00–15:30 **Distributed Optimization** (Mohler 453, IACOCCA)   *Chair: Charikleia Iakovidou*

1. **A flexible framework for noisy and nonconvex distributed optimization** *Charikleia Iakovidou*
2. **Scalable Compressed Communication in Distributed Inference and Optimization** *César A Uribe*
3. **Exploiting Shared Representations for Personalized Federated Learning** *Liam Collins*

14:00–15:30 **Robust and Stochastic Optimization under Uncertainty** (Mohler 304, WHITEHOUSE)   *Chair: Weijun Xie*

1. **Distributionally Robust Optimization with Markovian Data** *Mengmeng Li*
2. **Affine and Constant Policies in Adjustable Linear Robust Optimization: A New Perspective** *Ningji Wei*
3. **ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs** *Weijun Xie*

15:30–15:45 **Break**

15:45–16:45 **Plenary talk** (Mohler 451, Mohler 453, MONRO)   *Chair: Albert S. Berahas*

- **Strong formulations for Joint Chance-Constrained Programs** *Simge Küçükyavuz*

16:45–18:15 **First Order Methods** (Mohler 451, MONRO)   *Chair: Dimitri Papadimitriou*

1. **Stochastic nonlinear ADMM** *Dimitri Papadimitriou*
2. **A Communication Compression Decentralized Algorithm for Convex Composite Optimization** *Yao Li*
3. **Online Nash Social Welfare Maximization with Predictions** *Billy Zhengxu Jin*

16:45–18:15 **New Deterministic and Stochastic Methods for Derivative-Free Optimization II** (Mohler 375, LEHIGH)

Chair: Tommaso Giovannelli

1. **A linesearch-based derivative-free approach for mixed-integer nonsmooth constrained optimization** *Tommaso Giovannelli*
2. **Hessian-Aided Random Perturbation (HARP) Using Noisy Zeroth-Order Queries** *Jingyi Zhu*
3. **High Probability Complexity Bounds for Line Search Based on Stochastic Oracles** *Miaolan Xie*

16:45–18:15   **Robust Learning** (Mohler 453, IACOCCA)   *Chair: Guangyi Liu*

1. **Practical Convex Formulation of Robust Neural Network Training** *Yatong Bai*
2. **Robust Learning of Recurrent Neural Networks in Presence of Exogenous Noise** *Arash Amini*
3. **Robust Classification with Localized Observations Using Stable Recurrent Neural Networks** *Guangyi Liu*

16:45–18:15   **Optimization & Machine Learning** (Mohler 304, WHITEHOUSE)   *Chair: Matthias Chung*

1. **Optimal Decision in a Statistical Process Control with Cubic Loss** *Vladimir Turetsky*
2. **Big Data Inverse Problems** *Matthias Chung*

18:15–18:30   **AIMMS-MOPTA Modeling Competition & Plenary Award Session** (Mohler 451, Mohler 453, MONRO)

19:30–01:00   **Dinner**

08:30–09:00   **Registration**

09:00–10:00   **Plenary talk** (Mohler 451, Mohler 453, MONRO)   *Chair: Tamas Terlaky*

- **Towards a Fair and Efficient Air Transportation System** *Hamsa Balakrishnan*

10:00–10:15   **Break**

10:15–11:45   **Large Scale Optimization** (Mohler 451, MONRO)   *Chair: Salar Fattahi*

1. **Large-scale Inference of Sparsely-varying Markov Random Fields** *Salar Fattahi*
2. **Global optimization using random embeddings** *Estelle Massart*
3. **REX: Revisiting Budgeted Training with an Improved Schedule** *John Chen*

10:15–11:45   **Optimization and Energy** (Mohler 375, LEHIGH)   *Chair: Ben Wang*

1. **Building Load Control using Distributionally Robust Chance-Constrained Programs with Right-Hand Side Uncertainty and the Risk-Adjustable Variants** *Yiling Zhang*
2. **A Semidefinite Optimization-based Branch-and-Bound Algorithm for Several Reactive Optimal Power Flow Problems** *Miguel Anjos*
3. **Battery-Wind/Solar Farm Management: A Continuous Dynamic Programming Approach** *Ben Wang*

10:15–11:45   **Fairness in Machine Learning** (Mohler 453, IACOCCA)   *Chair: Suyun Liu*

1. **Automating Procedurally Fair Feature Selection in Machine Learning** *Clara Belitz*
2. **Balanced Districting on Grid Graphs withProvable Compactness and Contiguity** *Cyrus Hettle*
3. **A Stochastic Alternating Balance k-Means Algorithm for Fair Clustering** *Suyun Liu*

10:15–11:45   **Applications of Optimization** (Mohler 304, WHITEHOUSE)   *Chair: Robert Bosch*

1. **Variations on TSP Art** *Robert Bosch*
2. **Rare-event simulations in a chain of dynamical systems with small random perturbations** *Getachew Befekadu*
3. **Hungry for Equality: Fighting Food Deserts with Optimization** *Drew Horton*
4. **An alternative framework for the optimization of socially responsible portfolios applied to the Moroccan stock exchange** *Yahya Hanine*

11:45–12:00   **Break**

12:00–13:00   **Plenary talk** (Mohler 451, Mohler 453, MONRO)   *Chair: Martin Takac*

- **Incorporating second order ideas into first class machine learning methods** *Michael Mahoney*

13:00–14:00   **Lunch**

14:00–15:30   **Higher Order Optimization Methods** (Mohler 451, MONRO)   *Chair: Albert S. Berahas*

1. **A Trust Region-type Normal Map Semismooth Newton Method For Nonsmooth Large-scale Optimization** *Andre Milzarek*
2. **A New Multipoint Secant Method with a Dense Initial Matrix** *Jennifer Erway*
3. **A Subspace Acceleration Method for Minimization Involving a Group Sparsity-Inducing Regularizer** *Yutong Dai*

14:00–15:30   **Energy and Optimization** (Mohler 375, LEHIGH)   *Chair: Luis Zuluaga*

1. **On the Absence of Spurious Local Trajectories in Time-varying Nonconvex Optimization** *Cedric Josz*
2. **Revenue Adequate Prices for Chance-constrained Electricity Markets with Renewable Energy Sources** *Luis Zuluaga*
3. **Financially Adequate Environmental Pricing Designs** *Alberto Lamadrid*

14:00–15:30   **Optimization & Applications 1** (Mohler 453, IACOCCA)   *Chair: Gulcin Dinc Yalcin*

1. **The Quadratic Assignment Problem: Explaining Solvable Cases via Linear Reformulations** *Lucas Waddell*
2. **A Secretary Problem with Uncertain Offer Acceptance** *Sebastian Perez-Salazar*
3. **A new mathematical model and a greedy algorithm for the tourist trip design problem under new constraints: a real-world application** *Gulcin Dinc Yalcin*

15:30–15:45   **Closing Remarks**

ABSTRACTS

■ Nonlinear and Stochastic Optimization Algorithms

Room: **MONRO** (10:15 - 11:45) Chair: *Baoyu Zhou*

1. A Theoretical and Empirical Comparison of Gradient Approximations Methods in Derivative-Free Optimization

Liyuan Cao[1,], Albert S. Berahas[2], Katya Scheinberg[3]*

[1]Lehigh University; [*]`liyuancao7@gmail.com`;[2]University of Michigan;[3]Cornell University;

In our recently published paper, we analyze several methods for approximating gradients of noisy functions using only function values. These methods include finite differences, linear interpolation, Gaussian smoothing and smoothing on a sphere. The methods differ in the number of functions sampled, the choice of the sample points, and the way in which the gradient approximations are derived. For each method, we derive bounds on the number of samples and the sampling radius which guarantee favorable convergence properties for a line search or fixed step size descent method. We present numerical results evaluating the quality of the gradient approximations as well as their performance in conjunction with a line search derivative-free optimization algorithm.

2. Efficient algorithms for some variants of the extended trust-region subproblems

Maziar Salahi[1,], Saeid Ansary Karbasy[1], Abdelouahed Hamdi[2]*

[1]Department of Applied Mathematics, University of Guilan, Rasht, Iran; [*]`salahi.maziar@gmail.com`;[2]Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, Doha, Qatar;

In this talk, we discuss variants of the extended trust-region subproblems (eTRS), that are extensions of the well-known TRS. Even when we add one linear inequality constraint to TRS, the celebrated strong duality result may fail for eTRS. Thus several conditions are proposed in the literature implying strong duality and exact SDP relaxation. For example, exact SOCP/SDP relaxation is given for the case when extra linear constraints do not intersect inside the ball. We proposed the first efficient algorithm for eTRS when it has one linear constraint using the generalized eigenvalue problem which was also extended to two linear constraints. Taking advantage of the hard case and local nonglobal minimum of TRS , we were also able to develop an efficient algorithm for the case when linear constraints do not intersect side the ball, overperforming the SOCP/SDP relaxation.

3. Generalized Cyclic Stochastic Approximation and its Application in Multi-agent Systems

Jiahao Shi[1,], Karla Hernandez[2], Jiahao Shi[3], James C. Spall[3]*

[1]Johns Hopkins University Department of Applied Mathematics and Statistics; [*]`jiahaos@umich.edu`;[2]âĂŐWestern Digital;[3]Johns Hopkins University;

Stochastic approximation (SA) is a powerful class of iterative algorithms for minimizing a loss function, $L(\theta)$, when only noisy observations of $L(\theta)$ or its gradient are available. In this talk, we will present a generalized cyclic SA (GCSA) algorithm, a variant of SA procedures, where θ is divided into multiple subvectors that are updated one at a time. The subvector to update may be selected according to a random variable or according to a predetermined pattern. The convergence of GCSA, asymptotic normality of GCSA, and efficiency of GCSA relative to its non-cyclic counterpart are investigated. Finally, we apply the GCSA algorithm to a multi-agent stochastic optimization problem.

■ Constrained Optimization

Room: **LEHIGH** (10:15 - 11:45) Chair: *Negar Soheili Azad*

1. New notions of simultaneous diagonalizability of quadratic forms with applications to QCQPs

Alex Wang[1,]*

[1]Carnegie Mellon University; [*]`alw1@andrew.cmu.edu`;

A set of quadratic forms is simultaneously diagonalizable via congruence (SDC) if there exists a basis under which each of the quadratic forms is diagonal. This property appears naturally when analyzing quadratically constrained quadratic programs (QCQPs) and has important implications in this context. This talk will present a new weaker notion of simultaneous diagonalizability which extends the reach of the SDC property. Specifically, we say that a set of quadratic forms is d-restricted SDC (d-RSDC) if it is the restriction of an SDC set in up to d-many additional dimensions. Surprisingly, we will see that almost every pair of symmetric matrices is 1-RSDC.
We accompany our theoretical results with preliminary numerical experiments applying the RSDC property to QCQPs with a single quadratic constraint and additional linear constraints.
Based on joint work with Rujun Jiang.

2. On the rescaling and projection algorithm

Negar Soheili Azad[1,]*

[1]University of Illinois at Chicago; [*]`nazad@uic.edu`;

The projection and rescaling algorithm is a recently developed method that combines a basic procedure involving only low-cost operations with a periodic rescaling step. We propose a simple projection and rescaling algorithm that finds the "most interior" solutions to the pair of primal-dual polyhedral feasibility problems, an extension of the original projection and rescaling algorithm that finds a solution to one of these problems when it's feasible. We also present extensive numerical experiments on synthetic problem instances with varied levels of conditioning for both polyhedral and second-order cone feasibility problems. Our computational experiments provide promising evidence for the effectiveness of the projection and rescaling algorithm.

3. Second-Order Conic and Polyhedral Approximations of the Exponential Cone: Application to Mixed-Integer Exponential Conic Programs

Qing Ye[1,], Weijun Xie[2]*

[1]Ph.D. student, Virginia Tech; [*]`yqing1@vt.edu`;[2]Assistant Professor, Virginia Tech;

Exponents and logarithms exist in many important applications such as logistic regression, maximum likelihood, relative entropy and so on. Since the exponential cone can be viewed as the epigraph of perspective of the natural exponential function or the hypograph of perspective of the natural logarithm function, many mixed-integer convex programs involving exponential or logarithm functions can be recast as mixed-integer exponential conic programs (MIECPs). Recently, solver MOSEK is able to solve large-scale continuous exponential conic programs (ECPs). However, unlike mixed-integer linear programs (MILPs) and mixed-integer second-order conic programs (MISOCPs), MIECPs are far beyond development. To harvest the past efforts on MILPs and MISOCPs, this paper presents second-order

conic (SOC) and polyhedral approximation schemes for the exponential cone with application to MIECPs. To do so, we first extend and generalize existing SOC approximation approaches in the extended space, propose new scaling and shifting methods, prove approximation accuracies, and derive lower bounds of approximations. We then study the polyhedral outer approximation of the exponential cones in the original space using gradient inequalities, show its approximation accuracy, and derive a lower bound of the approximation. When implementing SOC approximations, we suggest learning the approximation pattern by testing smaller cases and then applying to the large-scale cases; and for the polyhedral approximation, we suggest using the cutting plane method when solving the continuous ECPs and branch and cut method for MIECPs. Our numerical study shows that the proposed scaling, shifting, and polyhedral outer approximation methods outperform solver MOSEK for both continuous ECPs and MIECPs and can achieve up to 20 times speed-ups compared to solver MOSEK when solving MIECPs.

■ Nonconvex optimization - Part (i)

Room: **IACOCCA** (10:15 - 11:45)   Chair: *Aida Khajavirad*

1. Proximity in Concave Integer Quadratic Programming

Alberto Del Pia[1,], Mingchen Ma[2]*

[1]University of Wisconsin at Madison;
[*]delpia@wisc.edu;[2]University of Wisconsin-Madison;

A classic result by Cook, Gerards, Schrijver, and Tardos provides an upper bound of $n\Delta$ on the proximity of optimal solutions of an Integer Linear Programming problem and its standard linear relaxation. In this bound, n is the number of variables and Δ denotes the maximum of the absolute values of the subdeterminants of the constraint matrix. Hochbaum and Shanthikumar, and Werman and Magagnosc showed that the same upper bound is valid if a more general convex function is minimized, instead of a linear function. No proximity result of this type is known when the objective function is nonconvex. In fact, if we minimize a concave quadratic, no upper bound can be given as a function of n and Δ. Our key observation is that, in this setting, proximity phenomena still occur, but only if we consider also approximate solutions instead of optimal solutions only. In our main result we provide upper bounds on the distance between approximate (resp., optimal) solutions to a Concave Integer Quadratic Programming problem and optimal (resp., approximate) solutions of its continuous relaxation. Our bounds are functions of n, Δ, and a parameter Î that controls the quality of the approximation. Furthermore, we discuss how far from optimal are our proximity bounds. This is joint work with Mingchen Ma.

2. Two-halfspace closure

Amitabh Basu[1,], Hongyi Jiang[1]*

[1]Johns Hopkins University; [*]abasu9@jhu.edu;

We define a new cutting plane closure for pure integer programs called the two-halfspace closure. It is a natural generalization of the well-known ChvÃątal-Gomory closure. We prove that the two-halfspace closure is polyhedral. We also study the corresponding 2-halfpsace rank of any valid inequality and show that it is at most the split rank of the inequality. Moreover, while the split rank can be strictly larger than the two-halfspace rank, the split rank is at most twice the two-halfspace rank. A key step of our analysis shows that the split closure of a rational polyhedron can be obtained by considering the split closures of all k-dimensional (rational) projections of the polyhedron, for any fixed $k \geq 2$. This result may be of independent interest.

3. Convexification of the Lennard-Jones Potential

Anatoliy Kuznetsov[1,], Nick Sahinidis[2]*

[1]School of Chemical & Biomolecular Engineering, Georgia Institute of Technology; [*]akuznetsov3@gatech.edu;[2]H. Milton Stewart School of Industrial & Systems Engineering, and School of Chemical & Biomolecular Engineering, Georgia Institute of Technology;

The Lennard-Jones potential is a semi-empirical relationship describing the energy between two particles due to van der Waals interactions. It has a wide range of applications in computational chemistry, including modeling nonbonded interactions in force fields used for protein folding calculations. We apply recently developed convexification techniques to derive the convex envelope of the Lennard-Jones potential for several geometries. Compared to convex relaxations generated using standard factorable relaxation techniques, the convex envelope introduces a significantly smaller relaxation gap.

■ Optimization and Healthcare

Room: **WHITEHOUSE** (10:15 - 11:45) Chair: *Karmel S. Shehadeh*

1. Distributionally Robust Optimization Approaches for a Mobile Facility Routing and Scheduling Problem

Karmel S. Shehadeh[1,]*

[1]Lehigh University; [*]kas720@lehigh.edu;

We study a mobile facility (MF) routing and scheduling problem in which probability distributions of the time-dependent demand for MF services is unknown. To address distributional ambiguity, we propose and analyze two distributionally robust MF routing and scheduling (DMFRS) models that seek to minimize the fixed cost of establishing the MF fleet and maximum expected transportation and unmet demand costs over all possible demand distributions residing within an ambiguity set. In the first model, we use a moment-based ambiguity set. In the second model, we use an ambiguity set that incorporates all distributions within a 1-Wasserstein distance from a reference distribution. To solve DMFRS models, we propose a decomposition-based algorithm and derive lower bound and twoâĂŞfamilies of symmetry-breaking inequalities to strengthen the master problem and speed up convergence. Finally, we present extensive computational experiments comparing the operational and computational performance of the proposed distributionally robust models and a stochastic programming model and drive insights into DMFRS.

2. Distributionally Robust Home Service Routing and Appointment Scheduling with Random Travel and Service Times

Man Yiu Tsang[1,], Karmel S. Shehadeh*

[1]Lehigh University; [*]mat420@lehigh.edu;

We study an integrated routing and appointment scheduling problem arising from home service practice. Specifically, given a set of customers within a service region that an operator needs to serve, we seek to find the operator's route and time schedule. The travel time between customers and the service time of each customer is random. The probability distributions of these random parameters are unknown. To address distributional ambiguity, we propose and analyze two distributionally robust home service routing and scheduling models that search for optimal routing and scheduling decisions to minimize the worst-case expectation of operational costs over all distributions residing within an ambiguity set. In the first model, we use a moment-based

ambiguity set. In the second model, we use a Wasserstein ambiguity set. We derive equivalent mixed-integer linear programming reformulations of both models. In an extensive numerical experiment, we investigate the proposed models computational and operational performance and derive insights into DHRAS.

3. A Voice base Nonsmooth Nonconvex optimization model for Parkinson Disease Severity Estimation

Habib Ghaffari Hadigheh[1,*]

[1]Ph.D. Student at McMaster University;
[*]ghaffh1@mcmaster.ca;

Parkinson's Disease (PD) is a common progressive neurodegenerative disorder characterized by several motor and non-motor features. It affects 1% of people older than 60, up to 4% of those over 80, and nearly one-half a million people over 50 in the U.S.. The healthcare-related cost due to PD is increasing as the population's longevity increases in developed countries, showing the importance of identifying patients with PD symptoms in the early stages. The most widely used tool in determining the symptoms and severity of PD is the Unified Parkinson's disease scale (UPDRS) which requires clinical expertise and experience. Many conventional ML models were introduced based on non-invasive measurements such as acoustic features using the patient's speech sample and UPDRS; However, most models identify the person as either a patient with PD or healthy controls (HS). There are challenges to generalizing the results of evaluations due to the limited number of data records that were collected from only one clinic in most cases. Another problem is that the UPDRS is scored based on highly qualitative data, which affects the reliability of PD's prediction in its early stages.

Having some experimental data may help one to find a reasonable regression function that relates the acoustic features of the patient's voice to the motor scores that an expert assigns to the case. Uncertainty theory initiated in 2007 is a mathematical framework devised to formalize human reasoning. Here, we use the facts of this theory to find an appropriate regression function to avoid overfitting commonly observed with interpolation methods. The problem of finding such a function is modeled as a non-smooth non-convex problem. This high-dimensional problem is unconstrained but computationally expensive. To avoid this cost, we devise a linearly constrained quadratic convex relaxation of the original model. We will discuss the computational complexity of the two models and the reasons we expect the relaxation to perform well despite the lower computational cost.

■ Methods for Large-Scale, Nonlinear and Stochastic Optimization

Room: **MONRO** (14:00 - 15:30)   Chair: *Jiahao Shi*

1. SQP for Nonlinear Equality Constrained Stochastic Optimization

Baoyu Zhou[1,*], *Albert S. Berahas*[2], *Frank E. Curtis*[1], *Daniel Robinson*[1]

[1]Lehigh University; [*]baz216@lehigh.edu;[2]University of Michigan;

Sequential quadratic optimization algorithms are proposed for solving smooth nonlinear optimization problems with equality constraints. The main focus is an algorithm proposed for the case when the constraint functions are deterministic, and constraint function and derivative values can be computed explicitly, but the objective function is stochastic. It is assumed in this setting that it is intractable to compute objective function and derivative values explicitly, although one can compute stochastic function and gradient estimates. We propose an adaptive stepsize selection scheme, since the objective is stochastic, for which it is assumed that line searches would be intractable. Under reasonable assumptions, convergence in expectation from remote starting points is proved for the proposed algorithm. The results of numerical experiments demonstrate the practical performance of our proposed techniques.

nite iteration and asymptotic sense. Numeral experiments also reveal its advantage in fast convergence in deep neural network problems with datasets like CIFAR-100.2

3. Complexity of Projected Newton Methods for Bound-constrained Optimization

Yue Xie[1,*]

[1]UW Madison; [*]xieyue1990@gmail.com;

Deriving complexity guarantees for nonconvex optimization problems are driven by long standing theoretical interests and by their relevance to machine learning and data science. This talk discusses complexity of algorithms for bound-constrained nonconvex optimization. We observe from the past work that pursuit of the state-of-art complexity guarantees can compromise the practicality of an algorithm. Therefore, we propose two practical projected Newton types of methods with complexity guarantees matching the best known. The first method is a scaled variant of Bertsekas' two-metric projection method, which can be shown to output an epsilon approximate first-order point in $O(\epsilon^{-2})$ iterations. The second is a projected Newton-Conjugate Gradient method, which locates an epsilon approximate second-order point with high probability in $O(\epsilon^{-3/2})$ iterations. Preliminary numerical experiments on Nonnegative Matrix Factorization indicate practicality of the latter algorithm.

■ Quantum Optimization I

Room: **LEHIGH** (14:00 - 15:30)   Chair: *Brandon Augustino*

1. An Inexact-Feasible Interior Point Method for Linear Optimization with High Adaptability to Quantum Computers

Mohammadhossein Mohammadisiahroudi[1,],
Ramin Fakhimi[2], TamÃąs Terlaky[2]*

[1]PhD student, Industrial & System Engineering Department, Lehigh University; [*]mom219@lehigh.edu;[2]Industrial & System Engineering Department, Lehigh University;

Quantum computing can speed up Interior Point Methods (IPMs) by using Quantum Linear System Algorithms (QLSAs) to solve the Newton systems. Since QLSAs inherently produce inexact solutions, an Inexact-Feasible IPM (IF-IPM) is proposed for linear optimization problems using a novel system that produces inexact but feasible steps. We also discuss how QLSAs can be used efficiently in an Iterative Refinement scheme to find an exact solution without excessive time of QLSAs. The results show that IF-QIPM has better time complexity than other quantum and classical IPMs w.r.t the dimension. The IF-IPM is implemented with both classical and quantum solvers to investigate its efficiency numerically.

2. Quantum Interior Point Methods for Semidefinite Optimization

Brandon Augustino[1,]*

[1]Department of Industrial and Systems Engineering, Lehigh University; [*]bra216@lehigh.edu;

We present two quantum interior point methods for semidefinite optimization problems, building on recent advances in quantum linear system solvers. The first scheme, more similar to a classical solver, computes an inexact search direction and is not guaranteed to stay feasible; the second scheme uses a nullspace representation of the Newton linear system to ensure feasibility even with inexact search directions. This second scheme would be impractical in the classical world, but it is well-suited for a hybrid quantum-classical setting. We show that both schemes converge to an optimal solution of the semidefinite optimization problem under standard assumptions. By comparing the theoretical performance of classical and quantum interior point methods with respect to various input parameters, we show that our second scheme obtains a speedup over classical algorithms in terms of the size of the problem, but has worse dependence on other numerical parameters.

3. Characterization and Mitigation of Errors in Quantum Computing via Consistent Bayesian

Muqing Zheng[1,], Ang Li[2], TamÃąs Terlaky[1], Xiu Yang[1]*

[1]Department of Industrial and Systems Engineering, Lehigh University; [*]muz219@lehigh.edu;[2]Advanced Computing, Mathematics and Data Division, Pacific Northwest National Laboratory;

Various noise models have been developed in quantum computing study to describe the propagation and effect of the noise which is caused by imperfect implementation of hardware. Identifying parameters such as gate and readout error rates are critical to these models. We use a Bayesian inference approach to identity posterior distributions of these parameters, such that they can be characterized more elaborately. By characterising the device errors in this way, we can further improve the accuracy of quantum error mitigation. Experiments conducted on IBM's quantum computing devices suggest that our approach provides better error mitigation performance than existing techniques used by the vendor. Also, our approach outperforms the standard Bayesian inference method in such experiments.

■ Nonconvex optimization, Part (ii)

Room: **IACOCCA** (14:00 - 15:30)   Chair: *Aida Khajavirad*

1. Sum-of-squares lower bounds for random combinatorial problems

Dmitriy Kunisky[1,], Afonso S. Bandeira[2]*

[1]Courant Institute of Mathematical Sciences; [*]dk3105@nyu.edu;[2]ETH Zurich;

I will present results showing that sum-of-squares semidefinite programming relaxations cannot certify strong bounds on random combinatorial problems in high dimension, focusing in particular on the Sherrington-Kirkpatrick model of statistical physics, the problem of optimizing a Gaussian quadratic form over the hypercube. I will describe how the proofs of these results are driven by understanding the geometric structure of high-dimensional sum-of-squares feasible points via the Gram matrix factorizations of associated pseudomoment matrices. We will see how these considerations suggest higher-degree analogs of the notions of "vector cuts" and "vector colorings" often offered as interpretations of the simpler Goemans-Williamson and Lovasz theta function relaxations of graph problems. The talk is based on joint work with Afonso Bandeira.

2. Sparse regression: decompositions, convexifications and algorithms

Andres Gomez[1,], Simge Küçükyavuz[2], Linchuan Wei[2]*

[1]University of Southern California; [*]gomezand@usc.edu;[2]Northwestern;

We study sparse regression problems, this is, regression problems where a small subset of the possible predictor variables should be used. Sparse regression problem can be easily modeled with indicator variables controlling which variables can be non-zero, although natural big-M formulations often result in poor relaxations. We propose novel convexifications for the epigraphs of a special class of quadratic functions with indicators with arbitrary constraints on the indicator variables.

3. Rank Pump: A Feasibility Heuristic For Polynomial Optimization

Chen Chen[1,], Daniel Bienstock[2]*

[1]Ohio State University; [*]chen.8018@osu.edu;[2]Columbia University;

The feasibility pump is a well-known primal heuristic for integer programming that involves two alternating sequences of projections. The original pump was designed for binary problems, and found such projections using linear programming and simple rounding. Unfortunately, the elegance of the pump may be lost in other settings. For instance, a natural extension of the pump to nonconvex MINLP involves NP-hard projection problems. We present our adaptation of the feasibility pump to polynomial optimization, called the rank pump. The rank pump has polynomial-time iterations, as all its projection problems can be solved in polynomial time.

■ OR in Healthcare

Room: **WHITEHOUSE** (14:00 - 15:30)   Chair: *Tugce Isik*

1. Solution Methods for Integrated Surgery Scheduling and Inventory Problem

Amogh Bhosekar[1,], Tugce Isik[1], Sandra Eksioglu[2]*

[1]Clemson University; *abhosek@g.clemson.edu;[2]University of Arkansas;

Considering the availability and cost of surgical instruments while creating operating rooms (ORs) schedule provides an opportunity to reduce the cost of healthcare. We propose a mixed integer programming (MIP) model for the integrated problem to determine the schedule of surgeries and assignments of instruments to surgeries over a week. The objective of the model is to minimize the cost of opening the ORs, overtime, idle-time, and the cost of using/borrowing instruments to satisfy the demand. To generate solutions in reasonable time, a Lagrangean decomposition-based heuristic is proposed in which the integrated problem is separated into a surgery scheduling problem and an instrument assignment problem. The surgery scheduling model has a special structure that has no resource sharing among surgeries assigned different (day,OR) tuples in the planning horizon. This renders the sequencing decisions for each (day,OR) assignment independent once a patient to (day,OR) schedule is given. A partitioning procedure based on Logic Based Benders Decomposition is used to solve the scheduling problem and results in MIP sequence optimization problem one for each (day,OR). The results of our experiments indicate that integrating decisions lowers the total system costs.

2. Optimal Nurse Allocation for the Surgical System: A Tandem Network with Flexible Servers

Tong Zhang[1,], Tugce Isik[1]*

[1]Clemson University; *tzhang5@g.clemson.edu;

Nursing shortage is a major challenge faced by the US healthcare system. Cross-training of highly skilled nursing staff has the potential to increase the nursing capacity at low cost. However, strategies to effectively allocate cross-trained nurses to tasks should be explored. In this paper, a tandem queueing system with flexible servers is considered to model a surgical system. When a patient needs surgery, they must go through three stages: pre-operative care, surgery, and post-operative care. Patients may move through a pre-operative area, the operating room (OR), and a post-anesthesia care unit (PACU) during these stages, or they may recover in the OR due to lack of PACU beds or staff. We consider different sets of operational rules regarding how patients flow through these locations. Considering holding costs and unit rewards for each patient departure, we model a Markov decision process to optimize the nurse allocation decisions under each set of rules. We investigate policies that maximize long-run average rewards and examine conditions under which nurse cross-training is beneficial.

3. Dynamic Tuberculosis Screening for Healthcare Employees

Mahsa Kiani[1,], Tugce Isik[2], Burak Eksioglu[3]*

[1]Data Scientist at Bayer Crop Science; *mkiani@g.clemson.edu;[2]Assistant Professor at Clemson University;[3]Professor at University of Arkansas;

Regular tuberculosis (TB) screening is required for healthcare employees since they can come into contact with infected patients. TB is a serious, contagious, and potentially deadly disease. Early detection of the disease, even when it is in latent form, prevents the spread of the disease and helps with treatment. Currently, there are two types of TB diagnostic tests on the market: skin test and blood test. The cost of the blood test is much higher than the skin test. However, the possibility of getting a false positive or false negative result in a skin test is higher especially for persons with specific characteristics, which can increase costs. In this study, we categorize healthcare employees into multiple risk groups based on the department they work in, the specific job they do, and their birth country. We create a Markov decision process (MDP) model to decide which TB test should be utilized for each employee group to minimize the total costs related to testing, undetected infections, employees' time lost. Due to the size of the problem, we use approximate dynamic programming (ADP) to obtain a near-optimal solution. By analyzing this solution to the ADP, we specify not only the type of the tests that should be used but also the frequency with which each test should be administered. Based on this analysis, we propose a simple policy that can be used by healthcare facilities since such facilities may not have the expertise or the resources to develop and solve sophisticated optimization models.

■ Advances in Large-Scale Nonlinear Optimization

Room: **MONRO** (16:45 - 18:15)   Chair: *Raghu Bollapragada*

1. Average Curvature FISTA for Nonconvex Smooth Composite Optimization Problems

Jiaming Liang[1,], Renato D. C. Monteiro[1]*

[1]Georgia Institute of Technology; *jliang76@gatech.edu;

A previous authors' paper introduces an accelerated composite gradient (ACG) variant, namely AC-ACG, for solving nonconvex smooth composite optimization (N-SCO) problems. In contrast to other ACG variants, AC-ACG estimates the local upper curvature of the N-SCO problem by using the average of the observed upper-Lipschitz curvatures obtained during the previous iterations, and uses this estimation and two composite resolvent evaluations to compute the next iterate. This paper presents an alternative FISTA-type ACG variant, namely AC-FISTA, which has the following additional features: i) it performs an average of one composite resolvent evaluation per iteration; and ii) it estimates the local upper curvature by using the average of the previously observed upper (instead of upper-Lipschitz) curvatures. These two properties acting together yield a practical AC-FISTA variant which substantially outperforms earlier ACG variants, including the AC-ACG variants discussed in the aforementioned authors' paper.

2. Implicit Regularization of Sub-Gradient Method in Robust Matrix Recovery: Don't be Afraid of Outliers

Jianhao Ma[1,]*

[1]University of Michigan; *jianhao@umich.edu;

It is well-known that simple short-sighted algorithms, such as gradient descent, generalize well in the over-parameterized learning tasks, due to their implicit regularization. However, it is unknown whether the implicit regularization of these algorithms can be extended to robust learning tasks, where a subset of samples may be grossly corrupted with noise. In this work, we provide a positive answer to this question in the context of robust matrix recovery problem. In particular, we consider the problem of recovering a low-rank matrix from a number of linear measurements, where a subset of measurements are corrupted with large noise. We show that a simple sub-gradient method converges to the true low-rank solution efficiently, when it is applied to the over-parameterized l1-loss function without any explicit regularization or rank constraint. Moreover, by building upon a new notion of restricted isometry property, called sign-RIP, we prove the robustness

of the sub-gradient method against outliers in the over-parameterized regime. In particular, we show that, with Gaussian measurements, the sub-gradient method is guaranteed to converge to the true low-rank solution, even if an arbitrary fraction of the measurements are grossly corrupted with noise.

3. Inexact Proximal Gradient Methods

Daniel Robinson[1,*]

[1]Lehigh University; [*]dpr219@lehigh.edu;

I will discuss our recent work on developing INEXACT proximal-gradient methods. In particular, we design termination conditions that indicate how accurately each proximal-gradient subproblem must be solved. Such termination conditions are crucial when the proximal-gradient subproblem does not have a closed form solution, as is the case for important regularizers such as the overlapping group regularizer and the latent group regularizer. Unlike previous work in this area, which has focused on developing complexity results by choosing the termination conditions IN ADVANCE, we base our termination conditions on information local to the current iterate. As such, our conditions are more practical and lead to different types of complexity results, which will be discussed.

■ Quantum Optimization II

Room: **LEHIGH** (16:45 - 18:15)   Chair: *Brandon Augustino*

1. Quantum-inspired formulations for the max k-cut problem

Ramin Fakhimi[1,*]*, Hamidreza Validi*[2]*, Illya V. Hicks*[3]*, Tamas Terlaky*[4]*, Luis F. Zuluaga*[5]

[1]PhD candidate; [*]fakhimi@lehigh.edu;[2]Postdoc at Rice University;[3]Professor of Computational and Applied Mathematics, Rice University;[4]Endowed Chair Professor Industrial and Systems Engineering, Lehigh University;[5]Associate Professor, Industrial and Systems Engineering Department, Lehigh University;

Solving combinatorial optimization problems on quantum computers has attracted many researchers since the emergence of quantum computing. The max k-cut problem is a challenging combinatorial optimization problem with multiple well-known optimization formulations. However, its mixed-integer linear optimization (MILO) formulations and mixed-integer semidefinite optimization formulation are all time-consuming to be solved. Motivated by recent progress in classic and quantum solvers, we study a binary quadratic optimization (BQO) formulation and two quadratic unconstrained binary optimization formulations. First, we compare the BQO formulation with the MILO formulations. Further, we propose an algorithm that converts any feasible fractional solution of the BQO formulation to a feasible binary solution whose objective value is at least as good as that of the fractional solution. Finally, we find tight penalty coefficients for the proposed quadratic unconstrained binary optimization formulations.

2. Improving QAOA with Warm-Start Initializations and Custom Mixers

Reuben Tate[1,*]*, Bryan Gard*[2]*, Swati Gupta*[1]*, Greg Mohler*[2]

[1]Georgia Institute of Technology; [*]rtate30@gatech.edu;[2]Georgia Tech Research Institute;

In this talk, we consider bridging classical optimization techniques with quantum algorithms. We propose using classical "warm-starts" (obtained via solutions to low-rank semidefinite programming relaxations) in order to initialize the starting state of the Quantum Approximate Optimization Algorithm (QAOA) in the context of the MAX-CUT problem. In addition to changing the initial state, we also consider changing the mixing Hamiltonian in a way that allows us analyze QAOA through the lens of quantum adiabatic algorithms. Our experiments suggest that this modified version of QAOA is robust against quantum noise and is able to yield higher quality cuts (compared to standard QAOA or the classical Goemans-Williamson algorithm) even with low-circuit depth and limited training time for most instances. We provide simulation and theoretical results on the performance of the proposed framework. This is based on joint work with Bryan Gard, Swati Gupta and Greg Mohler.

3. Qubo Reformulations of Combinatorial Optimization Problems

Rodolfo Alexander Quintero Ospina[1,*]*, David Bernal*[2]*, TamÃąs Terlaky*[3]*, Luis F. Zuluaga*[3]

[1]Department of Industrial and Systems Engineer, Lehigh University; [*]roq219@lehigh.edu;[2]Carnegie Mellon University;[3]Lehigh University;

Adiabatic quantum computers have shown to outperform classical computers in solving some particular instances of NP-hard problems, like the Graph partitioning problem. To do this, a Quadratic Unconstrained Binary Optimization (QUBO) formulation is needed. Given that many combinatorial problems, in particular, NP-hard problems, can be formulated as QUBO instances, the interest in getting implementable QUBO formulations of such problems has grown in recent years. In this presentation, we will focus on the QUBO formulations of the independent set and the maximum k-colorable subgraph problems, and some possible limitations to implement them in quantum computers.

■ Advances in Nonconvex Optimization

Room: **IACOCCA** (16:45 - 18:15)   Chair: *Meisam Razaviyayn*

1. A Riemannian Block Coordinate Descent Method for Computing the Projection Robust Wasserstein Distance

Minhui Huang[1,*]*, Shiqian Ma*[1]*, Lifeng Lai*[1]

[1]UC Davis; [*]mhhuang@ucdavis.edu;

The Wasserstein distance has become increasingly important in machine learning and deep learning. Despite its popularity, the Wasserstein distance is hard to approximate because of the curse of dimensionality. A recently proposed approach to alleviate the curse of dimensionality is to project the sampled data from the high dimensional probability distribution onto a lower-dimensional subspace, and then compute the Wasserstein distance between the projected data. However, this approach requires to solve a max-min problem over the Stiefel manifold, which is very challenging in practice. The only existing work that solves this problem directly is the RGAS (Riemannian

Gradient Ascent with Sinkhorn Iteration) algorithm, which requires to solve an entropy-regularized optimal transport problem in each iteration, and thus can be costly for large-scale problems. In this paper, we propose a Riemannian block coordinate descent (RBCD) method to solve this problem, which is based on a novel reformulation of the regularized max-min problem over the Stiefel manifold. We show that the complexity of arithmetic operations for RBCD to obtain an ϵ-stationary point is $O(\epsilon^{-3})$. This significantly improves the corresponding complexity of RGAS, which is $O(\epsilon^{-12})$. Moreover, our RBCD has very low per-iteration complexity, and hence is suitable for large-scale problems. Numerical results on both synthetic and real datasets demonstrate that our method is more efficient than existing methods, especially when the number of sampled data is very large.

2. A Smoothing Scheme for Nonconvex-Concave Min-Max Problems

Weiwei Kong[1,], Renato D.C. Monteiro[2]*

[1]Oak Ridge National Labs; [*]wwkong92@gmail.com;[2]Georgia Institute of Technology;

This talk presents a smoothing scheme for obtaining an approximate stationary point of a composite nonconvex-concave min-max problem by applying a well-known algorithm to the composite smooth approximation of the original problem. More specifically, approximate stationary points of the original problem are obtained by applying (to its composite smooth approximation) an accelerated inexact proximal point method presented in a previous paper by the authors. Iteration complexity bounds for the smoothing scheme are also given for two notions of approximate stationarity. Finally, numerical results are given to demonstrate the efficiency of the scheme.

3. Smooth nonconvex-nonconcave min-max optimization problems with small inner maximization constraint set

Meisam Razaviyayn[1,], Dmitrii Ostrovskii[2], Babak Barazandeh[2], Tianjian Huang[2]*

[1]University of Southern California; [*]razaviya@usc.edu;[2]USC;

Recent applications that arise in machine learning have surged significant interest in solving min-max optimization problems. This problem has been extensively studied in the convex-concave regime for which a global optimal solution solution can be computed efficiently. However, in the nonconvex-nonconcave (smooth) regime, most problems cannot be solved to any reasonable notion of stationarity. In this work, we identify a class of smooth nonconvex-nonconcave min-max problems that can be solved efficiently up to first-order stationarity of its Moreau envelope. In particular, we propose an efficient algorithm for finding (first-order) stationary solutions to nonconvex-nonconcave min-max problems when the radius of the constraint set for the inner maximization problem is comparable to the desired accuracy in the outer problem. Our results are the first of its kind that find stationary solutions to nonconvex-nonconcave min-max problems without assuming any restriction on the objective function (other than standard smoothness assumptions). We also discuss the validity of our assumptions and evaluate the performance of our algorithm on the problem of training robust neural networks against adversarial attacks.

■ Optimization with Applications in Healthcare

Room: **WHITEHOUSE** (16:45 - 18:15)   Chair: *Secil Sozuer*

1. Data-Driven Distributionally Robust Surgery Planning in Flexible Operating Rooms Over a Wasserstein Ambiguity

Karmel S. Shehadeh[1,]*

[1]Lehigh University; [*]kas720@lehigh.edu;

We study elective surgery planning in flexible operating rooms where emergency patients are accommodated in the existing elective surgery schedule. Probability distributions of surgeries durations are unknown, and only a small set of historical realizations is available. To address distributional ambiguity, we first construct an ambiguity set that encompasses all possible distributions of surgery duration within a Wasserstein distance from the empirical distribution. We then define a data-driven distributionally robust surgery assignment (DSA) problem, which seeks to determine optimal elective surgery assigning decisions to available surgical blocks in multiple ORs to minimize the sum of patient-related costs and the expectation of OR overtime and idle time costs over all distributions residing in the ambiguity set. Using DSA structural properties, we derive an equivalent mixed-integer linear programming (MILP) reformulation of the min-max DSA model, that can be solved efficiently using off-the-shelf optimization software. Using real-world surgery data, we conduct extensive numerical experiments comparing the operational and computational performance of our approach with two state-of-the-art approaches.

2. Simulation-Based Optimization of Dynamic Appointment Scheduling Problem with Patient Unpunctuality and Provider Lateness

Secil Sozuer[1,], Robert H. Storer[1], Miao Bai[2]*

[1]Lehigh University; [*]ses515@lehigh.edu;[2]University of Connecticut;

Healthcare providers are under growing pressure to improve efficiency due to an aging population and increasing expenditures. This research is designed to address a particular healthcare scheduling problem, dynamic and stochastic appointment scheduling with patient unpunctuality and provider lateness. We consider that the stochasticity is coming from uncertain patient requests, uncertain service duration, patient unpunctuality and provider lateness. The aim is to find the optimal schedule start time for the patients in order to minimize the expected cost incurred from patient waiting time, server idle time, and server overtime. By conducting perturbation analysis for the gradient estimation, a Sample Average Approximation (SAA), a Robust Stochastic Approximation (Robust SA) and adaptive Stochastic Approximation (ad-SA) algorithms are used. The structural properties of the sample path cost function and expected cost function are studied. Numerical experiments show the computational advantages of using perturbation based gradient information over CPLEX and interior point methods for our problem.

3. Supply Side Flexibility in Revenue Management: An Application in Medical Wire and Device Manufacturing

Cigdem Gurgur[1,]*

[1]Purdue University Fort Wayne; [*]gurgurc@pfw.edu;

Revenue management is a concept aimed to maximize capacity utilization and through that maximize revenues. While the main developments in revenue management have taken place in the fields of service industries, relatively little research has been done for the manufacturing sector. We consider revenue management for a medical wire and device manufacturing company with make-to-order mode of operation, complex alloy market, broad-product mix and limited inventory capacity. Orders with different profit margins arrive stochastically and the company has to decide which orders to accept and which orders to reject. In the "Make-to-Order," the allocation of the finite capacity to certain orders is complicated because the processes that are used to make a product are consistent, but vary in the manner in which product goes through the process. Furthermore, the number of iterative loops in which a particular order may go through the same process varies. We model the problem with a Markov decision process and propose a value iteration heuristic. In numerical tests we show the potential benefit of using revenue management instead of a first-come-first-serve policy and assess the performance of the heuristic procedure.

■ Advances in Nonlinear ADMM and Related Computational Methods

Room: **MONRO** (10:15 - 11:45)   Chair: *Dimitri Papadimitriou*

1. Escaping from Zero Gradient: Revisiting Action-Constrained Reinforcement Learning via Frank-Wolfe Policy Optimization

Ping-Chun Hsieh[1,], Jyun-Li Lin[1], Wei Hung[1], Shang-Hsuan Yang[1], Xi Liu[2]*

[1]National Chiao Tung University (NCTU), Hsinchu, Taiwan; [*]pinghsieh@nctu.edu.tw;[2]Facebook AI, Menlo Park, CA, USA;

This talk revisits action-constrained reinforcement learning (RL), a widely-used learning setting in various real-world applications, such as scheduling in networked systems with resource constraints and control of a robot with kinematic constraints. While the existing projection-based approaches ensure zero constraint violation, they could suffer from the zero-gradient problem due to the tight coupling of the policy gradient and the projection, which results in sample-inefficient training and slow convergence. To tackle this issue, we propose a learning algorithm that decouples the action constraints from the policy parameter update by leveraging state-wise Frank-Wolfe and a regression-based policy update scheme. Moreover, we show that the proposed algorithm enjoys convergence and policy improvement properties in the tabular case as well as generalizes the popular DDPG algorithm for action-constrained RL in the general case. Through experiments, we demonstrate that the proposed algorithm significantly outperforms the benchmark methods on a variety of control tasks.

2. Relaxed alternating minimization algorithm for convex optimization problems in image denoising

Yuchao Tang[1,], Shuangshuang Wu[2], Tieyong Zeng[3]*

[1]NCU; [*]yctang@ncu.edu.cn;[2]Nanchang UniversitNanchang University;[3]The Chinese University of Hong Kong;

In this paper, we propose a relaxed alternating minimization algorithm for solving two-block separable convex minimization problems with linear equality constraints, where one block in the objective functions is strongly convex. This algorithm is derived from the relaxed proximal gradient algorithm. We prove that the proposed algorithm converges to an optimal primal-dual solution of the original problem. Furthermore, we study asymptotic $\circ\left(\frac{1}{k}\right)$ convergence rate of the primal feasibility residual, where k is the number of iterations. As applications, we apply the proposed algorithm to solve several composite convex minimization problems arising in image denoising and evaluate the numerical performance of the proposed algorithm on a novel image denoising model. Numerical results on both artificial and real noisy images demonstrate the efficiency and effectiveness of the proposed algorithm.

3. Power of Alternating Direction Method of Multipliers (ADMM) in Deep Learning

Junxiang Wang[1,], Liang Zhao[1], Yue Chen[2], Zheng Chai[2]*

[1]Emory University; [*]jwan936@emory.edu;[2]George Mason University;

The Alternating Direction Method of Multipliers (ADMM) has been demonstrated powerful performance in many conventional machine learning applications and is recently considered to be a potential alternative to Stochastic Gradient Descent (SGD) as a deep learning optimizer. However, several challenges remain in this emerging domain, including slow convergence towards solutions, the expensive computational cost, and the lack of theoretical convergence guarantees. In this talk, I introduce a novel optimization framework for deep learning via ADMM (dlADMM) to address them simultaneously. Specifically, the parameters in each layer are updated backward and then forward so that the parameter information in each layer is exchanged efficiently; the computational cost is reduced from cubic to quadratic via a dedicated algorithm design for subproblems that enhances them utilizing iterative quadratic approximations and backtracking. Moreover, we provide proof of convergence to a critical point for an ADMM-based method (dlADMM) in a neural network problem under mild conditions. In order to achieve model parallelism, we extend our dlADMM framework to parallel deep learning ADMM framework (pdADMM): parameters in each layer of neural networks can be updated independently in parallel. Extensive experiments on multiple benchmark datasets demonstrate that our proposed dlADMM and pdADMM outperform most of the comparison methods, and the pdADMM can lead to more than 10 times speedup for training large-scale deep neural networks.

■ Algorithms for Derivative-Free Optimization

Room: **LEHIGH** (10:15 - 11:45)   Chair: *Baoyu Zhou*

1. Zeroth-Order Riemannian Optimization

Jiaxiang Li[1,]*

[1]UCDavis; [*]jxjli@ucdavis.edu;

We consider stochastic zeroth-order optimization over Riemannian submanifolds embedded in Euclidean space, where the task is to solve Riemannian optimization problem with only noisy objective function evaluations. Towards this, our main contribution is to propose estimators of the Riemannian gradient and Hessian from noisy objective function evaluations, based on a Riemannian version of the Gaussian smoothing technique. The proposed estimators overcome the difficulty of the non-linearity of the manifold constraint and the issues that arise in using Euclidean Gaussian smoothing techniques when the function is defined only over the manifold. We use the proposed estimators to solve Riemannian optimization problems in the following settings for the objective function: (i) stochastic and gradient-Lipschitz (in both nonconvex and geodesic convex settings), (ii) sum of gradient-Lipschitz and non-smooth functions, and (iii) Hessian-Lipschitz. For these settings, we analyze the oracle complexity of our algorithms to obtain appropriately defined notions of e-stationary point or e-approximate local minimizer. Notably, our complexities are independent of the dimension of the ambient Euclidean space and depend only on the intrinsic dimension of the manifold under consideration. We demonstrate the applicability of our algorithms by simulation results and real-world applications on black-box stiffness control for robotics and black-box attacks to neural networks.

2. New Hybrid Algorithms for Global and Local Derivative-Free Optimization

Ahmad Almomani[1,], Jonathan McCart[2]*

[1]SUNY Geneseo mathematics department; [*]almomani@geneseo.edu;[2]SUNY Geneseo;

Derivative-Free methods are highly demanded in the last three decades for solving optimization problems. In many practical applications, the derivatives are not available or hard to compute due to a "black-box"

or simulation-based formulation, or you cannot trust the approximation. Derivative-Free Optimization (DFO) methods are applicable for these kinds of problems compared to methods that employ derivatives. The need for DFO arises extensively across all engineering and science disciplines. Particle Swarm Optimization (PSO) is considered one of the best (DFO) global solvers and has been efficient and robust compared with the other DFO algorithms. Hybrid optimization recently grew due to the high demand for more efficient algorithms than the original optimization method used. In this talk, I will introduce new hybrid algorithms inspired by PSO that show high efficiency over other hybrid algorithms.

3. Manifold Sampling for Optimizing Nonsmooth Nonconvex Compositions

Baoyu Zhou[1,]*

[1]Lehigh University; [*]baz216@lehigh.edu;

We propose a manifold sampling algorithm for minimizing a nonsmooth composition function $f = h \circ F$, assuming h is piecewise smooth with known subdifferential, and F is smooth but with unavailable Jacobian. While exact first-order information of f may be unavailable, approximate first-order information can be obtained by using the gradients of models approximating the components of F combined with known information about h. By collecting such manifold information in a trust region around an iterate, our algorithm can evaluate a function decrease condition and measure Clarke stationarity at the same time. While each manifold sampling iteration contains many subproblems, we show each is tractable. We prove that all cluster points of the sequence of iterates generated by the algorithm are Clarke stationary. Numerical results demonstrate manifold sampling building models of F and using knowledge about h is competitive with other algorithms when they utilize first-order information about f.

■ AIMMS-MOPTA Optimization Modeling Competition

Room: **IACOCCA** (10:15 - 11:45) Chair: *Karmel Shehadeh*

■ Advances in Stochastic Optimization

Room: **MONRO** (14:00 - 15:30) Chair: *Dimitri Papadimitriou*

1. A Stochastic Decoupling Method for Minimizing the Sum of Smooth and Non-Smooth Functions

Konstantin Mishchenko[1,], Peter Richtarik[1]*

[1]KAUST; [*]konsta.mish@gmail.com;

We consider the problem of minimizing the sum of three convex functions: i) a smooth function f in the form of an expectation or a finite average, ii) a non-smooth function g in the form of a finite average of proximable functions g_j, and iii) a proximable regularizer R. We design a variance-reduced method which is able to progressively learn the proximal operator of g via the computation of the proximal operator of a single randomly selected function g_j in each iteration only. Our method can provably and efficiently accommodate many strategies for the estimation of the gradient of f, including via standard and variance-reduced stochastic estimation, effectively decoupling the smooth part of the problem from the non-smooth part. We prove a number of iteration complexity results, including a general $O(1/t)$ rate, $O(1/t^2)$ rate in the case of strongly convex smooth f, and several linear rates in special cases, including accelerated linear rate. For example, our method achieves a linear rate for the problem of minimizing a strongly convex function f subject to linear constraints under no assumption on the constraints beyond consistency. When combined with SGD or SAGA estimators for the gradient of f, this leads to a very efficient method for empirical risk minimization. Our method generalizes several existing algorithms, including forward-backward splitting, Douglas-Rachford splitting, proximal SGD, proximal SAGA, SDCA, randomized Kaczmarz and Point-SAGA. However, our method leads to many new specific methods in special cases; for instance, we obtain the first randomized variant of the Dykstra's method for projection onto the intersection of closed convex sets.

2. An Online Algorithm for Maximum-Likelihood Quantum State Tomography

Yen-Huan Li[1,], Chien-Ming Lin[1], Yu-Ming Hsu[1]*

[1]National Taiwan University; [*]yenhuan.li@csie.ntu.edu.tw;

Quantum state tomography, the task of estimating an unknown quantum state given measurement outcomes, is essential to building reliable quantum computing devices. We propose, to the best of our knowledge, the first online algorithm for maximum-likelihood quantum state tomography. Suppose the quantum state to be estimated corresponds to a D-by-D density matrix. The per-iteration computational complexity of the algorithm is $O(D^3)$, independent of the data size. The expected numerical error of the algorithm is $O(\sqrt{(1/T)D \log D})$, where T denotes the number of iterations. The algorithm is a quantum extension of Soft-Bayes, a recent algorithm for online portfolio selection (Orseau et al. Soft-Bayes: Prod for mixtures of experts with log-loss. Int. Conf. Algorithmic Learning Theory. 2017) and a provably converging stochastic version of an expectation maximization-type method called RρR that does not always converge (Rehacek et al. Iterative algorithm for reconstruction of entangled states. Phys. Rev. A. 2001.).

3. Perturbed Fenchel duality and first-order methods

Javier Pena[1,]*

[1]Carnegie Mellon University; [*]jfp@andrew.cmu.edu;

We show that the iterates generated by a generic first-order meta-algorithm satisfy a canonical perturbed Fenchel duality inequality. The latter in turn readily yields a unified derivation of the best known convergence rates for various popular first-order algorithms including the conditional gradient method as well as the main kinds of Bregman proximal methods: subgradient, gradient, fast gradient, and universal gradient methods.
This is joint work with David H. Gutman at Texas Tech University.

■ New Deterministic and Stochastic Methods for Derivative-Free Optimization I

Room: **LEHIGH** (14:00 - 15:30)   Chair: *Tommaso Giovannelli*

1. Full-Low Evaluation Methods for Derivative-Free Optimization

Oumaima Sohab[1,*], Luis Nunes Vicente[1], Albert S. Berahas[2]

[1]Lehigh University; [*]ous219@lehigh.edu;[2]University of Michigan;

We propose a new class of directional methods for Derivative-Free Optimization that considers two types of iterations. The first type is expensive in function evaluations, but exhibits good performance in the smooth, non-noisy case. The instance considered is BFGS computed over gradients approximated by finite differences. The second type is cheap in function evaluations, more appropriate under the presence of noise or non-smoothness. The instance considered is probabilistic direct search with 1 or 2 random directions.
The resulting Full-Low Evaluation method is globally convergent even in the non-smooth case, and yields the appropriate rates in the smooth case. Results show that is efficient and robust across problems with different levels of smoothness and noise.

2. Randomized DFO Methods for Fitting Numerical Physics Models

Matt Menickelly[1,*], Raghu Bollapragada[2], Witold Nazarewicz[3], Jared O'Neal[1], Paul-Gerhard Reinhard[2], Stefan Wild[1]

[1]Argonne National Laboratory; [*]mmenickelly@anl.gov;[2]University of Texas at Austin;[3]Michigan State University;[2]Friedrich-Alexander-UniversitÃd't Erlangen-Nürnberg;

We address the calibration of a computationally expensive nuclear physics model for which derivative information with respect to the fit parameters is not readily available. Of particular interest is the performance of optimization-based training algorithms when dozens, rather than millions or more, of training data are available and when the expense of the model places limitations on the number of concurrent model evaluations that can be performed. Our initial experiments inspired the development of a new randomized variant of the derivative-free optimization solver, POUNDerS, for which we present some preliminary numerical results.

3. A derivative-free method for structured optimization problems

Andrea Cristofari[1,*], Francesco Rinaldi

[1]University of Padua; [*]andrea.cristofari@unipd.it;

In this talk, a derivative-free algorithm is proposed to minimize a black-box objective function over the convex hull of a given set of points, focusing on problems with sparse optimal solutions. At each iteration, the proposed approach solves a reduced problem, based on an inner approximation of the feasible set. This inner approximation is dynamically updated by using rules that in general allow us to keep the dimension of the reduced problem small. We show global convergence to stationary points and, under suitable assumptions, identification properties of the algorithm. Finally, numerical results are provided.

■ Distributed Optimization

Room: **IACOCCA** (14:00 - 15:30)   Chair: *Charikleia Iakovidou*

1. A flexible framework for noisy and nonconvex distributed optimization

Charikleia Iakovidou[1,*]

[1]Northwestern University; [*]chariako@u.northwestern.edu;

We present a flexible framework for distributed optimization, where the amounts of computation and communication executed at each iteration can be tailored on a case-by-case basis to balance convergence accuracy and application cost. We then investigate the performance of this framework under three major concerns encountered in modern distributed systems: costly gradient evaluations, communication bottlenecks and nonconvex objective functions. First, we demonstrate that when cheaper stochastic gradient approximations are used instead of the true gradients and the communication between nodes is probabilistically quantized to alleviate bandwidth demands, our algorithm achieves geometric convergence to a neighborhood of the optimal solution for strongly convex functions, and can outperform state-of-the-art methods depending on the implementation of quantized consensus. Next we show that when the global objective function is nonconvex, our method successfully evades saddle points and approaches the minimizers of the original problem under mild assumptions.

2. Scalable Compressed Communication in Distributed Inference and Optimization

César A Uribe[1,*], Mohammad Taha Toghani[1]

[1]Rice University; [*]cauribe@rice.edu;

We propose a new decentralized consensus algorithm with compressed communication that scales linearly with the network size n. Moreover, we present variations of the proposed method for distributed optimization and distributed inference tasks. In both applications, we prove convergence guarantees and convergence rates for a wide class of compression operators in the local communication between agents, and arbitrary static undirected and connected networks. We further present numerical experiments that confirm our theoretical results and illustrate the scalability and communication-efficiency of our algorithms.

3. Exploiting Shared Representations for Personalized Federated Learning

Liam Collins[1,*], Hamed Hassani[2], Aryan Mokhtari[1], Sanjay Shakkottai[1]

[1]University of Texas at Austin; [*]liamc@utexas.edu;[2]University of Pennsylvania;

Neural networks have shown the ability to extract universal feature representations from data such as images and text that have been useful for a variety of learning tasks. However, the fruits of representation learning have yet to be fully-realized in federated learning (FL). In this talk, we propose FedRep: a novel federated learning framework and algorithm for learning a shared data representation across clients and unique local heads for each client. We prove that FedRep learns the ground-truth representation with per-user sample complexity that diminishes with the number of users in a linear setting, demonstrating that FedRep harnesses the benefits of collaboration in FL. Finally, we discuss experimental results showing that FedRep outperforms a variety of personalized FL methods on multiple data-heterogeneous FL benchmarks.

■ Robust and Stochastic Optimization under Uncertainty

Room: **WHITEHOUSE** (14:00 - 15:30) Chair: *Weijun Xie*

1. Distributionally Robust Optimization with Markovian Data

Mengmeng Li[1,], Tobias Sutter[1], Daniel Kuhn[1]*

[1]EPFL; [*]mengmeng.li@epfl.ch;

We study a stochastic program where the probability distribution of the uncertain problem parameters is unknown and only indirectly observed via finitely many correlated samples generated by an unknown Markov chain with d states. We propose a data-driven distributionally robust optimization model to estimate the problem's objective function and the corresponding optimal decision. By leveraging results from large deviations theory, we derive statistical guarantees on the quality of this estimator. The distributionally robust optimization problem is a nonconvex program of size $O(d^2)$. By exploiting the underlying problem structure we propose a customized Frank-Wolfe algorithm to solve it with simple convex oracle subproblems of size $O(d)$. Numerical experiments show that our approach statistically outperforms existing methods from the literature.

2. Affine and Constant Policies in Adjustable Linear Robust Optimization: A New Perspective

Ningji Wei[1,]*

[1]Heinz, CMU; [*]ningjiwei@cmu.edu;

We provide general conditions under which constant (a.k.a. static) policies and affine policies (a.k.a., linear decision rules) are optimal for adjustable robust linear optimization. Our results provide a unifying framework to reinterpret and extend several existing results in the robust optimization literature and provide new geometric insights to understand this class of problems.

3. ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs

Weijun Xie[1,]*

[1]Virginia Tech; [*]wxie@vt.edu;

Chance constrained programs (CCPs) are generic frameworks for decision-making under uncertain constraints. The objective of a CCP is to find the best decision that violates the uncertainty constraints within the prespecified risk level. A CCP is often nonconvex and is difficult to solve to optimality. This paper studies and generalizes the ALSO-X, originally proposed by Ahmed, Luedtke, SOng, and Xie (2017), for solving a CCP. We first show that the ALSO-X resembles a bilevel optimization, where the upper-level problem is to find the best objective function value and enforce the feasibility of a CCP for a given decision from the lower-level problem, and the lower-level problem is to minimize the expectation of constraint violations subject to the upper bound of the objective function value provided by the upper-level problem. This interpretation motivates us to prove that when uncertain constraints are convex in the decision variables, ALSO-X always outperforms the CVaR approximation. We further show (i) sufficient conditions under which ALSO-X can recover an optimal solution to a CCP; (ii) an equivalent bilinear programming formulation of a CCP, inspiring us to enhance ALSO-X with a convergent alternating minimization method (ALSO-X+); (iii) extensions of ALSO-X and ALSO-X+ to solve distributionally robust chance constrained programs (DRCCPs) under Wasserstein ambiguity set. Our numerical study demonstrates the effectiveness of the proposed methods.

■ First Order Methods

Room: **MONRO** (16:45 - 18:15) Chair: *Dimitri Papadimitriou*

1. Stochastic nonlinear ADMM

Dimitri Papadimitriou[1,]*

[1]3NLab; [*]dpapadimitriou@3nlab.org;

This talk investigates a class of structured nonconvex optimization problems of the form min F(x) = f(x) + h(x) + g(c(x)-b) where f and g are proper l.s.c convex functions that are not necessarily differentiable, h is a differentiable nonconvex function and c a nonlinear differentiable mapping. For this purpose, we first develop the alternating direction method of multipliers (ADMM) which yields a nonlinear ADMM. Further, application to statistical learning from large datasets, requires extending the deterministic ADMM to the stochastic setting. Recently, the solving of nonconvex stochastic ADMMs with equality constraints has been proposed including the setting with h nonconvex and smooth, and g convex and possibly nonsmooth. Some of the recently proposed methods such as SVRG-ADMM and SAGA-ADMM further combine Variance Reduction with ADMM. However, in all them, the mapping c remains nevertheless linear. As we well demonstrate, the involvement of nonlinear function c is critical for solving practical instances of the general problem.

2. A Communication Compression Decentralized Algorithm for Convex Composite Optimization

Yao Li[1,], Xiaorui Liu[1], Ming Yan[1], Kun Yuan[2]*

[1]Michigan State University; [*]liyao6@msu.edu;[2]Alibaba Group;

Decentralized optimization problem has extensive applications in large-scale machine learning. Recently, several compression techniques are proposed and combined with some existing decentralized algorithms to tackle the communication bottleneck. However, there is no algorithm considering decentralized problems with nonsmooth regularizer to date. We propose a stochastic algorithm to solve it with the communicated information compressed by unbiased stochastic operator of arbitrary precision, where the strong convexity is assumed on the smooth component of the objective function and the linear convergence up to the neighborhood of the solution is guaranteed. For the problems with finite-sum structure, we accelerate the algorithm by two well-known variance reduction schemes, Loopless-SVRG and SAGA, to exploit the exact linear convergence. Detailed complexity is investigated to show the effect of the inexact communication on the compressed algorithm compared to the vanilla one.

3. Online Nash Social Welfare Maximization with Predictions

Billy Zhengxu Jin[1,]*

[1]Cornell University; [*]bzj3@cornell.edu;

We consider the problem of allocating a set of divisible goods to N agents in an online manner, aiming to maximize the Nash social welfare, a widely studied objective which provides a balance between fairness and efficiency. The goods arrive in a sequence of T periods and the value of each agent for a good is adversarially chosen when the good arrives. We first observe that no online algorithm can achieve a competitive ratio better than the trivial $O(N)$, unless it is given additional information about the agents' values.

We consider a setting where for each agent, the online algorithm is only given a prediction of her *monopolist utility, i.e., her utility if all goods were given to her alone (corresponding to the sum of her values over the T periods)*. Our main result is an online algorithm whose competitive ratio is parameterized by the multiplicative errors in these predictions. The algorithm achieves a competitive ratio of $O(\log N)$

and $O(\log T)$ if the predictions are perfectly accurate. Moreover, the competitive ratio degrades smoothly with the errors in the predictions, and is surprisingly robust: the logarithmic competitive ratio holds even if the predictions are very inaccurate. Our bounds are essentially tight: no online algorithm, even if provided with perfectly accurate predictions, can achieve a competitive ratio of $O(\log^{1-\epsilon} N)$ or $O(\log^{1-\epsilon} T)$ for any constant $\epsilon \geq 0$.

■ New Deterministic and Stochastic Methods for Derivative-Free Optimization II

Room: **LEHIGH** (16:45 - 18:15)   Chair: *Tommaso Giovannelli*

1. A linesearch-based derivative-free approach for mixed-integer nonsmooth constrained optimization

Tommaso Giovannelli[1,]*

[1]Lehigh University; [*]tog220@lehigh.edu;

Mixed-integer nonsmooth optimization problems with black-box objective and constraint functions frequently arise in many real-world applications. Since first-order information is unavailable, derivative-free optimization algorithms are required. To this end, a novel linesearch-based approach is first developed for bound-constrained problems and then extended to the nonlinearly nonsmooth constrained case by using an exact penalty approach. The nonsmoothness of the objective function is handled by using a dense sequence of directions, while primitive vectors are adopted to deal with the discrete variables. Extensive numerical results highlight the effectiveness of the approach.

2. Hessian-Aided Random Perturbation (HARP) Using Noisy Zeroth-Order Queries

Jingyi Zhu[1,]*

[1]Resource Optimization Team, Decision Intelligence Lab, DAMO Academy, Alibaba Group; [*]jingyi.zhu@jhu.edu;

In stochastic optimization problems where only noisy zeroth order (ZO) oracles are available, the Kiefer-Wolfowitz algorithm and its randomized counterparts are widely used as gradient estimator. Existing algorithms generate the randomized perturbation from a zero mean and unit-covariance distribution. In contrast, this work considers the generalization where the perturbations may have a non-isotropic covariance constructed from the history of the ZO queries. We propose to feed the second-order approximation into the covariance matrix of the random perturbation, so it is dubbed as Hessian-Aided Random Perturbation (HARP). HARP collects several (two or more, depending on the specific estimator form) zeroth-order queries per iteration to form approximations for both the gradient and the Hessian. We show the convergence (in almost surely sense) and derive the convergence rate for HARP under mild assumptions. We demonstrate, with theoretical guarantees and numerical experiments, that HARP is less sensitive to ill-conditioning and more query-efficient than other gradient approximation schemes whose random perturbation has an identity covariance.

3. High Probability Complexity Bounds for Line Search Based on Stochastic Oracles

Miaolan Xie[1,], Billy Jin[2], Katya Scheinberg[1]*

[1]Cornell University; [*]mx229@cornell.edu;[2]Cornell University;

We consider a line-search method for continuous optimization under a stochastic setting where the function values and gradients are available only through inexact probabilistic zeroth and first-order oracles. These oracles capture multiple standard settings including expected loss minimization and zeroth-order optimization. Moreover, our framework is very general and allows the function and gradient estimates to be biased. The proposed algorithm is simple to describe, easy to implement, and uses these oracles in a similar way as the standard deterministic line search uses exact function and gradient values. Under fairly general conditions on the oracles, we derive a high probability tail bound on the iteration complexity of the algorithm when applied to non-convex smooth functions. These results are stronger than those for other existing stochastic line search methods and apply in more general settings.

■ Robust Learning

Room: **IACOCCA** (16:45 - 18:15)   Chair: *Guangyi Liu*

1. Practical Convex Formulation of Robust Neural Network Training

Yatong Bai[1,]*

[1]Department of Mechanical Engineering, University of California, Berkeley; [*]yatong_bai@berkeley.edu;

Recent work has shown that the training of a one-hidden-layer, scalar-output fully-connected ReLU neural network can be reformulated as a finite-dimensional convex program, enabling tractable global optimization of such a neural network. However, the scale of this convex program grows exponentially in data size. In this work, we prove that a stochastic procedure with a linear complexity well approximates the exact formulation. We also derive a convex optimization approach to efficiently solve the "adversarial training" problem, which trains neural networks that are robust to adversarial input perturbations. Our method can be applied to binary classification and regression, and provides an alternative to the current adversarial training methods, such as Fast Gradient Sign Method (FGSM) and Projected Gradient Descent (PGD). The proposed method achieves a noticeably better adversarial robustness and performance than the existing methods.

2. Robust Learning of Recurrent Neural Networks in Presence of Exogenous Noise

Arash Amini[1,]*

[1]Lehigh Univeristy; [*]ara416@lehigh.edu;

Recurrent Neural networks (RNN) have shown promising potential for learning dynamic features of sequential data. However, artificial neural networks are known to exhibit poor robustness in presence of noise input, where the sequential architecture of RNNs exacerbates the problem. In this paper, we will use ideas from control and estimation theories to propose a tractable robustness analysis for RNN models that are subject to noisy inputs. The variance of the output of the noisy system is adopted as a robustness measure to quantify the impact of noise on learning. It is shown that the robustness measure can be estimated efficiently using linearization techniques. Using these results, we proposed a learning method to enhance robustness of a RNN with respect to exogenous Gaussian noise with known statistics. Several theoretical upper bounds are also presented to show how the robustness measure depends on the trainable parameters. Our extensive

simulations on benchmark problems reveal that our proposed methodology significantly improves robustness of recurrent neural networks in a systematic manner.

3. Robust Classification with Localized Observations Using Stable Recurrent Neural Networks

Guangyi Liu[1,*]*, Arash Amini*[1]*, Nader Motee*[1]

[1]Lehigh University; [*]gul316@lehigh.edu;

We consider the problem of distributed multi-robot classification, where a network of robots collect localized observations from an environment, communicate their beliefs with their neighboring robots, and learn to classify the underlying environment. We show that robots will conclude convergent classifications on specific paths (i.e., localized observations) if trained to classify with an exponentially stable recurrent neural network. The same convergent result is also valid for a network of robots that share and fuse their beliefs. Moreover, we show that the robustness of classification can be drastically improved by using a network of robots. Finally, our theoretical results are validated with extensive simulations of map and image classification problems via Reinforcement Learning.

∎ Optimization & Machine Learning

Room: **WHITEHOUSE** (16:45 - 18:15)   Chair: *Matthias Chung*

1. Optimal Decision in a Statistical Process Control with Cubic Loss

Vladimir Turetsky[1,*]*, Valery Y. Glizer*[1]

[1]Ort Braude College of Engineering, Karmiel, Israel;
[*]turetsky1@braude.ac.il;

The Statistical Process Control (SPC) is a method for control of a process quality by monitoring its state based on statistical samples of a process characteristic index, taken in some time intervals. The SPC is widely used in industry, medicine, veterinary, environment control, etc. Its objective is to minimize losses which can be caused by delay in the detection of undesirable process changes. During many years, the traditional SPC practice was to take the samples with a fixed time interval. The use of a variable time sampling interval was suggested for the first time in [1]. Then, this approach was developed in a number of works (see e.g. [2] and references therein). In [1], the detection delay was considered as a criterion for optimality of the time-sampling interval. More general criterion, an expected loss caused by such a delay, was used in the authors' works [2,3,4], where this loss was chosen as a quadratic function of the delay. However, there exist real-life problems with another, more general, relations between the delay and the loss. Thus, in [5] it was shown that the relationship between the delay in the detection of some cow's disease and an economic loss, caused by this delay, is close to a cubic function.

In this talk, we consider the SPC time-sampling optimization with the loss as a pure cubic function of the detection delay. The mathematical model of this problem is an optimal control problem, treated by the Pontryagin's Maximum Principle.

References
[1] M. R. Reynolds, R. W. Amin, J. C. Arnold and J. Nachlas, X charts with variable sampling intervals. Technometrics, Vol. 30, pp. 181âĂŞ192, 1988.
[2] V. Y. Glizer and V. Turetsky, Optimal time-sampling problem in a statistical control with a quadratic cost functional: analytical and numerical approaches. Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics, Porto, Portugal, Vol. 1, pp. 21âĂŞ32, 2018.
[3] E. Bashkansky and V. Y. Glizer, Novel approach to adaptive statistical process control optimization with variable sampling interval and minimum expected loss. International Journal of Quality Engineering and Technology, Vol. 3, pp. 91âĂŞ107, 2012.
[4] V. Y. Glizer, V. Turetsky and E. Bashkansky, Statistical process control optimization with variable sampling interval and nonlinear expected loss. Journal of Industrial and Management Optimization, Vol. 11, pp. 105âĂŞ133, 2015.
[5] T. E. Carpenter, J. M. O'Brien, A. Hagerman and B. McCarl, Epidemic and economic impacts of delayed detection of foot-and-mouth disease: a case study of a simulated outbreak in California. Journal of Veterinary Diagnostic Investigation, Vol. 23, pp. 26-33, 2011

2. Big Data Inverse Problems

Matthias Chung[1,*]

[1]Virginia Tech; [*]mcchung@vt.edu;

Emerging fields such as data analytics, machine learning, and uncertainty quantification heavily rely on efficient computational methods for solving inverse problems. With growing model complexities and ever increasing data volumes, state of the art inference method exceeded their limits of applicability and novel methods are urgently needed. Hence, new inference method need to focus on the scalability to large dimension and to address model complexities.

In this talk, we discuss massive least squares problems arising in training neural networks and for imaging applications where the size of the forward model exceeds the storage capabilities of computer memory or the data is simply not available all at once. We consider randomized row-action methods that can be used to approximate the solution. We introduce a sampled limited memory row-action method for least squares problems, where an approximation of the global curvature of the underlying least squares problem is used to speed up the initial convergence and to improve the accuracy of iterates. Our proposed methods can be applied to ill-posed inverse problem, where we establish sampled regularization parameter selection methods. Numerical experiments on superresolution, tomographic reconstruction, and neural networks demonstrate the efficiency of these sampled limited memory row-action methods. This is joint work with Julianne Chung, Elizabeth Newman, Lars Ruthotto, Tanner Slagel, and Luis Tenorio.

■ Large Scale Optimization

Room: **MONRO** (10:15 - 11:45)

Chair: *Salar Fattahi*

1. Large-scale Inference of Sparsely-varying Markov Random Fields

Salar Fattahi[1,], Andres Gomez[2]*

[1]University of Michigan; [*]`fattahi@umich.edu`;[2]University of Southern California;

We study the problem of inferring time-varying Markov random fields (MRF), where the underlying graphical model is both sparse and changes sparsely over time. Most of the existing methods for the inference of time-varying MRFs rely on the regularized maximum likelihood estimation, that typically suffer from weak statistical guarantees and high computational time. Instead, we introduce a new class of constrained optimization problems for the inference of sparsely-changing MRFs. The proposed optimization problem is formulated based on the exact L0 regularization, and can be solved in near-linear time and memory. Moreover, we show that the proposed estimator enjoys a provably small estimation error. Our proposed method is extremely efficient in practice: it can accurately estimate time-varying graphical models with more than 500 million variables within one hour.

2. Global optimization using random embeddings

Estelle Massart[1,], Coralia Cartis[2], Adilet Otemissov[2]*

[1]University of Oxford - NPL;
[*]`estelle.massart@maths.ox.ac.uk`;[2]Mathematical Institute, University of Oxford, and The Alan Turing Institute;

We present a general random subspace algorithmic framework for global optimization and analyse its convergence using tools from conic integral geometry and random matrix theory. We then particularise this framework and analysis for the class of functions with low effective dimension. We show that its convergence does not depend on the ambient dimension, and are able to estimate the effective dimension in the run of the algorithm. Encouraging numerical results are also presented that use local or global solvers in the subspace.

3. REX: Revisiting Budgeted Training with an Improved Schedule

John Chen[1,]*

[1]Rice University; [*]`jc114@rice.edu`;

Deep learning practitioners often operate on a computational and monetary budget. Thus, it is critical to design optimization algorithms that perform well under any budget. The linear learning rate schedule is considered the best budget-aware schedule [**?**], as it outperforms most other schedules in the low budget regime. On the other hand, learning rate schedules –such as the `30-60-90` step schedule– are known to achieve high performance when the model can be trained for many epochs. Yet, it is often not known a priori whether one's budget will be large or small; thus, the optimal choice of learning rate schedule is made on a case-by-case basis. In this paper, we frame the learning rate schedule selection problem as a combination of i) selecting a profile (i.e., the continuous function that models the learning rate schedule), and ii) choosing a sampling rate (i.e., how frequently the learning rate is updated/sampled from this profile). We propose a novel profile and sampling rate combination called the Reflected Exponential (REX) schedule, which we evaluate across seven different experimental settings with both SGD and Adam optimizers. REX outperforms the linear schedule in the low budget regime, while matching or exceeding the performance of several state-of-the-art learning rate schedules (linear, step, exponential, cosine, step decay on plateau, and OneCycle) in both high and low budget regimes. Furthermore, REX requires no added computation, storage, or hyperparameters.

■ Optimization and Energy

Room: **LEHIGH** (10:15 - 11:45)

Chair: *Ben Wang*

1. Building Load Control using Distributionally Robust Chance-Constrained Programs with Right-Hand Side Uncertainty and the Risk-Adjustable Variants

Yiling Zhang[1,]*

[1]University of Minnesota; [*]`yiling@umn.edu`;

Aggregation of heating, ventilation, and air conditioning (HVAC) loads can provide reserves to absorb volatile renewable energy, especially solar photo-voltaic (PV) generation. However, the time-varying PV generation is not perfectly known when the system operator decides the HVAC control schedules. To consider the unknown uncertain PV generation, in this talk, we consider a distributionally robust chance-constrained (DRCC) building load control problem under two typical ambiguity sets: the moment-based and Wasserstein ambiguity sets. We derive mixed integer linear programming (MILP) reformulations for DRCC problems under both sets. Especially, for the DRCC problem under the Wasserstein ambiguity set, we utilize the right-hand side (RHS) uncertainty to derive a more compact MILP reformulation than the commonly known MILP reformulations with big-M constants. All the results also apply to general individual chance constraints with RHS uncertainty. Furthermore, we propose an adjustable chance-constrained variant to achieve a trade-off between the operational risk and costs. We derive MILP reformulations under the Wasserstein ambiguity set and second-order conic programming (SOCP) reformulations under the moment-based set. Using real-world data, we conduct computational studies to demonstrate the efficiency of the solution approaches and the effectiveness of the solutions.

2. A Semidefinite Optimization-based Branch-and-Bound Algorithm for Several Reactive Optimal Power Flow Problems

Miguel Anjos[1,]*

[1]University of Edinburgh; [*]`miguel.f.anjos@ed.ac.uk`;

The Reactive Optimal Power Flow (ROPF) problem consists in computing an optimal power generation dispatch for an alternating current transmission network that respects power flow equations and operational constraints. Some means of action on the voltage are modelled in the ROPF problem such as the possible activation of shunts, which implies discrete variables. The ROPF problem belongs to the class of nonconvex MINLPs (Mixed-Integer Nonlinear Problems), which are NP-hard problems. In this paper, we solve three new variants of the ROPF problem by using a semidefinite optimization-based Branch-and-Bound algorithm. We present results on MATPOWER instances and we show that this method can solve to global optimality most instances. On the instances not solved to optimality, our algorithm is able to find solutions with a value better than the ones obtained by a rounding algorithm. We also demonstrate that applying an appropriate clique merging algorithm can significantly speed up the resolution of semidefinite relaxations of large ROPF instances.

3. Battery-Wind/Solar Farm Management: A Continuous Dynamic Programming Approach

Ben Wang[1,], Jeffery Choy[2], Abdullah Alshelahi[2], Romesh Saigal[2]*

[1]Department of Industrial & Operations Engineering, University of Michigan, Ann Arbor, Michigan, USA; [*]papaver@umich.edu;[2]Department of Industrial & Operations Engineering, University of Michigan, Ann Arbor, Michigan, USA;

The intermittent and volatile nature of production of renewable energy raises significant challenges for the wind/solar farm owners and operators to meet their customers' demand. One way to address this challenge is to utilize an energy storage facility such as a battery to store the excess energy and supply it when the production is in deficit. In this paper, we propose a single principal and single agent framework to achieve this goal. In our setting, the principal is the local energy distributor and owner of the farm is the agent. At each epoch of time, the principal compensates the agent for the supplied energy and the agent optimizes his operations to determine the quantity of energy to supply from both the output of the farm and the energy stored in the battery. The principal and agent problems are formulated as two interconnected stochastic optimal control problems (SOCPs). The objective of the SOCPs is to maximize the profits of the distributor and the farm in the continuous time over a given horizon. Using the Martingale representation theorem, we prove the existence of a control variable that allows principal to decouple these two SOCPs. Further, the optimality condition of the agent's problem is derived using dynamic programming techniques. While the principle optimal policy is determined by solving a Hamilton–Jacobi–Bellman (HJB) equation. In the case of finite horizon with a relaxed boundary condition on the battery charge, we compute an analytical solution with affine structure for the partial differential equation encountered in the HJB. The proposed approach is generic and can be extended to problems with several (two or more) farms providing energy.

■ Fairness in Machine Learning

Room: **IACOCCA** (10:15 - 11:45) Chair: *Suyun Liu*

1. Automating Procedurally Fair Feature Selection in Machine Learning

Clara Belitz[1,]*

[1]University of Illinois - Urbana Champaign; [*]cbelitz2@illinois.edu;

In recent years, machine learning has become more common in everyday applications. Consequently, numerous studies have explored issues of unfairness against specific groups or individuals in the context of these applications. Much of the previous work on unfairness in machine learning has focused on the fairness of outcomes rather than process. I will discuss these different approaches and their implications for how we optimize fairness and accuracy. I will also discuss a novel feature selection method inspired by fair process (procedural fairness) in addition to fair outcome. It specifically introduces the the notion of unfairness weight, which indicates how heavily to weight unfairness versus accuracy when measuring the marginal benefit of adding a new feature to a model. The goal is to maintain accuracy while reducing unfairness, as defined by six common statistical definitions. This approach selects unfair features and sensitive features for the model less frequently as the unfairness weight increases. As such, this procedure is an effective approach to constructing classifiers that both reduce unfairness and are less likely to include unfair features in the modeling process.

2. Balanced Districting on Grid Graphs with Provable Compactness and Contiguity

Cyrus Hettle[1,], Shixiang Zhu[1], Swati Gupta[1], Yao Xie[1]*

[1]Georgia Institute of Technology; [*]chettle@gatech.edu;

Given a graph $G = (V, E)$ with vertex weights $w(v)$ and a desired number of parts k, the goal in graph partitioning problems is to partition the vertex set V into parts $V_1, , V_k$. Metrics for compactness, contiguity, and balance of the parts V_i are frequent objectives, with much existing literature focusing on compactness and balance. Revisiting an old method known as striping, we give the first polynomial-time algorithms with guaranteed contiguity and provable bicriteria approximations for compactness and balance for planar grid graphs. We consider several types of graph partitioning, including when vertex weights vary smoothly or are stochastic, reflecting concerns in various real-world instances. We show significant improvements in experiments for balancing workloads for the fire department and reducing over-policing using 911 call data from South Fulton, GA.

3. A Stochastic Alternating Balance k-Means Algorithm for Fair Clustering

Suyun Liu[1,], Luis Nunes Vicente[1]*

[1]Department of Industrial and Systems Engineering, Lehigh University; [*]sul217@lehigh.edu;

In the application of data clustering to human-centric decision-making systems, such as loan applications and advertisement recommendations, the clustering outcome might discriminate against people across different demographic groups, leading to unfairness. A natural conflict occurs between the cost of clustering (in terms of distance to cluster centers) and the balance representation of all demographic groups across the clusters, leading to a bi-objective optimization problem that is nonconvex and nonsmooth. To determine the complete trade-off between these two competing goals, we design a novel stochastic alternating balance fair k-means (SAfairKM) algorithm, which consists of alternating classical mini-batch k-means updates and group swap updates. The number of k-means updates and the number of swap updates essentially parameterize the weight put on optimizing each objective function. Our numerical experiments show that the proposed SAfairKM algorithm is robust and computationally efficient in constructing well-spread and high-quality Pareto fronts both on synthetic and real datasets. Moreover, we propose a novel companion algorithm, the stochastic alternating bi-objective gradient descent (SA2GD) algorithm, which can handle a smooth version of the considered bi-objective fair k-means problem, more amenable for analysis. A sublinear convergence rate of O(1/T) is established under strong convexity for the determination of a stationary point of a weighted-sum function parameterized by the number of steps or updates on each function.

■ Applications of Optimization

Room: **WHITEHOUSE** (10:15 - 11:45) Chair: *Robert Bosch*

1. Variations on TSP Art

Robert Bosch[1,]*

[1]Oberlin College; [*]rbosch@oberlin.edu;

TSP Art is produced by (1) applying a stippling algorithm to a grayscale image, (2) considering the resulting point set to be the cities of a TSP instance, (3) finding a high-quality tour of the cities, and finally, (4) drawing (or laser cutting or 3D printing) the tour. One well known example is the Mona Lisa TSP Challenge. In this talk, we will discuss several variants of TSP Art. One variant arises when one wants

to force two city-free regions to be on the same side of the salesperson's tour (or alternatively, on opposite sides of it). Another allows for variations in how the distances between cities are computed. (For example, it might be that in one region, it is easier to travel north-south than east-west, while in another, the opposite holds.) We will also discuss techniques for designing structured knight's tours.

2. Rare-event simulations in a chain of dynamical systems with small random perturbations

Getachew Befekadu[1,*]

[1]Morgan State University;
[*]getachew.befekadu@morgan.edu;

In this talk, we consider an importance sampling problem for rare-event simulations involving the behavior of a diffusion process pertaining to a chain of dynamical systems with small random perturbations. Here, we assume the overall dynamical system is formed by n-subsystems in which a small random perturbation enters in the first subsystem and then subsequently transmitted to the other subsystems. We provide an efficient importance sampling estimator for the asymptotic probabilities of certain rare-events involving such a diffusion process that are difficult to observe in simulations. The approach for such an analysis basically relies on the connection between the probability theory of large deviations and that of the values functions for a family of stochastic control problems associated with the underlying dynamical system with small random perturbations, where such a connection also provides a computational framework for constructing efficient importance sampling estimators for rare- event simulations. Moreover, the framework also allows us to derive a family of Hamilton-Jacobi-Bellman equations for which we also provide a solvability condition for the corresponding optimal control problem. (Joint work with Dan Anyumba, Department of Electrical & Computer Engineering, Morgan State University).

3. Hungry for Equality: Fighting Food Deserts with Optimization

Drew Horton[1,*]

[1]University of Colorado Denver;
[*]drew.horton@ucdenver.edu;

Food deserts are a form of food insecurity related to a lack of access to healthy, fresh, and affordable food. According to the United States Department of Agriculture (USDA), 13.7 million households in the U.S. experienced food insecurity in 2019. This problem has only been exacerbated by the ongoing COVID-19 pandemic, and disproportionately affects marginalized communities. In one traditional approach where we seek to minimize the expected distance of the population to grocery stores, the worst-off members in our communities tend to be ignored in the solution as outliers. To address these food insecurities, and the existing inequities, we demonstrate how the Kolm-Pollak equally-distributed equivalent function (EDE) can be minimized over a facility location integer program to minimize not only expected distance but also the inequality of the distribution. The EDE is a nonlinear function making the problem computationally significantly harder than the traditional model, therefore we discuss various ways to approach the optimization including a piecewise linear relaxation of the model. We present results demonstrating how our model works on real-world data to produce an optimal distribution of grocery store locations. In minimizing the inequality, we are ensuring that we are prioritizing relief in disproportionately affected communities.

4. An alternative framework for the optimization of socially responsible portfolios applied to the Moroccan stock exchange

Yahya Hanine[1,*], *Yahya Hanine*[1], *Mohamed Tkiouat*[1], *Younes Lahrichi*[2]

[1]IFE-Lab, Laboratory of Applied Mathematical Studies and Research (LERMA), Mohammadia School of Engineering (EMI), Mohammed V University, Rabat, Morocco;
[*]yahyahanine@gmail.com;[2]Laboratory of Research in Finance (LAREF), ISCAE Business School, Casablanca, Morocco;

The purpose of this article is to propose an alternative approach for portfolio optimization combining financial and ethical constraints as well as objective and subjective preferences of investors. This approach intends to support investors in the selection and optimization of the performance of financial and social portfolios. More precisely, we introduce the Analytic Hierarchy Process (AHP) to measure the ethical performance (EP) score of each asset considering the ethical criteria. Fuzzy multiple criteria decision making (FMCDM) is used to determine the overall financial quality score of the assets with respect to key financial criteria, i.e., short-term return, long-term return, and risk. The interactive fuzzy programming approach is also applied to support the investor's decision, considering his subjective preferences. The robustness of our approach is tested through an empirical study involving the case of the Casablanca Stock Exchange (CSE). The results give evidence that the Socially Responsible (SR) portfolio performed similarly to the conventional one, as no significant differences were found in terms of return. However, the SR portfolio allows the investor to achieve their ethical goals with a slight financial sacrifice. Keywords: Portfolio optimization; SRI; MCDM; fuzzy set theory; AHP

■ Higher Order Optimization Methods

Room: **MONRO** (14:00 - 15:30)   Chair: *Albert S. Berahas*

1. A Trust Region-type Normal Map Semismooth Newton Method For Nonsmooth Large-scale Optimization

Andre Milzarek[1,*]

[1]The Chinese University of Hong Kong, Shenzhen;
[*]andremilzarek@cuhk.edu.cn;

We propose a normal map-based semismooth Newton method with a trust region mechanism for solving composite problems involving smooth nonconvex and nonsmooth convex terms in the objective function. The considered class of problems comprises a large variety of applications such as large-scale problems arising in machine learning or imaging. Our method uses semismooth Newton steps for a normal map-based formulation of the first-order optimality conditions. We combine the Newton steps with a trust region-type globalization to ensure global convergence. We further show that the Kurdyka-Lojasiewicz framework is applicable and that transition to fast local convergence can be obtained. Finally, extensions using approximate Hessian information and numerical results are discussed.

2. A New Multipoint Secant Method with a Dense Initial Matrix

Jennifer Erway[1,], Oleg Burdakov[2], Mostafa Rezapour[1]*

[1]Wake Forest University; [*]erwayjb@wfu.edu;[2]Linköping University;

In this presentation, we discuss a new multipoint symmetric secant (MSS) that uses a dense initial matrix rather than a multiple of the identity initial matrix. We discuss the convergence analysis of the new method and compare the numerical results of applying the new method with the standard MSS, which uses a multiple of the identity initial matrix, on several problems from the CUTEst test problem set.

3. A Subspace Acceleration Method for Minimization Involving a Group Sparsity-Inducing Regularizer

Yutong Dai[1,], Daniel P. Robinson[1], Frank E. Curtis[1]*

[1]Industrial and Systems Engineering, Lehigh University; [*]yud319@lehigh.edu;

We consider the problem of minimizing an objective function that is the sum of a convex function and a group sparsity-inducing regularizer. Problems that integrate such regularizers arise in modern machine learning applications, often for the purpose of obtaining models that are easier to interpret and that have higher predictive accuracy. We present a new method for solving such problems that utilize subspace acceleration, domain decomposition, and support identification. Our analysis provides the global iteration complexity on obtaining an ϵ accurate solution and shows that, under common assumptions, the iterates locally converge superlinearly. Numerical results on regularized logistic regression and regularized linear regression show that our approach is efficient and robust, with the ability to outperform state-of-the-art methods.

■ Energy and Optimization

Room: **LEHIGH** (14:00 - 15:30)   Chair: *Luis Zuluaga*

1. On the Absence of Spurious Local Trajectories in Time-varying Nonconvex Optimization

Cedric Josz[1,], Salar Fattahi[2], Yuhao Ding[3], Reza Mohammadi[3], Javad Lavaei[3], Somayeh Sojoud[3]*

[1]Columbia University; [*]cj2638@columbia.edu;[2]University of Michigan;[3]University of California;

We study the landscape of a time-varying nonconvex optimization problem, for which the input data vary over time and the solution is a trajectory rather than a single point. A motivating example will be the alternating current optimal power flow problem where the demand varies throughout the day. To understand the complexity of finding a global solution of such a problem, we introduce the notion of spurious (i.e., non-global) local trajectory as a generalization to the notion of spurious local solution in nonconvex (time-invariant) optimization. We develop an ordinary differential equation (ODE) associated with a time-varying nonlinear dynamical system which, at limit, characterizes the spurious local solutions of the time-varying optimization problem. We prove that the absence of spurious local trajectory is closely related to the transient behavior of the developed system. In particular, we provide sufficient conditions for the ODE to escape spurious local minima due to time variations.

2. Revenue Adequate Prices for Chance-constrained Electricity Markets with Renewable Energy Sources

Luis Zuluaga[1,], Xin Shi[2], Alberto Lamadrid[1]*

[1]Lehigh University; [*]luis.zuluaga@lehigh.edu;[2]Emovi;

In a commodity market, revenue adequate prices refer to compensations that ensure that a market participant has a non-negative profit. We study the problem of deriving revenue adequate prices for an electricity market-clearing model with uncertainties resulting from the use of renewable energy sources (RES). To handle the uncertainty, we use a chance-constrained optimization (CCO) approach. Then, we show how prices that satisfy revenue adequacy in expectation for the market operator, and cost recovery in expectation for all conventional and RES generators, can be obtained from the optimal dual variables associated with the deterministic equivalent of the CCO market-clearing model.

3. Financially Adequate Environmental Pricing Designs

Alberto Lamadrid[1,]*

[1]Lehigh University; [*]ajlamadrid@lehigh.edu;

The regular conduct of these business activities sometimes has associated byproducts that have unintended environmental consequences and may pose risks to human health. A potential way to manage these ramifications can be done by organizing a market. For example, businesses may trade pollution permits, as was the case with Title IV of the 1990 Clean Air Act Amendments (CAAA) in the United States for sulfur dioxide (SO2). In fact, business subject to environmental regulation could innovate in ways that lead them to reduce inefficiencies, improve the total factor productivity, and increase competitiveness, the Porter Hypothesis (Porter 1991).

This would leave a fundamental question for stakeholders. How to design markets to determine prices, first, with uncertainty, and second, that explicitly account for the reciprocal environmental effects occurring in the course of business?

To answer these questions, we derive a new financially adequate market clearing pricing scheme. By financial adequacy we mean that the market administrator does not have losses, and the overall market is non-confiscatory, allowing for example, firms to recover their costs. This financial adequacy is sometimes called revenue adequacy for the market administrator and cost recovery for producers in the context of electricity systems.

Unlike related financially adequate pricing schemes that only take into account the marginal market costs associated with market clearing commodity demands, the proposed pricing scheme also takes into account the marginal market costs associated with the market participants operating at maximum capabilities. The proposed pricing scheme guarantees revenue adequacy in expectation for the market administrator, and cost recovery in expectation for all producers. More importantly, the proposed pricing scheme allows to analyze the effects of environmental limitations in the market, like CO2 emissions, as it internalizes, in the pricing signals, the costs or benefits associated with compliance of these limits by the market participants; and in particular, conventional technology producers.

■ Optimization & Applications 1

Room: **IACOCCA** (14:00 - 15:30) Chair: *Gulcin Dinc Yalcin*

1. The Quadratic Assignment Problem: Explaining Solvable Cases via Linear Reformulations

Lucas Waddell[1,*]

[1]Bucknell University; [*]lucas.waddell@bucknell.edu;

The quadratic assignment problem is a well-known, NP-hard discrete optimization program that has been extensively studied for over 60 years. This talk presents a recently discovered connection between special objective function structures that permit the QAP to be solved in polynomial time, and a mixed 0-1 linear reformulation of the QAP that is commonly used in state-of-the-art exact solution algorithms. Specifically, we show how these special solvable cases can be explained in terms of the dual region to the continuous relaxation of the level-1 reformulation-linearization technique (RLT) representation.

2. A Secretary Problem with Uncertain Offer Acceptance

Sebastian Perez-Salazar[1,*], *Mohit Singh*[1], *Alejandro Toriello*[1]

[1]Georgia Institute of Technology; [*]sperez@gatech.edu;

Motivated by the problem of displayed online advertisement where customers may fail to click on an ad, and by the problem of hiring candidates that can turn down job offers, we consider an online selection problem where a candidate may or may not accept an offer according to a known probability. Candidates arrive online, in random order. Upon an arrival, a decision maker observes the candidate's partial rank compared to previously observed candidates and either extends an offer to the candidate or moves to the next candidate, without the possibility of recalling previously observed candidates. If an offer is extended, the corresponding candidate accepts the offer with probability p, in which case the process ends, or rejects it, in which case the decision maker moves to the next candidate.

Because the decision maker does not know the top candidate willing to accept an offer, the goal is to maximize a robust objective defined as the minimum over integers k of the probability of choosing one of the top k candidates given that one of these candidates will accept an offer. This robust objective compares an online algorithm against an adversary that knows which candidates will accept the offer. Using Markov decision process theory, we derive an exact linear program for this max-min objective that characterizes the policy space. We further relax this linear program into an infinite counterpart, which we use to provide non-trivial asymptotic upper bounds for our objective. For values of p at least 0.594 we provide tight lower bounds and optimal policies, while for p less than 0.594 we show that our robust objective is at least 0.466.

3. A new mathematical model and a greedy algorithm for the tourist trip design problem under new constraints: a real-world application

Gulcin Dinc Yalcin[1,*], *Hilal Malta*[2], *Seher Saylik*[2]

[1]Eskisehir Technical University / Lehigh University; [*]gdinc@eskisehir.edu.tr / gud221@lehigh.edu;[2]Eskisehir Technical University;

The aim of the tourist trip design problem (TTDP) is to generate routes for tourists to maximize the points of interest (POIs) visited within specific time windows. In this study, we considered new constraints: budget, weather, and break. First, the budget is required for entrance fees and the distance between two points where a taxi has to be used. Additionally, the expense of the break was taken into account. Then, the weather was considered for summer and other seasons. On a summer day, tourists are likely to prefer visiting POIs, which are indoor areas, between specific times e.g. 11 a.m. to 3 p.m. to protect against the side effects of the sun. Furthermore, tourists need to take a break to relax during the trip. We formulated TTDP with these new constraints (TTDP-BWB) using a mixed-integer nonlinear programming (MINLP) approach. Then, a greedy algorithm was developed with a new greedy function that took the new constraints into account. We coded the algorithm using Android Studio and developed a mobile application for the case of Eskisehir in Turkey. We generated problems on the small and medium scale for the case of Eskisehir and used large-scale problems from published literature. We compared the results of the algorithm with MINLP results for the small-scale problems. Computational results showed that the algorithm is promising.

Index of Authors