## How to Navigate the Technical Sessions

There are four primary resources to help you understand and navigate the Technical Sessions:

- This Technical Session listing, which provides the most detailed information. The listing is presented chronologically by day/time, showing each session and the papers/abstracts/authors within each session.

### The Session Codes

**TA01**

**Room number.** Room locations are also indicated in the listing for each session.

**Time Block.** Matches the time blocks shown in the Program Schedule.

**The day of the week**

### Time Blocks

#### Thursday

**A** — 9:00am - 10:30am
**B** — 11:00am - 12:30pm
**C** — 1:30pm - 3:00pm
**D** — 3:00pm - 4:00pm
**E** — 4:00pm - 5:30pm

#### Friday

**A** — 8:30am - 10:00am
**B** — 10:30am - 11:00pm
**C** — 1:00pm - 2:30pm
**D** — 2:40pm - 3:40pm

#### Saturday

**A** — 8:30am - 10:00am
**B** — 10:30am - 12:00pm
**C** — 1:30pm - 2:30pm

# Thursday, 9:00am – 10:30am

## ■ TA01

Dodds Auditorium

**Tutorial – Machine Learning in Policy Search**

Contributed Session

**Lifelong Learning in Policy Search**

Chair: Haitham Bou Ammar, Princeton University,
465 Meadow Road, Apt 4203, Princeton, NJ, United States, 08540

**1. Machine Learning in Policy Search**

Haitham Bou Ammar, Princeton University, Princeton, NJ,
Contact: hammar@princeton.edu

Policy search approaches represent a promising method for solving sequential decision making problems. In the policy search setting, we assume that we are given a class of policies, parameterized by an unknown vector of parameters, mapping from the states to the actions. The goal is to determine the optimal parameters that maximize the total expected reward. A common problem with policy search is that the search through the policies can be difficult and computationally expensive. This problem is further manifested when including complex policy forms (e.g., when adding forecasts to the state space to facilitate learning) typical for real-world applications.In this tutorial we will target the above framework and shed-the-light on a new method for leveraging some of the problems inherent to policy search. First, a novel method for incorporating forecasts into the state variables is introduced. This method considers complex policy forms represented by neural networks. The neural network weight connections are then trained upon environmental interactions to maximize the total expected return.Though successful in some settings, it becomes clear that complex policy forms are hard to train due to a variety of problems including stochasticity, non-stationarity, and learning from scratch. Here, we focus on the third problem, where instead of learning policies from scratch, as in standard methods, a novel method based on lifelong machine learning is adopted. In lifelong learning, the agent has to operate within an online multi-task learning setting, where it is required to learn a series of sequential decision making problems over its lifetime. The agent will learn the tasks consecutively, acquiring multiple trajectories within each task before moving to the next. The tasks may be interleaved, giving the agent the opportunity to revisit earlier tasks for further experience, but the agent has no control over the task order. To reduce computational and sample complexities, we present a method which supports transfer by assuming the existence of shared knowledge repositories between tasks which are learnt (online) upon trajectory observations. These repositories are fixed across tasks thus capturing a non-stationary shared structure which is then tailored using task-specific projections to suit the peculiarities of each problem domain.In a generalization of this previous method, we finally show that such transferred knowledge can arrive from highly dissimilar sources. Namely, we present a framework for lifelong transfer in policy search where knowledge arrives from tasks belonging to different domains (i.e., exhibiting varying state and action representations).We present successful results of these methods in a set of experiments varying from simulated battery storage problems to real-world robotics including an application of controlling a quadcopter unmanned aerial vehicle.

**2. On the Implementation of Lifelong Learning Algorithms**

Jose Marcio Luna, University of Pennsylania, Penn Engineering, Philadelphia, PA, joseluna@seas.upenn.edu

Lifelong learning is of great importance to creating general-purpose machines. It enables us to rapidly learn new tasks by building upon old knowledge. Our recent work explored this notion yielding algorithms for learning consecutive tasks. In this talk, we validate our findings through experimental results in disturbance rejection scenarios. We also present a technique to improve performance using task descriptions that allow for model prediction without data i.e., zero-shot learning.

**3. Stochastic Optimization using Parametric Cost Function Approximations**

Raymond Perkins, Princeton University
Co-authors: Warren B. Powell, Princeton University

A widely used heuristic for solving stochastic optimization problems is to use a deterministic rolling horizon procedure, which has been modified to handle uncertainty (e.g. buffer stocks, schedule slack). We formalize this strategy and call it a parametric cost function approximation where parameterized modifications of costs or constraints can be adaptively learned using the principles of policy search. We will also present stochastic algorithms for tuning the policy in a simulator

## ■ TA02

Bowl 1

### Convex Optimization and Monotone Operators

Contributed Session

Chair: Jonathan Eckstein, MSIS and RUTCOR, 100 Rockafeller Road, Room 5145, Piscataway, NJ, United States, 08854

**1. A Proximal Partially-parallel Splitting Method for Separable Convex Programs**

Kai Wang, Singapore, Singapore, 639798
Co-Author: Jitamitra Desai, Singapore, Singapore

We propose a "hybrid" proximal partially-parallel splitting method for solving separable convex minimization problems. A worst-case O(1/k) convergence rate is established for this algorithm under both ergodic and nonergodic conditions. Computational efficiency is also demonstrated by solving several instances of the robust PCA problem.

**2. Conditional Gradient Type Methods for Composite Nonlinear and Stochastic Optimization**

Saeed Ghadimi, School of Mathematics, Tehran, Iran, Islamic Republic of,

We present a conditional gradient (CG) type method for minimizing a weakly smooth (possibly nonconvex) term with a strongly convex regularization. Iteration cost and complexity of this method fall between those of the classic CG and proximal type methods. We also extend this method for stochastic optimization and present a few new complexity results.

**3. Gradient Sliding for Saddle Point Problems**

Yuyuan Ouyang, Clemson, SC, United States,
Co-Author: Guanghui Lan, Gainesville, FL, United States,

We consider a class of nonsmooth optimization problem with saddle point structure. In order to compute an approximate solution to the saddle point problem with a smooth component, we propose novel first-order methods that could skip the computation of the gradient of the smooth component from time to time.

**4. Asynchronous Projective Monotone Operator Splitting Algorithms**

Jonathan Eckstein, Professor, MSIS and RUTCOR, 100 Rockafeller Road, Room 5145, Piscataway, NJ, United States, 0885
Co-Author: Patrick Combettes, Professor, Laboratoire Jacques-Louis Lions, Paris, France,5.

We describe a new family of decomposition algorithms for finding zeroes of sums of maximal monotone operators. Only a subset of the subsystems need to be evaluated between coordination steps, and subproblems and coordination steps may be overlapped asynchronously. Applications include asynchronous ADMM-like convex optimization methods.

## ■ TA03

Bowl 2

### First-order Methods for Optimization under Uncertainty

Contributed Session

Chair: Fatma Kilinc-Karzan, Tepper School of Business, 5000 Forbes Ave, Pittsburgh, PA, United States, 15213

**1. Convex and Stochastic Optimization under Indirect Observations**

Niao He, Champaign, IL, United States

We examine convex optimization problems involved with misspecified parameters or distributions, while only a limited number of indirect observations are available. We establish safe and computationally tractable approximations of these problems and equip them with efficient first order algorithms.

**2. Incremental Methods for Additive Convex Cost Optimization**

Mert Gurbuzbalaban, Cambridge, MA, United States
Co-Authors: Asuman Ozdaglar, 77 Massachusetts Avenue, 32-D630, Cambridge, MA, United States, 02139, Pablo Parrilo, 77 Massachusetts Avenue, Room 32D-726, Cambridge, MA, United States, 02139-4307

Motivated by machine learning problems over large data sets, we consider the problem of minimizing the sum of a large number of convex functions. We develop and study incremental methods for solving such problems, in particular for the random reshuffling method, we provide a sharp convergence rate result which answers a long-standing open question.

**3. Convex Computation of the Reachable Set for Uncertain Hybrid Systems**

Ram Vasudevan, Mechanical Engineering, 1231 Beal Avenue, 204, Automotive Laboratory, Ann Arbor, MI, United States, 48109

This talk presents a convex optimization method to efficiently compute the set of configurations of a polynomial hybrid dynamical system that are able to safely reach a user defined target set despite parametric uncertainty in the model. This class of models describes, for example, legged robots moving over uncertain terrains.

**4. Conditional Gradient Sliding for Convex Optimization**

Yi Zhou, Industrial and Systems Engineering, Atlanta, GA, United States
Co-Author: Guanghui Lan, Atlanta, GA, United States

We present a new conditional gradient type method for convex optimization. The conditional gradient sliding algorithm in the paper skips the computation of gradients from time to time, and as a result, achieves the optimal complexity bounds in terms of not only the number of calls to the linear oracle, but also the number of gradient evaluations.

## ■ TA04

Bowl 16

### New Paradigms for Cut Generation

Contributed Session

Chair: Thiago Serra, 5000 Forbes Avenue, Pittsburgh, PA, United States, 15213

**1. Cutting Planes from Extended LP Formulations**

Merve Bodur, ISyE, Atlanta, GA, United States,

For mixed integer sets, we study extended formulations of their LP relaxations. For any 0-1 mixed integer set with n integer and k continuous variables, we construct an extended formulation with 2n+k-1 variables whose split closure is integral. We extend this to general mixed integer sets and construct the best extended formulation w.r.t. split cuts.

**2. Final Point Generalized Intersection Cuts**

Aleksandr M Kazachkov, 5000 Forbes Avenue, Pittsburgh, PA, United States, 15217
Co-Author: Francois Margot, Tepper School of Business, 5000 Forbes Avenue, Pittsburgh, PA, United States, 15213-3890, Egon Balas, Tepper School of Business Rm 232B, 5000 Forbes Avenue, Pittsburgh, PA, United States, 15213-3890

We introduce a new class of generalized intersection cuts for mixed integer programming called final point cuts. We report on computational experiments that show these cuts to close, on average, more than twice the size of the integrality gap closed by standard intersection cuts, and roughly a third of that closed by the split closure.

**3. On Polyhedral Relaxations for Bilinear Terms in Nonconvex Quadratic Optimization**

Akshay Gupte, Assistant Professor, Mathematical Sciences, O-321 Martin Hall, Clemson, SC, United States, 29634

We study the set $\{(x,y): y_ij = x_i x_j \forall (i,j) \in I \times J, (x_I, x_J) \in P_1 \times P_2 \}$ for disjoint $I,J$ and polytopes $P_1,P_2$. Its convex hull is described when $P_1$ is some simplex in R^m or a rectangle in R^2. For some general structures of $P_1$, a sequential convexification procedure is used to generate linear valid inequalities.

**4. Reformulating the Disjunctive Cut Generating Linear Program**

Thiago Serra, 5000 Forbes Avenue, Pittsburgh, PA, United States, 15213

In conventional lift-and-project, CGLP optima may yield dominated cuts due to distortions caused by normalization. This work proposes the Reverse Polar CGLP, a reformulation where normalization defines separability. Cuts derived from RP-CGLP optima define supporting hyperplanes of the disjunctive hull, hence never being strongly dominated.

## ■ TA05

Room 023

### Logistics

Contributed Session

Chair: Esra Buyuktahtakin, Industrial and Manufacturing Engineering, 1845 N Fairmount, Wichita, KS, 67260-0035, Wichita, KS, United States, 67260

**1. Data Analytics for Process Improvement of Warehouse Operations**

Fatma Gzara, Management sciences, Waterloo, ON, Canada, N2l 3g1
Co-Authors: Ugur Yildiz, Management Sciences, Waterloo, ON, Canada, N2l3g1, Samir Elhedhli, Management Sciences, Waterloo, ON, Canada, N2l 3g1

We report on a data analytics project carried out in partnership with Dematic Ltd, a global leader in warehouse management and control systems, to develop descriptive, predictive and prescriptive tools for warehouse operations. We present data visualization tools, statistical analysis of operations, and optimization-simulation tool.

**2. Increasing Supply Chain Robustness through Process Flexibility and Inventory**

He Wang, Cambridge, MA, United States,
Co-Author: Yehua Wei, 100 Fuqua Drive, Durham, NC, United States, 27713
Co-Author: David Simchi-Levi, Dept of Civil and Environmental Engineering, 77 Massachusetts Avenue Rm 1-171, Cambridge, MA, United States, 02139

We develop a delayed constraint generation algorithm for solving two-stage robust optimization models, in which the second stage recourse problem has some network structure. This algorithm helps us study a supply chain risk mitigation strategy that combines firm's production and inventory decisions.

**3. Optimization In Transportation Planning: Assigning Trips On Routes**

EunSu Lee, Assistant Professor, Management, 160 Harborside Plaza 2, #234H, Jersey City, NJ, United States, 07311

The state of North Dakota has been experiencing significant traffic increase due to recent oil development and changes of agricultural logistics patterns, resulting in needs of heavy capital investment on roads for pavement and maintenance. Based on the quickest routes, the model run 20-year life-cycle traffic assignment on major and local roads.

**4. Dynamic Programming Approximation Algorithms for the Capacitated Lot-sizing Problem**

Esra Buyuktahtakin, Industrial and Manufacturing Engineering, 1845 N Fairmount, Wichita, KS, 67260-0035, Wichita, KS, United States, 67260
Co-Author: Ning Liu, Industrial and Manufacturing Engineering, Wichita, KS, United States,

We present a new idea for approximating the inventory cost function in a truncated dynamic program for solving the capacitated lot-sizing problem. Computational results are promising: our approximation approaches could significantly decrease the computational time required by exact methods, while capturing the optimal solution in most cases.

## ■ TA06

Sherrerd

### Optimization on Networks I

Contributed Session

Chair: John P Dickerson, Computer Science Department, 9219 Gates-Hillman Center, Carnegie Mellon University, Pittsburgh, PA, United States, 15213

**1. Improved Relaxation Analysis for a Dynamic Knapsack Problem**

Alejandro Toriello, Industrial and Systems Engineering, 765 Ferst Dr NW, Atlanta, GA, United States, 30332
Co-Authors: Weihong Hu, Industrial and Systems Engineering, 755 Ferst Dr. NW, Atlanta, GA, United States, 30332, Daniel Blado, Industrial and Systems Engineering, Atlanta, GA, United States,

We consider a knapsack problem in which item sizes are stochastic and realized after an attempted insertion, and the decision maker chooses an item to insert dynamically based on remaining capacity. We analyze a semi-infinite relaxation and provide theoretical and experimental evidence of its tightness.

**2. Weighted Target Set Selection On Social Networks**

Rui Zhang, DOIT, College Park, MD, United States, 20742
Co-Author: S. Raghavan, DOIT, College Park, MD, United States, 20742

We consider the weighted target set selection problem, which arises in the context of influence maximization on social networks. Motivated by the desire for exact approaches, a tight and compact extended formulation is presented on trees. Furthermore, based on the extended formulation, a branch-and-cut approach is proposed for general networks.

**3. Heuristics for the Wireless Backhaul NetworkDesign Problem**

Carlos Eduardo de Andrade, Advanced Technology, 200 S Laurel Avenue, Room A5-1E33, Middletown, NJ, United States, 07748

We present a network design problem used to serve mobile devices and backhaul the traffic to the core network. We develop a biased random-key genetic algorithm to best deploy the equipment. The algorithm is tested on a real world study case in the vicinity of a large city in the United States.

**4. Fast Optimal Clearing of Capped-chain Barter Exchanges**

John P Dickerson, Computer Science Department, 9219 Gates-Hillman Center, Carnegie Mellon University, Pittsburgh, PA, United States, 15213
Co-Author: Tuomas W Sandholm, Gates Center for Computer Science, Carnegie Mellon University, Pittsburgh, PA, United States, 15213

We present new integer programming models for clearing kidney exchanges with small cycles and long (but not infinite) chains—-a constraint that mimics fielded exchanges. We show on real data from two nationwide kidney exchanges as well as generated data that our models scale significantly better than the prior state of the art.

# Thursday, 11:00am – 12:30pm

## ■ TB01

Dodds Auditorium

### Optimization in Machine Learning

Contributed Session

Chair: Yan Li, Operations Research and Financial Engineering, Princeton, NJ, United States, 08540

**1. A Bregman-Lagrangian Perspectiveon Accelerated Gradient Methods**

Wibisono Andre, 387 Soda Hall #1776, Berkeley, CA, United States, 94720-1776
Co-Authors: Michael Jordan, Berkeley, CA, United States, 94720-1776, Ashia Wilson, 5521 Research Park Drive, Suite 200, Berkeley, CA, United States, 94720-1776

What is the natural scope of the acceleration concept in optimization theory? We take a Lagrangian perspective on this problem, defining a functional that we refer to as a "Bregman Lagrangian," and showing that this functional generates a large class of accelerated methods in continuous time.

**2. A Hybrid Interior-point Alternating Direction Algorithm for Knowledge-based SVM and Feature Selection**

Ioannis Akrotirianakis, Senior Scientist, Business Analytics and Monitoring, 755 College Road East, Princeton, NJ, United States, 08540

We consider classification tasks in the regime of scarce labeled data in high dimensional feature space. We propose a new hybrid optimization algorithm that solves the elastic-net SVM through an ADMM in the first phase, followed by an IPM for the classical SVM in the second phase. Numerical results show the power of the new method.

**3. Perturbed Iterate Analysis for Asynchronous Stochastic Optimization**

Horia Mania, Berkeley, CA, United States, 94720
Co-Authors: Dimitris Papailiopoulos, Berkeley, CA, United States, Ben Recht, Berkeley, CA, United States, Xinghao Pan, Berkeley, CA, United States, 94720, Michael I Jordan, Berkeley, CA, United States, Kannan Ramchandran, Berkeley, CA, United States

We introduce and analyze stochastic optimization methods where the input to each gradient update is perturbed by bounded noise. We show that this framework forms the basis of a unified approach to analyze asynchronous implementations of stochastic optimization algorithms.

**4. Optimal Learning On Regularized Trees**

Yan Li, Operations Research and Financial Engineering, Princeton, NJ, United States, 08540
Co-Authors: Warren B Powell, Operations Research and Financial Engineering, 230 Sherrerd Hall, Dept of Operations Research and Financial Eng, Princeton, NJ, United States, 08544, Han Liu, Operations Research and Financial Engineering, 35 Olden Street, Princeton, NJ, United States, 08544

We derive a knowledge gradient policy for noisy global optimization with regularized trees. Besides selecting the best alternative before the budget is exhausted, we also aim to learn the important features. Experimental work shows that the method is scalable to high dimension and efficiently learns the underlying low dimensional structures.

## ■ TB02

Bowl 1

### Tutorial Session: Polynomial Optimization

Contributed Session

Chair: Amir Ali Ahmadi, Dept. of Operations Research & Financial Eng., Sherrerd Hall (Room 329), Charlton Street, Princeton, NJ, United States, 08544

**1. Tutorial: Optimization over Nonnegative Polynomials: Algorithms and Applications**

Amir Ali Ahmadi, Department of Operations Research and FInancial Engineering, Princeton University, Sherrerd Hall (room 329), Charlton Street, Princeton, NJ 08544 United States, a_a_a@princeton.edu

This tutorial is about a problem that is very easy to describe: How to impose conditions on the coefficients of a multivariate polynomial so that it becomes nonneagtive on a set defined by polynomial equations and inequalities? This problem has surprisingly many applications in applied and computational mathematics including polynomial optimization, combinatorial optimization, Lyapunov analysis of dynamical systems, regression and classification in machine learning, formal verification of safety-critical systems, and automated theorem proving in geometry, among many others. I will explain how nonnegative polynomials arise in all these applications and then focus on the computational complexity of the problem and some of the algorithms we currently have for approaching it. In particular, I will cover the theory of sum of squares polynomials and its relation to (i) classical Positivstellensatz results in algebraic geometry, and (ii) modern computational tools involving semidefinite optimization (as shown by Parrilo and others). This tutorial will be followed by a tutorial by Etienne de Klerk who will speak about the Lasserre hierarchy and it connection to moment problems. This will be "dual" (in the convex analysis sense) to my presentation.

Time permitting, I will also touch upon my recent work with Anirudha Majumdar to make the algorithms in this field more scalable by moving from semidefinite programming to linear or second order cone programming. There will be a follow-up session with talks by Anirudha Majumdar and Georgina Hall who will speak further about this direction of research.

The theory of nonnegative polynomials has recently found applications also in robotics, and this will be touched upon in the plenary talk by Russ Tedrake.

**2. Tutorial: On Lasserre-type Hierarchies of Upper Bounds for Bound-constrained Polynomial Optimization**

Etienne de Klerk, Tilburg, Netherlands
Co-author: Amir Ali Ahmadi, Dept. of Operations Research & Financial Eng., Sherrerd Hall (Room 329), Charlton Street, Princeton, NJ, United States, 08544

We consider minimizing a given polynomial f over the hypercube [-1,1]n. An idea of Lasserre is to find a probability distribution on [-1,1]n with polynomial density function of fixed degree that minimizes the expectation of the random variable f(X). We show improved error bounds for this approach.

## ■ TB03

Bowl 2

### Stochastic Mixed-integer Programming Models

Contributed Session

Chair: Saravanan Venkatachalam, Detroit, MI, United States, 48377

**1. A Polyhedral Study of the Static Probabilistic Lot-sizing Problem**

Xiao Liu, PhD Candidate, Integrated Systems Engineering, 1971 Neil Ave Columbus, OH 43210, 320 Baker Systems Building, Columbus, OH, United States, 43210
Co-Author: Simge Küçükyavuz, Associate Professor, Integrated Systems Engineering, 210 Baker Systems Building, 1971 Neil Av, Columbus, OH, United States, 43210

We study the static probabilistic lot-sizing problem. We propose a new formulation based on the simple recourse structure of this problem. We propose two classes of valid inequalities and prove their strength. We discuss several extensions of the proposed methods. The computational results show the effectiveness of the proposed methods.

**2. An Asynchronous Incremental Bundle Method for Dual Decomposition of Stochastic Mixed-integer Programming**

Kibaek Kim, Lemont, IL, United States
Co-Authors: Victor M Zavala, Madison, WI, United States, Cosmin G Petra, 9700 South Cass Avenue, Building 240, Lemont, IL, United States, 60439

We present an incremental bundle method for dual decomposition of stochastic mixed-integer programming, enabling fully asynchronous computations on high performance computing clusters. We have implemented the method in an open-source software DSP and also report numerical results.

**3. Irreducible Infeasible Subsystem Approach for Stochastic Integer Programs with Probabilistic Constraints**

Lewis Ntaimo, Associate Professor, Industrial and Systems Engineering, 3131 TAMU, College Station, TX, United States, 77843
Co-Authors: Julian Gallego, Chicago, IL, United States, 60606, Bernardo Pagnoncelli, Associate Professor, School of Business, Chile, Chile

Probabilistically constrained stochastic integer programs (PC-SIPs) are challenging to solve and linear programming (LP) provides very weak bounds on the optimal value. This work introduces irreducible infeasible subsystem (IIS) inequalities for strengthening the LP-relaxation of PC-SIPs.

**4. Algorithm for Two-stage Stochastic 0-1 Programs with Absolute Semi-deviation Risk Measure**

Saravanan Venkatachalam, Detroit, MI, United States, 48377
Co-Author: Lewis Ntaimo, 3131 TAMU, College Station, TX, United States, 77843

We present a methodology for absolute semi-deviation (ASD) risk-measure for stochastic 0-1 programs. The proposed methodology uses information from expected excess, and uses cutting planes based on sub-gradient information from (ASD) model constraints. Computational results for a supply chain application will be presented.

## ■ TB04

Bowl 16

### Mixed Integer Second Order Conic Optimization

Contributed Session

Chair: Alexander Vinel, Industrial and Systems Engineering, 3308 Shelby Center, Auburn, AL, United States, 36849

**1. Applying Disjunctive Conic and Cylindrical Cuts to Solve Quantitative Asset Allocation Problems**

Sertalp B Cay, PhD Candidate, Industrial and Systems Engineering, 200 W Packer Ave, Bethlehem, PA, United States, 18015
Co-Authors: Tamas Terlaky, Industrial and Systems Engineering, 200 West Packer Avenue, Bethlehem, PA, United States, 18015, Julio Cezar Goez, Bergen, Norway, 5034

The methodology of disjunctive conic and cylindrical cuts (DCC) was developed to solve mixed integer second order cone optimization problems. In this talk, we show that DCCs have significant positive impact when solving realistic asset allocation problems and their impact highly depends on where they are added inside the Branch and Conic Cut solver.

**2. On the Identification of Optimal Partition in Semidefinite Optimization**

Ali Mohammad Nezhad, Industrial and Systems Engineering, 200 W. Packer Ave., Bethlehem, PA, United States, 18015
Co-Author: Tamas Terlaky, 200 West Packer Avenue, Dept ISE, Bethlehem, PA, United States, 18015

We extend the concept of optimal partition to semidefinite optimization. A condition number is defined to set bounds on the eigenvalues of solutions near the central path. Using the optimal partition, we establish a set of conditions for a speedup in the convergence of an interior point method close to optimality.

**3. Generalizing Higher Order Cone Programming Problems and Related Optimization Techniques**

Alexander Vinel, Industrial and Systems Engineering, 3308 Shelby Center, Auburn, AL, United States, 36849

A class of convex nonlinear optimization problems which is a direct generalization of p-order cone programming is studied. We show that some techniques used in cone programming can be naturally extended to this class. We consider finding projections onto the defined convex nonlinear set and dimensionality reduction through lifting.

## ■ TB05

Room 023

### Optimization and Learning in Finance

Contributed Session

Chair: John M Mulvey, ORFE, Sherrerd Hall, Princeton, NJ, United States, 08540

**1. Linear Solution Schemes for Mean Semivariance Project Portfolio Selection Problems**

Onur Babat, PhD Candidate, Industrial and Systems Engineering, 217 West Packer Avenue Apt: 106, Bethlehem, PA, United States, 18015
Co-Author: Luis Zuluaga, Assistant Professor, Industrial and Systems Engineering, Bethlehem, PA, United States

In this paper, we propose two linear solution schemes to solve the Mean-SemiVariance Project (MSVP) portfolio selection problem; that is, the proposed schemes avoid the use of MIQP solvers and only require the use of MILP techniques. We illustrate our solution schemes by solving a real MSVP problem arising in a Latin American oil and gas company.

**2. Outperforming with Machine Learning Algorithms**

Abhishek Saxena, Quantitative Research Analyst, Quantitative Research, 3301 C Street, Suite 500, Anchorage, AK, United States, 99503

We use CART and AdaBoost in Mean-Variance and CVaR optimization setups to construct portfolios with consistent and statistically significant outperformance over the market. We also introduce a new machine learning algorithm based on neighborhood selection in all possible Euclidean spaces of a fixed dimensionality from a given selection of factors.

**3. Resource Allocation under Ambiguous Risk Preferences**

Chaithanya Bandi, Managerial Economics and Decision Sciences, Evanston, IL, United States

Modern resource allocation problems have three features: 1) Large Scale Nature, (2) Real time or instantaneous decision making;(3) Ambiguous Risk preferences. In this talk, we consider such high dimensional problems and present closed form insightful solutions. Joint work with Paat Rusmevichientong.

**4. Linking Machine Learning and Optimal Decisions: Applications in Financial Planning**

John M Mulvey, Professor, ORFE, Sherrerd Hall, Princeton, NJ, United States, 08540

Machine learning methods are largely ill suited to discovering optimal investment strategies for investors. Classification algorithms ignore portfolio elements, especially dynamic policies in the face of nonlinear risk aversion. We integrate the ML algorithm with a decision model and illustrate the approach via a university endowment.

## ■ TB06

Sherrerd

### Optimal Learning

Contributed Session

Chair: Jialei Wang, Operations Research, Ithaca, NY, United States, 14853

**1. Derivative-free Optimization of Expensive Noisy Functions using Stratified Bayesian Optimization**

Saul Toscano-Palmerin, 134 Summerhill Drive Apartment 2, Ithaca, NY, United States, 14850
Co-Author: Peter Frazier, ORIE, 232 Rhodes Hall, Cornell University, Ithaca, NY, United States, 14853

We consider simulation optimization, and noisy derivative-free optimization of expensive functions, when most of the randomness in the objective is produced by a few scalar random inputs. We present a new global optimization algorithm, which uses this strong dependence to improve performance.

**2. Optimal Learning with Discrete Priors**

Weidong Han, Operations Research and Financial Engineering, Princeton, NJ, United States, 08544

We consider an optimal learning problem with a random parameter following a discrete prior distribution. We formulate the problem into a dynamic program, and propose a novel lookahead policy called the OC-correction. Empirical experiments show that the proposed policy outperforms many existing approximation algorithms.

**3. Regret Analysis of the Finite Horizon Gittins Index Strategy for Multi-armed Bandits**

Tor Lattimore, Department of Computing Science, Edmonton, AB, Canada, T6E2J1

I prove near-optimal frequentist regret guarantees for the finite-horizon Gittins index strategy for multi-armed bandits. The new results explicate the exceptional performance of the Gittins index strategy and provide new insights about both Bayesian optimal strategies and the design of frequentist algorithms for multi-armed bandits.

**4. Multifidelity Optimization Exploiting Correlation Across Information Sources**

Jialei Wang, Operations Research, Ithaca, NY, United States, 14853
Co-Authors: Peter Frazier, 232 Rhodes Hall, Cornell University, Ithaca, NY, United States, 14853, Matthias Poloczek, 37 Uptown Road, Apt 1D, Ithaca, NY, United States, 14850

We consider multifidelity optimization, where multiple information sources for evaluating the objective function with varying fidelity and cost are available and correlated. Our approach selects an information source to sample adaptively in a way of maximizing information gain from each sample.

# Thursday, 1:30 PM - 3:00 PM

## ■ TC01

Dodds Auditorium

### Computational Optimization in Learning – I

Contributed Session

Chair: Somayeh Moazeni, Assistant Professor, 1 Castle Point Terrace on Hudson, Hoboken, NJ, United States, 07030Distributed Energy Storage Locations For Grid Reliability Maximization

**1. Distributed Energy Storage Locations for Grid Reliability Maximization**

Ricardo A Collado, Assistant Professor, School of Business, Hoboken, NJ, United States,
Co-Author: Somayeh Moazeni, Assistant Professor, 1 Castle Point Terrace on Hudson, Hoboken, NJ, United States, 07030

We propose a stochastic programming approach to maximize system reliability under uncertainty in tie-lines due to natural disasters. Our method optimally places energy storage devices to aid the network flow in case of natural disasters. We propose robust and risk-averse solutions for the problem as well as solution methods based on cutting planes.

**2. Distributed Optimization with Arbitrary Local Solvers**

Martin Takac, H.S. Mohler Laboratory, 200 West Packer Avenue, Bethlehem, PA, United States, 18015
Co-Authors: Michael I. Jordan, Berkeley, CA, United States, Martin Jaggi, Universitaetsstr 6, Zurich, Switzerland, 8092, Peter Richtarik, 6317 James Clerk Maxwell Building, Kings Buildings, Edinburgh, United Kingdom, EH93JZ, Jakub Konecny, Edinburgh, United Kingdom, Chenxin Ma, Bethlehem, PA, United States, Virginia Smith, Berkeley, CA, United States

We present a framework for distributed optimization that allows a flexibility of arbitrary solvers to be used on each node locally and yet maintains competitive performance against other state-of-the-art special-purpose distributed methods. We give strong primal-dual convergence rate guarantees for our framework that hold for arbitrary solvers.

**3. L1 Regularized Least Squares for Support Recovery of High Dimensional Single Index Models with Gaussian Designs**

Matey Neykov, ORFE, Princeton, NJ, United States, 08544
Co-Authors: Jun S. Liu, Professor, Statistics, Cambridge, MA, United States, 02138, Tianxi Cai, Professor, Biostatistics, Boston, MA, United States, 02115

In this paper we demonstrate, that LASSO succeeds optimally up to a scalar in recovering the signed support of 0 under a single index model Y=f(X 0, ) if X N(0, ), and the covariance satisfies the irrepresentable condition. Our work extends existing results on the support recovery of LASSO for the linear model, to a certain class of SIM.

**4. Non-Stationary Parametric Cost Function Approximation for Risk Averse Stochastic Optimization**

Somayeh Moazeni, Assistant Professor, Hoboken, NJ, United States, 07036
Co-Authors: Boris Defourny, Industrial and Systems Engineering, 200 W Packer Ave, Bethlehem, PA, United States, 18015, Warren B Powell, 230 Sherrerd Hall, Dept of Operations Research and Finncial Eng, Princeton, NJ, United States, 08544

We present an algorithmic strategy to non-stationary policy search for MDPs with constrained action sets and a risk-sensitive optimality criterion. The method relies on modeling time-variant policy parameters by an interpolating surface model for parametric cost function approximation policies. It is illustrated by an energy storage problem.

## ■ TC02

Bowl 1

### Convex Relaxations for Nonconvex Polynomial Programs

Contributed Session

Chair: Jeffrey Zhang, Princeton, NJ, United States, 08540

**1. DC Decomposition of Nonconvex Polynomials with Algebraic Techniques**

Georgina Hall, Operations Research and Financial Engineering, Princeton, NJ, United States, 08540
Co-Author: Amir Ali Ahmadi, Dept. of Operations Research&Financial Eng., Sherrerd Hall (Room 329), Charlton Street, Princeton, NJ, United States, 08544

The concave-convex procedure is a majorization-minimization algorithm for difference of convex (DC) optimization where the constraints and objective function are given as the difference of two convex functions. In this talk, we focus on polynomial optimization and introduce LP, SOCP and SDP-based algorithms for finding optimal DC decompositions.

**2. Optimization Techniques for Controlling Highly Agile Robots with Formal Safety Guarantees**

Anirudha Majumdar, Cambridge, MA, United States, 02139
Co-Author: Russ Tedrake, Cambridge, MA, United States

This talk will present work on using tools from convex optimization such as sums-of-squares (SOS) programming for pushing highly agile robotic systems to the brink of their hardware limits while ensuring that they operate in a safe manner despite uncertainty in the environment and dynamics.

**3. Semidefinite Programming for Bimatrix Games**

Jeffrey Zhang, Princeton, NJ, United States, 08540
Co-Author: Amir Ali Ahmadi, Sherrerd Hall 329, Princeton, NJ, United States, 08540

We approach the problem of finding an approximate Nash equilibrium for bimatrix games with semidefinite programming relaxations. We propose and analyze the performance of a series of algorithms from both a theoretical and computational perspective.

## ■ TC03

Bowl 2

### Stochastic Programming Approaches in Advanced Statistical Applications

Contributed Session

Chair: Stan Uryasev, AOD, 303 Weilhall, Gainesville, FL, United States, 32611

**1. Identification of Effective Scenarios in Distributionally Robust Stochastic Programs with Variation Distance**

Hamed Rahimian, 1971 Neil Avenue, 210 Baker Systems Building, Columbus, OH, United States, 43210
Co-Authors: Tito Homem-de-Mello, School of Business, Diagonal Las Torres 2640, Santiago, Chile, Guzin Bayraksan, Columbus, OH, United States, Guzin Bayraksan, Columbus, OH, United States,

Instead of a known distribution, an uncertainty set of distributions is constructed using the variation distance. We examine the model properties and propose a decomposition-based algorithm to solve it. We characterize a minimal scenario tree, where the presence of every scenario is critical in determining the optimal objective function.

**2. Modeling with Multistage Stochastic Programming**

Alan King, Yorktown Heights, NY, United States

The modeling and solution of multistage stochastic programs will be presented. Production and inventory management, and financial portfolio optimization are use cases. The uncertainty is a hidden Markov Chain evolution of the distribution parameters. Stochastic dual dynamic programming will be presented as a solution algorithm.

**3. Inverse Optimization with Noisy Data**

Auyon Siddiq, Berkeley, CA, United States
Co-Authors: Zuo-Jun (Max) Shen, Berkeley, CA, United States, Anil Aswani, Berkeley, CA, United States,

We present an approach for inverse optimization with noisy solution data. The estimates attain risk and parameter estimation consistency under reasonable conditions. While the formulation is non-convex, we present an approximation algorithm that yields consistent estimates. Numerical performance is competitive with existing techniques.

## ■ TC04

Bowl 16

### First-order Methods for Convex Optimization

Contributed Session

Chair: Angelia Nedich, University of Illinois, Urbana, IL, Contact: angelia@illinois.edu

**1. Decentralized Directed Formation Control Based on First-order Methods**

Wei Shi, Postdoctoral Research Associate, Industrial and Enterprise Systems Engineering, 1308 W Main St, 128 Coordinated Science Lab, Urbana, IL, United States, 61801
Co-Authors: Alex Olshevsky, Assistant Professor, Industrial and Enterprise Systems Engineering, 1308 W Main Street, 145 Coordinated Science Lab, Urbana, IL, United States, 6180, Angelia Nedich, Associate Professor, Industrial and Enterprise Systems Engineering, 104 S Mathews Ave, 201D Transportation Building, Urbana, IL, United States, 61801

We design decentralized protocols for directed multi-agent formation control. To cope with asymmetric information flow, the system is modeled as a convex/nonconvex game. (Stochastic) First-order methods are applied. We show the convex system converges linearly. Simulation shows the nonconvex system achieves a global optimizer in most trials.

**2. Distributed Linearized Alternating Direction Method of Multipliers for Composite Convex Consensus Optimization**

Necdet Serhat Aybat, Industrial Engineering, Dept. of Industrial Engineering, University Park, PA, United States, 16802

We propose distributed proximal (stochastic) gradient methods to compute a consensual decision among agents that minimizes the sum of agent-specific private convex composite functions. We show rate results for ergodic convergence in both sub-optimality error and consensus violation, and examine the effect of the underlying topology.

**3. On Stochastic Mirror Descent Algorithms: Self-tuned Stepsizes, and Optimal Averaging Schemes for L1 Regularized Loss Minimization**

Farzad Yousefian, Industrial and Engineering Management, Industrial Eng. & Managemt, 317D Engineering North, Stillwater, OK, United States, 74078
Co-Author: Arash Pourhabib, Assistant Professor, Industrial and Engineering Management, 322 Engineering North, Stillwater, OK, United States, 74078

We first develop a self-tuned stepsize rule for the stochastic mirror descent (SMD) algorithm under which the optimal convergence is attained and an upper bound of MSE is minimized. Second, we develop an optimal weighted averaging scheme with increasing weights for the truncated SMD algorithm for solving L1 regularized loss minimization problems.

**4. On Convergence Properties of Fast Nesterov Method**

Angelia Nedich, 12, University of Illinois, Urbana, IL, United States

We present some new results regarding the convergence properties of the iterates generated by the fast Nesterov algorithm for a general convex optimization problem involving the sum of a nonsmooth convex function and a smooth convex function. We show that both iterate sequences converge to the same solution for a generic choice of overshoot parameters.

## ■ TC05

Room 023

### Stochastic Optimization in Energy

Contributed Session

Chair: Stephan Meisel, University of Muenster, Muenster, Germany

**1. An Approximation Algorithm to Generate Future Wind Power Scenarios**

Bita Analui, Ira A. Fulton Schools of Engineering, Tempe, AZ, United States, 85287
Co-Author: Anna Scaglione, Professor, Ira A. Fulton Schools of Engineering, Tempe, AZ, United States

In this work, we propose a compression and quantization scheme to approximate wind power trajectories in a finite subspace of wavelet basis and use stochastic approximation algorithm to directly quantize the realizations of the wavelet coefficients, while retaining key features of the high order statistics, over the optimization horizon.

**2. Distributed Energy Storage Locations for Grid Reliability Maximization**

Ricardo A Collado, Assistant Professor, School of Business, Hoboken, NJ, United States,
Co-Author: Somayeh Moazeni, Assistant Professor, 1 Castle Point Terrace on Hudson, Hoboken, NJ, United States, 07030

We propose a stochastic programming approach to maximize system reliability under uncertainty in tie-lines due to natural disasters. Our method optimally places energy storage devices to aid the network flow in case of natural disasters. We propose robust and risk-averse solutions for the problem as well as solution methods based on cutting planes.

**3. Robust Lot Sizing and its Application to Heat-and-power Co-generation**

Arie Koster, Professor of Discrete Optimization, Lehrstuhl II für Mathematik, Pontdriesch 14-16, Aachen, Germany, 52062
Co-Authors: Stefano Coniglio, Lehrstuhl II für Mathematik, Aachen, Germany, 52062, Nils Spiekermann, Ph.D. student, Lehrstuhl II für Mathematik, Aachen, Germany, D-52056, Alexander Hein, Aachen, Germany, Olaf Syben, Aachen, Germany

We study lot sizing problems with storage deterioration arising in the day-ahead planning of combined heat-and-power plants. We show that the lot sizing problem under (heat) demand uncertainty can be polynomially reduced to a single static lot-sizing problem, provided optimization over the uncertainty set can be done in polynomial time.

**4. Learning Policies for Energy Storage Problems**

Stephan Meisel, Muenster, Germany
Co-Author: Warren B Powell, 230 Sherrerd Hall, Dept of Operations Research and Financial Eng, Princeton, NJ, United States, 08544

Four elementary classes of policies exist for solving sequential decision problems. The performance of a policy depends on the problem instance to be solved. Some instances require hybrid policies. We propose an approach to learning the best policy for instances of an energy storage problem.

## ■ TC06

Sherrerd

### Tutorial Session: Online Learning for Revenue Management

Contributed Session

Chair: David Simchi-Levi, 77 Massachusetts Avenue, Cambridge, MA, United States

**1. Tutorial: Applications of Machine Learning and Optimization in Online Revenue Management**

David Simchi-Levi, Professor of Engineering Systems, MIT, Cambridge, MA, United States, dslevi@mit.edu

In a dynamic pricing problem where the demand function is unknown a priori, price experimentation can be used for demand learning. In practice, however, online sellers are faced with a few business constraints, including the inability to conduct extensive experimentation, limited inventory and high demand uncertainty. In this talk we discuss models and algorithms that combine machine learning and price optimization that significantly improve revenue. Specifically, we start by considering a dynamic pricing model where the demand function is unknown but belongs to a known finite set. The seller is allowed to make limited number of price changes during a finite time horizon. The objective is to minimize the regret, i.e. the expected total revenue loss compared to a clairvoyant who knows the demand distribution in advance. We demonstrate a pricing policy that incurs the smallest possible regret, up to a constant factor. In the second part of the presentation we extend the model to a network revenue management problem where an online retailer aims to maximize revenue from multiple products with limited inventory. As common in practice, the retailer does not know the expected demand at each price and must learn the demand information from sales data. We propose an efficient and effective dynamic pricing algorithm, which builds upon the Thompson sampling algorithm used for multi-armed bandit problems by incorporating inventory constraints into the pricing decisions. The algorithm proves to have both strong theoretical performance guarantees as well as promising numerical performance results when compared to other algorithms developed for the same setting. Throughout the presentation, we report results from live implementations at companies such as Rue La La, Groupon and a large European Airline carrier.

**2. One by One Solicitation by Approximate Dynamic Programming for Online Advertising**

Dmitri I Arkhipov, Computer Science, Office 4081, 6210 Donald Bren Hall, Irvine, CA, United States, 92603
Co-Authors: John G Turner, The Paul Merage School of Business, Room SB2 338, Irvine, CA, United States, 92697-3125, Amelia C Regan, Bren Hall 4068, Department of Computer Science, Irvine, CA, United States, 92697-3435, Michael B Dillencourt, Professor, Computer Science, Office 4086, 6210 Donald Bren Hall, Irvine, CA, United States, 92697-3425

**3. The Knowledge Gradient with Logistic Belief Models for Binary Classification**

Yingfei Wang, Computer Science, Computer Science Dept., Princeton University, Princeton, NJ, United States, 08540
Co-Author: Warren B Powell, 230 Sherrerd Hall, Dept of Operations Research and Financial Eng, Princeton, NJ, United States, 08544

We consider sequential decision making problems for binary classifications with the goal to identify the best alternative with the highest response. We use Bayesian logistic regression to predict the label. By formulating the problem as a MDP, we develop a knowledge-gradient policy. Experiments demonstrate the effectiveness of the proposed method.

# Thursday 3:00pm – 4:00pm

## ■ TD01

Dodds Auditorium

### Plenary: Optimization Challenges in Electric Power Markets – Richard O'Neill

Federal Energy Regulatory Commission

# Thursday 4:00pm – 5:30pm

## ■ TE01

Dodds Auditorium

### Machine Learning and Optimization – I

Contributed Session

Chair: Gah-Yi Vahn, Management Science & Operations, London, United Kingdom

**1. Bayesian Dynamic Learning and Pricing with Strategic Customers**

Xi Chen, Stern School of Business, 1 Washington Square Village, Apt 4H, New York, NY, United States, 10012
Co-Author: Zizhuo Wang, 1009 5th Street SE, Minneapolis, MN, United States, 55414

We study the dynamic learning and pricing problem when the customer behave strategically in making his purchase decision. By using a Bayesian setting, we first show that a naive myopic Bayesian policy (MBP) by the seller may lead to incomplete learning. To resolve this issue, we propose a randomized Bayesian policy (RBP).

**2. On Predicting Home Telemonitoring Patients' Adverse Events**

Andrea Lodi, MAGI, Montréal, QC, Canada
Co-Authors: Louis-Martin Rousseau, MAGI, Montréal, QC, Canada, Jonathan Vallee, Montréal, QC, Canada, Nicolas Chapados, MAGI, Montréal, QC, Canada

We review home healthcare agency's challenges addressable by optimization and machine learning. This includes a risk prediction framework for home telemonitoring patients where we use maxout neural networks to predict patient's adverse events. Extended tests show dramatic improvements over logistic regression and manually engineered alerts.

**3. Optimization and Learning for Sequential Decision Making**

Shipra Agrawal, Riverside, New York, NY, United States, 10027

In this talk, I will present techniques that combine optimization and learning for decision making in complex, uncertain, online environments. Much of this work is motivated by challenges faced in modern revenue management problems, namely, a) unknown or difficult to estimate demand distributions, b) multiple complex nonlinear constraints and objectives, c) the need for fast large-scale algorithms, and d) personalized decisions. Formulating these problem aspects into an "online stochastic convex programming" framework, we devise fast algorithms that combine primal-dual paradigm with online learning to achieve provably optimal performance bounds. When applied to the special case of online packing, our ideas yield simpler and faster algorithms with optimal competitive ratio for this widely studied problem.

**4. Performance-based Regularization in Data-driven Portfolio Optimization**

Gah-Yi Vahn, Management Science & Operations, London, United Kingdom
Co-Authors: Andrew Lim, Professor, Decision Science, Singapore, Singapore, Noureddine El Karoui, Associate Professor, Statistics, Berkeley, CA, United States

We investigate performance-based regularization (PBR), on the portfolio optimization problem. The idea is to constrain the sample variances of the estimated quantities in the problem; for portfolio optimization they are the portfolio risk and return. We consider PBR for mean-variance and mean-CVaR portfolio optimization problems.

## ■ TE02

Bowl 1

### Games and Distributed Optimization

Contributed Session

Chair: Andrew L Liu, Industrial Engineering, West Lafayette, IN, United States, 47906

**1. Tropical Optimization Problems. Solution Techniques and Application Examples**

Nikolai Krivulin, Professor, Faculty of Mathematics and Mechanics, 28 Universitetsky Ave., St. Petersburg, Russian Federation, 198504

We consider unconstrained and constrained problems formulated in the framework of tropical algebra to minimize nonlinear functions defined on vectors over an idempotent semifield. We give a brief overview of known problems and existing solution methods. Applications in project scheduling, location analysis and decision making are discussed.

**2. Witsenhausen's Counterexample and Learning In Teams of Agents**

Abhishek Gupta, Electrical and Computer Engineering, 205 Dreese Laboratories, 2015 Neil Avenue, Columbus, OH, United States, 43210

Consider a team, in which each agent observes a random variable and then takes an action. The team acts to minimize a common cost function. A novel approach to establish the existence of a team-optimal solution will be discussed. Further, the challenges in devising a learning algorithm to compute an optimal solution will also be discussed.

**3. Benders Decomposition for the Optimum Communication Spanning Tree Problem**

Ivan Contreras, Concordia University, Montreal, QC, Canada;.icontrer@encs.concordia.ca
Co-authors: Carlos Luna-Mota, Elena Fernández, Technical University of Catalonia, Barcelona, Spain

In this talk we present an exact algorithm based on a Benders decomposition of an arc-based formulation for the optimum communication spanning tree problem. The standard algorithm is enhanced through the use of a Benders-branch-and-cut scheme and the generation of strong optimality cuts. Computational experiments are reported.

**4. Bridging Game Theory and Large Scale Optimization through Potential Games**

Andrew L Liu, Assistant Professor, Industrial Engineering, West Lafayette, IN, United States, 47906

We showed that under player-wise convexity, an equilibrium of a potential game is equivalent to a stationary point of the potential function optimization problem, and designed convergent sequential and parallel algorithms to solve both stochastic large-scale potential games and block-structured optimization problems.

## ■ TE03

Bowl 2

### Robust Optimization

Contributed Session

Chair: Ekaterina Kostina, Institute for Applied Mathematics, Im Neuenheimer Feld 293, Heidelberg, Germany, 69120

**1. Distributionally Robust Inventory Control When Demand Is a Martingale**

Linwei Xin, Industrial and Enterprise Systems Engineering, 104 S. Mathews Ave., Urbana, IL, United States, 61801
Co-Author: David Goldberg, 755 Ferst Drive, Atlanta, GA, United States, 30332-0205

Independence of random demands across periods is typically assumed in multi-period inventory models. We consider a distributionally robust model in which the sequence of demands must take the form of a martingale with given mean and support. We explicitly solve this problem. We also compare to the setting in which demand is independent across periods.

**2. Robust Optimal Feedback for Optimal Control Problems under Uncertainties**

Ekaterina Kostina, Institute for Applied Mathematics, Im Neuenheimer Feld 293, Heidelberg, Germany, 69120

We are interested in computation of worst-case feedback optimal controls. Such problems have a very high degree of computational complexity, and their solution is often too expensive or slow for
complex problems, especially optimal control in real-time. Our aim is development of quick approximate methods and their efficient implementation.

## TE04

Bowl 16

### Tutorial: Discrete Stochastic Optimization

Contributed Session

Chair: Ted Ralphs, 200 W. Packer Avenue, Bethlehem, PA, United States,

**1. Tutorial: Discrete Stochastic Optimization**

Ted Ralphs, Lehigh University, Bethlehem, PA, United States, ted@lehigh.edu

In a traditional optimization modeling framework, the class of optimization problems one can consider is restricted to those in which a fixed decision must be made by a single decision-maker (DM) at a fixed point in time with a single objective. Such models assume implicitly that the DM has deterministic knowledge of all inputs to the model and that there is no possibility for "recourse" to counterbalance the possible adverse effects of unpredicted future events on the effectiveness of the determined strategy. There are a wide variety of generalizations of this basic framework that account for additional features present in many real-world optimization problems. Multilevel/multistage optimization is a framework that allows both for multiple (possibly competing) DMs and for decisions made at at multiple points in time. The framework subsumes both game theoretic models, in which multiple DMs with competing objectives make decisions sequentially, and recourse models, in which a single DM must make a sequence of decisions over time in order to react to changing conditions. The uncertainty resulting from a competitive decision-making environment and the uncertainty arising from exogenous sources of stochasticity can be handled in a unified way, both from the perspective of modeling and from the perspective of algorithm development.

In the first part of this tutorial session, we present a general framework for modeling and solution of this broad class of problems. The overview will consist primarily of a high-level survey of solution methodology, but the focus will not be on details of particular algorithms. Rather, we focus on development of a conceptual framework within which one can understand the theoretical basis for existing solution approaches in terms of the ingredients necessary for development of convergent algorithms.

In the second part of the session, we'll introduce a software framework implementing the algorithmic components introduced in the first part. In particular, we'll describe details of two software packages: SYMPHONY and MibS. MibS is an overarching framework for solution of bilevel integer linear programs. SYMPHONY is an MILP solver that has been extended to include some unique capabilities that allow it to be used effectively as a sub-solver within MibS, including warm-starting, exporting of dual functions, and solution of continuous bilevel integer linear programs. Computational results will be presented.

**2. Bilevel Optimization and the SYMPHONY MILP Solver**

Suresh Bolusani, Lehigh University, H.S. Mohler Laboratory, 200 West, Bethlehem, PA, 18015
Co-Authors: Ted Ralphs, Lehigh University, Bethlehem, PA, 18015), Ashutosh Mahajan, IIT Bombay, Menal Guzelsoy, SAS

In this talk, we describe recent advances in SYMPHONY, an open-source solver framework for mixed integer linear optimization, now with support both for the solution of (continuous) bilevel linear optimization problems and with other extended capabilities that support its use as a subsolver within the MibS framework for solving bilevel integer linear optimization problems. Computational results will be presented.

**3. The MibS Solver for Bilevel Integer Optimization Problems**

Sahar Tahernajad, Lehigh University · H.S. Mohler Laboratory, 200 West Packer Avenue, Bethlehem, PA, 18015
Co-Authors: Ted Ralphs, Lehigh University, Bethlehem, PA, 18015, Scott DeNegre, Hospital for Special Surgery

We discuss an algorithmic framework that extends the branch-and-cut framework of Denegre and Ralphs for solving bilevel integer linear optimization problems. The framework includes a variety of cut generation techniques and other algorithmic components, such as branching strategies and primal heuristics. We discuss the implementation of these components within the open source solver framework MibS, built with modules available in COIN-OR repository. Computational results will be presented.

## TE05

Room 023

### Approximate Dynamic Programming for Managing Energy Operations

Contributed Session

Chair: Selvaprabu Nadarajah, Chicago, IL, United States,

**1. A Multi-scale Decision Rule Approach for Multi-market Multi-reservoir Management**

Napat Rujeerapaiboon, Risk Analytics and Optimization Chair, Lausanne, Switzerland,
Co-Authors: Daniel Kuhn, EPFL CDM MTEI RAO, ODY 1 01 B (Odyssea), Lausanne, Switzerland, 1015, Wolfram Wiesemann, South Kensington Campus, London, United Kingdom, SW7 2AZ

Peak/off-peak spreads on electricity spot market are eroding due to nuclear phaseout and growth in photovoltaics. The shrinking peak/off-peak arbitrage thus forces hydropower plants to participate in the balancing market. We propose a two-layer stochastic model for the operation of hydropower plants engaging in both markets simultaneously.

**2. Approximate Dynamic Programming for Dynamic Quantile-based Risk Measures**

Daniel R. Jiang, Operations Research and Financial Engineering, 8 Lawrence Dr #104, Princeton, NJ, United States, 08540
Co-Author: Warren B Powell, 230 Sherrerd Hall, Dept of Operations Research and Financial Eng, Princeton, NJ, United States, 08544

We propose a new approximate dynamic programming algorithm and a companion sampling procedure to solve risk-averse Markov decision processes under a class of dynamic quantile-based risk measures. Convergence results are proven and an application to energy storage is shown.

**3. Dynamic Programming with Inaccurate System State Data**

Zheng Wang, School of Automation, Nanjing, China

Dynamic programming (DP) with inaccurate system state data is an issue of great significance and challenge in optimal control. To solve the problem, we develop two methods, i.e., robust optimization-based DP and state probability-based DP, and apply them to inventory control under inaccurate record.

**4. Merchant Energy Trading in a Network**

Selvaprabu Nadarajah, Chicago, IL, United States
Co-Author: Nicola Secomandi, Tepper School of Business, 5000 Forbes Avenue, Pittsburgh, PA, United States, 15213

We formulate the merchant trading of energy in a network of storage and transport assets as a Markov decision process with uncertain energy prices. We develop tractable heuristics and both lower and upper bounds on the optimal policy value, achieving tractability by applying linear optimization in novel ways for approximate dynamic programming.

## TE06

Sherrerd

### Optimization Tools for Extracting Hidden Structure in Networks

Contributed Session

Chair: Emmanuel Abbe, Princeton University, Princeton, NJ, United States, 08540

**1. Strong SOCP Relaxations for the Optimal Power Flow Problem**

Santanu Dey, Georgia Institute of Technology, Atlanta, GA 30080 United States, santanu.dey@isye.gatech.edu

We propose three strong SOCP relaxations for the AC OPF problem. These relaxations are incomparable to each other and two of them are incomparable to the SDP relaxation. Computational studies show that these relaxations have several advantages over existing convex relaxations and provide a practical way to obtain feasible solutions with good quality

**2. Finding Planted Subgraphs using the Schur Horn Relaxation**

Venkat Chandrasekaran, Pasadena, CA, United States, 90068

Extracting structured planted subgraphs from large graphs is a fundamental question that arises in a range of application domains. We describe a tractable approach based on convex optimization for this problem. Our method relies on semidefinite descriptions of majorization inequalities on the spectrum of a matrix. (Joint work with Utkan Candogan.)

**3. The Vertex Separator Problem**

Renata Sotirov, Tilburg, Netherlands
Co-Author: Franz Rendl, Institut fuer Mathematik, Klagenfurt, Australia, A-9020

The vertex separator problem for a graph is to find a subset of vertices (called vertex separator) whose removal disconnects the graph into two components. In this talk we present a new class of semidefinite relaxations for the related graph partition problem and show how to exploit those SDP bounds to obtain bounds on the size of the vertex separators.

**4. Information-Computation Tradeoffs in Community Detection**

Emmanuel Abbe, Princeton, NJ, United States, 08540

The stochastic block model (SBM) has recently gathered major attention as a canonical model for community detection. This talk overviews recent results establishing the thresholds for partial and exact recovery, with focus on the presence/absence of information-computation gaps.

# Friday, 8:30am – 10:00am

## ■ FA01

Dodds Auditorium

### Asynchronous Parallel Optimization and Learning

Contributed Session

Chair: Qihang Liu, Rutgers University, 100 Rockafeller Road, Piscataway, New Jersey

**1. DSCOVR: A Randomized asynchronous Algorithm for Distributed Learning with Parameter Server**

Adams Wei Yu, Machine Learning Department, 5722 Beacon Street, Pittsburgh, PA, United States, 15217

Parameter server is a computation framework that could handle big model learning with big data. We propose an algorithm on the this framework, which exploits the double partitions in both data and model to gain parallelism, and applies periodic variance reduction to achieve linear convergence.

**2. Parallel Asynchronous Stochastic Coordinate Descent with Auxiliary Variables**

Cho-Jui Hsieh, Assistant Professor, Computer Science and Statistics, Davis, CA, United States, 95616
Co-Authors: Inderjit Dhillon, Computer Science, Austin, TX, United States, Hsiang-Fu Yu, Computer Science, AUSTIN, TX, United States

Most popular stochastic coordinate descent (SCD) algorithms require a set of auxiliary variables to be maintained to facilitate efficient computation. We study a family of asynchronous parallel SCD algorithms with auxiliary variables. We provide a theoretical guarantee and show the algorithms are much faster than existing solvers.

**3. Asynchronous Parallel Stochastic Gradient for Nonconvex Optimization**

Ji Liu, Computer Science, Rochester, NY, United States, 14627

This paper studies two asynchronous parallel algorithms: one is over a computer network and the other is on a shared memory system. We establish the convergence rate and the linear speedup property for both algorithms. Our results generalize and improve existing analysis for convex minimization.

**4. Object-parallel Solution of Large-scale Lasso Problems**

Gyorgy Matyasfalvi, PhD Student, MSIS, 100 Rockafeller Road, Piscataway, NJ, United States, 08854
Co-Author: Jonathan Eckstein, 100 Rockafeller Road, Room 5145, Piscataway, NJ, United States, 08854

We describe an "object-parallel" C++ approach to implementing first-order optimization methods. As an example application, we solve large-scale Lasso problems on a distributed-memory supercomputer with the spectral projected gradient (SPG) method. We can efficiently accommodate highly unbalanced sparsity patterns.

## ■ FA02

Bowl 1

### Tutorial Session: Nonlinear Programming Challenges in the Optimal Power Flow Problem

Contributed Session

Chair: Javad Lavaei, Columbia, NY, United States

**1. Tutorial: Near-Global Solutions of Nonlinear Mixed-Integer Power Optimization Problems: Theory, Numerical Algorithm, and Case Studies**

Javad Lavaei, Columbia University, New York, NY, United States, lavaei@ee.columbia.edu

The planning and operation of power networks depend heavily on various large-scale optimization problems solved from every few minutes to every several months. Some of these power optimization problems are: state estimation, optimal power flow (OPF), security-constrained OPF, unit commitment, and network reconfiguration. It is very challenging to solve these problems efficiently due to the nonlinearities imposed by two different sources: (i) discrete variables such as the ratio of a tap-changing transformer, the on/off status of a line switch, or the commitment parameter of a generator, (ii) laws of physics. Since 1962, the non-linearity of the OPF-based problems has been studied extensively, and various heuristic and local-search algorithms have been proposed. Recently, there has been a growing interest in using semidefinite programming (SDP) relaxations for finding global or near-global solutions of large-scale non-convex power problems. In this talk, we present several results on this topic.

The existence of a rank-1 matrix solution to the SDP relaxation of a given non-linear mixed-integer power optimization problem enables the recovery of a global solution of the original problem. We prove that the relaxation is exact for a large class of problems due to the physics of power grids. In the case where the relaxation is inexact, we show that since real-world power networks have low treewidth, the SDP relaxation always has a low-rank solution. By leveraging this property, we design a penalized SDP relaxation to reduce the rank to 1, leading to a near-global solution. As a case study, we demonstrate that the proposed relaxation is able to find feasible solutions for the optimal power flow problem for various IEEE and Polish test systems with a global optimality guarantee of at least 99%. Moreover, we propose a computationally-cheap distributed algorithm to solve the penalized SDP relaxation at scale. Unlike the existing second-order methods that cannot handle large SDP problems, this approach leverages the low treewidth of power grids and is able to solve problems with over 30 million parameters in less than 30 minutes on a modest computing platform. As an example, we demonstrate that solving the SDP relaxation of OPF over the NY Grid using the proposed algorithm requires only a series of eigenvalue decompositions over small matrices of size at most 40, which runs very fast.

**2. Convexification of Power Flow Problem over Arbitrary Networks**

Ramtin Madani, Berkeley, CA, United States, 94709

Classical power flow (PF) problem aims to find the unknown parameter of a power network. This problem is solved via approximation in practice, under the assumption that bus voltages have small angles. In this talk, we propose a convex program with the property that it solves the PF problem precisely rather than approximately, as long as angles are small.

**3. Convexification of Unit Commitment Problem**

Morteza Ashraphijuo, UC Berkeley, Industrial Engineering & Operations Research, 1780 Le Roy Avenue, Berkeley, CA 94709 United States, ashraphijuo@berkeley.edu

The unit commitment (UC) is a mixed-integer nonlinear program. We develop a strengthened semidefinite programming (SDP) relaxation of the UC problem and study some hard instances of the UC problem for which the globally optimal values of the discrete and continuous variables of the UC problem are found by solving the proposed convex problem only once.

## ■ FA03

Bowl 2

### Parallel and Distributed Stochastic Optimization

Contributed Session

Chair: Alejandro Ribeiro, University of Pennsylvania, Electrical and Systems Engineering,200 South 33rd Street, Philadelphia, PA, 19104

**1. Parallel Asynchronous Algorithms for Nonconvex Large Scale Optimization**

Gesualdo Scutari, Associate Professor, Industrial Engineering, West Lafayette, IN, United States, 47907
Co-Authors: Loris Cannelli, PhD student, Industrial Engineering, West Lafayette, IN, United States, 47907, Francisco Facchinei, C/O Dipart Informatica E Sistem, Via Buonarroti 12, Roma, Italy, 00185, Paolo Scarponi, Industrial Engineering, West Lafayette, IN, United States, Vyacheslav Kungurtsev, Industrial Engineering, Prague, Czech Republic

We propose a parallel asynchronous lock-free algorithm for the minimization of the sum of a nonconvex smooth function and a nonsmooth convex one. The algorithm is proved to converge to stationary solutions of the problem while being scalable with the number of cores.

**2. DSA: A Decentralized Stochastic Averaging Method with Linear Convergence Rate**

Aryan Mokhtari, 200 South 33rd Street, Philadelphia, PA, United States, 19143
Co-Author: Alejandro Ribeiro, Room 203 Moore Building, Wynnewood, PA, United States, 19104

We propose decentralized double stochastic averaging gradient (DSA) as a stochastic decentralized method to solve consensus convex optimization problems. We show that the iterates generated by DSA converges linearly in expectation. Numerical results verify the convergence properties of DSA and showcase improvement relative to alternatives.

**3. Distributed Newton for the General Consensus Problem**

Rasul Tutunov, 465 Meadow Road, Apt. 4203, Princeton, NJ, United States, 08540
Co-Authors: Ali Jadbabaie, 200 South 33rd Street, Department of Electrical and Systems Engineer, Philadelphia, PA, United States, 19104, Haitham Bou Ammar, Princeton, NJ, United States, 19104

In this paper, we propose a distributed Newton method for the general consensus problem. Our method accurately approximates the Newton direction up-to-any arbitrary precision. The experimental results show that our technique outperforms state-of-the-art method including those from distributed ADMM.

## ■ FA04

Bowl 16

### Advances in Convexification Techniques for Structured Nonconvex Optimization

Contributed S ession

Chair: Fatma Kilinc-Karzan, Industrial and Systems Engineering, 227 Durham hall, 1145 Perry Street, Blacksburg, VA, United States, 24060

**1. On the Polyhedrality of the T-branch Closure and Other Closures**

Diego Moran, Industrial and Systems Engineering, 227 Durham Hall, 1145 Perry Street, Blacksburg, VA, United States, 24060
Co-Author: Oktay Gunluk, Ibm T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY, United States, 10598, Sanjeeb Dash, 1101 Kitchawan Road, Yorktown Heights, NY, United States, 10598

A t-branch split cut is an inequality that is valid for the set obtained by removing from a polyhedron the union of t split sets. Cook et al. (1990) and Dash et al (2013) showed that t-branch split cuts yield closures that are rational polyhedra for t=1,2. We further extend these results and show that the t-branch split closure is a polyhedron for all t=1,2,3...

**2. Intersection Cuts for Rank-constrained Semidefinite Programming**

Chen Chen, Postdoctoral Research Scientist, IEOR, New York, NY, United States, 10027
Co-Authors: Gonzalo Munoz, 500 W 120th Street, New York, NY, United States, 10027, Daniel Bienstock, Dept of IEOR, 342 Mudd, New York, NY, United States, 10027

We introduce intersection cuts for rank-one constrained semidefinite programming, applicable to polynomial programming. This stems from a general perspective of the intersection cut as a valid inequality for some generic closed point set F. As in integer programming, large F-free convex sets are crucial to generating effective intersection cuts.

**3. First-order Methods for Robust Convex Optimization**

Nam Ho-Nguyen, Tepper School of Business, Pittsburgh, PA, United States,

We present several efficient first-order methods to approximately solve robust convex optimization problems. We also introduce the notion of weighted regret online learning and the online saddle-point problem, which form key building blocks for our methods. Finally, we discuss some proximal-type algorithms for these problems.

**4. Approximation Tools for Structured Nonconvex Optimization Problems**

Fatma Kilinc-Karzan, Assistant Professor of Operations Research, Tepper School of Business, 5000 Forbes Ave, Pittsburgh, PA, United States, 15213
Co-Author: Levent Tuncel, Professor, Department of Combinatorics
and Optimization, Waterloo, ON, Canada

We study maximizing certain convex functions over convex domains. Such problems generalize computing matrix norms and arise in robust optimization and machine learning. We build tractable convex relaxations, study the properties, approximation quality, and exactness of these relaxations, and establish connections to existing results .

## ■ FA05

Room 023

### Stochastic Math Programming

Contributed Session

Chair: Kuo-Ling Huang, Northwestern University, Evanston, IL, United States, jupiters1117@northwestern.edu

**1. Decomposition for Two-stage Stochastic Unit Commitment Models**

Semih Atakan, Industrial and Systems Engineering, Los Angeles, CA, United States
Co-Author: Suvrajeet Sen, Industrial and Systems Engineering, Los Angeles, CA, United States, 90089-0193

We study a two-stage stochastic unit commitment problem with an embedded network capable of switching decisions. We show that this problem is amenable to decomposition algorithms with binary decisions in both stages.

**2. Stochastic Dynamic Linear Programming Algorithm**

Harsha Gangammanavar, Industrial Engineering, 134 Freeman Hall, Clemson University, Clemson, SC, United States, 29634
Co-Author: Suvrajeet Sen, Professor, Daniel J. Epstein Dept. of Industrial and Systems Engineering, 3715 McClintock Ave, GER 240, Los Angeles, CA, United States, 90089-0193

We present Stochastic Dynamic Linear Programming: a sequential sampling algorithm for multistage linear programs with interstage independent random parameters. We provide its convergence properties, extensions to special correlation structures and Markov processes, and computational comparison with stochastic dual dynamic programming.

**3. Computational Experiences with Benders Decomposition for Stochastic Mixed Integer Programs**

Sanjay Mehrotra, Professor, IEMS, Evanston, IL, United States, 60093
Co-Author: Kuo-Ling Huang, IEMS, Evanston, IL, United States

We present computational experience with an implementation of a multi-cut integer L-shaped method for solving two-stage stochastic MIPs. Advanced starting techniques, scenario clustering, convex function linearization strategies, and convexification techniques will be discussed within the iOptimize software package.

## ■ FA06

Sherrerd

### Nonconvex Optimization

Contributed Session

Chair: Jon Lee, Ann Arbor, MI, United States

**1. Greedy Approaches for Column Subset Selection**

Jason Altschuler, Computer Science, 0028 Frist Campus Center, Princeton, NJ, United States, 08544
Co-Authors: Morteza Zadimoghaddam, Gang (Thomas) Fu, Afshin Rostamizadeh, Vahab Mirrokni, New York, NY, United States, Aditya Bhaskara, School of Computing, Salt Lake City, UT, United States

Column subset selection is a powerful approach for approximating large matrices while preserving the inherent meaning of columns. We show novel approximation guarantees for the natural greedy algorithm. We further present bounds on a distributed coreset framework that allows the greedy algorithm to scale to massive data sets. 2.

**2. Projection Algorithms for Nonconvex Minimization with Application to Sparse Principal Component Analysis**

Dzung Phan, Research Staff Member, 1101 Kitchawan Road, P.O. Box 218, Yorktown Heights, NY, United States, 10598
Co-Authors: William Hager, Professor, Department of Mathematics, Gainesville, FL, United States, 32611, Jiajie Zhu, Department of Mathematics, Gainesville, FL, United States, 32611

We minimize a concave function over nonconvex sets, and propose a gradient projection algorithm (GPA) and an approximate Newton algorithm (ANA). Convergence results are established. In sparse principal component analysis, the performance of GPA is very similar to the fastest current methods. Sometimes, ANA is much faster and gives a better solution.

**3. Virtuous Smoothing for Global Optimization**

Jon Lee, Management Science, Iowa City, IA, United States, 52245, Jon Lee, Ann Arbor, MI, United States
Co-Author: Daphne Skipper, Augusta, GA, United States

Exact solvers for global-opt (GO) and mixed-integer nlp (MINLP) rely on NLP solvers. Convergence of most NLP solvers requires that functions be twice continuously differentiable. Yet many models naturally utilize functions with some limited non-differentiability. We propose and analyze a smoothing method in the context of GO/MINLP models.

# Friday, 10:30am – 11:30 am

## FB01

Dodds Auditorium

### Plenary Session: Optimization Challenges in Statistics and Machine Learning

Han Liu - Princeton University

**1. Optimization Challenges in Statistics and Machine Learning**

Optimization lies at the heart of machine learning. In this talk, I will use the graphical model learning problem to illustrate several representative optimization challenges in statistics and machine learning. These challenges range from high dimensionality, nonconvexity, to massive sample size. In addition to presenting these challenges, I will focus on explaining the origins of these challenges and show that they come from very natural formulation of the learning problems. We will also introduce several new research directions lying at the intersection between optimization and machine learning.

# Friday, 1:00pm – 2:30pm

## ■ FC01

Dodds Auditorium

### Computational Optimization in Learning - II

Contributed Session

Chair: Yuying Li, Computer Science, 200 University Avenue West, Waterloo, ON, Canada, N2L 3G1

**1. Optimal Learning of Binary Programs**

Boris Defourny, ISE, 200 W Packer Ave, Bethlehem, PA, United States, 18015

We consider the expected improvement criterion for an integer program with binary decisions and unknown coefficients, learned by Bayesian updating using trial solutions. This leads to a learning problem which is a MIP with second-order cone constraints. Comparable results for related optimal design of experiment criteria are established as well.

**2. Distributed Stochastic Variance Reduced Gradient Methods and a Lower Bound for Communication Complexity**

Qihang Lin, The University of Iowa, Iowa City, IA, Contact: qihang-lin@uiowa.edu, Tengyu Ma, Princeton University, Princeton, NJ, Tianbao Yang, Jason Li, UC Berkeley, Berkeley, CA,

We design a distributed stochastic variance reduced gradient algorithm for minimizing the average of convex functions, which needs fewer rounds of communication than others. We also prove a lower bound for the rounds of communication for a broad class of methods including the proposed algorithms. We show that our method achieves this lower bound.

**3. Update Algorithms for the Roundoff-error-free Lu and Cholesky Factorizations**

Adolfo R Escobedo, PhD Candidate, Industrial & Systems Engineering, 3131 TAMU, College Station, TX, United States, 77843
Co-Author: Erick Moreno-Centeno, Industrial & Systems Engineering, 3131 TAMU, College Station, TX, United States, 77843

We introduce efficient update algorithms for the Roundoff-Error-Free (REF) LU and Cholesky factorizations. The updates are addition, deletion, and replacement of rows and columns of a basis. Combined with REF substitution, the featured algorithms provide a complete framework for solving LPs exactly and efficiently.

**4. Unbalanced Learning with RankRC**

Yuying Li, Computer Science, 200 University Avenue West, Waterloo, ON, Canada, N2L 3G1

We propose RankRC to address challenges associated with both rare class problems and large scale learning, by 1) optimizing area under curve of the receiver of operator characteristic in the training process, and 2) using a rare class kernel representation. We provide computational and theoretical justifications for the rare class representation.

## ■ FC02

Bowl 1

### Degeneracy in Optimization

Contributed Session

Chair: Dimitri Papadimitriou, Copernicuslaan 50, Antwerp, Belgium, 2018

**1. On the Solution of Some Very Large-scale Highly Degenerate Combinatorial Optimization Problems: Applications to Clustering**

Claudio Contardo, Associate Professor, Management and Technology, Montreal, QC, Canada, H2X 3X2
Co-Author: Daniel Aloise, Natal, Brazil, 59088650

We consider the solution of some combinatorial problems presenting high degrees of degeneracy. Our algorithm starts with a small sample of nodes and proceeds iteratively to enlarge that sample in a way that ensures the exactness and the convergence of the method. We exemplify the use of our algorithm by solving the minimax diameter clustering problem.

### 2. The Many Faces of Degeneracy in Conic Optimization

Henry Wolkowicz, Waterloo, ON, Canada
Co-Author: Dmitriy Drusvyatskiy, Seattle, WA, United States

Slater condition, existence of a strictly feasible solution, is at the heart of convex optimization. We describe the various reasons for the loss of strict feasibility,and describe ways to cope with it and, in many pronounced cases, how to use it as an advantage.

### 3. Vector Space Decomposition for Linear Programs

Jacques Desrosiers, Decision Sciences, 3000, Cote-Sainte-Catherine, Montréal, QC, Canada, H3T 2A7
Co-Authors: Marco Luebbecke, Operations Research, Kackertstr. 7, Aachen, Germany, D-52072, Jean Bertrand Gauthier, Decision Sciences, Montréal, QC, Canada

It's all about degeneracy! We present a primal framework for which special cases are the Primal Simplex, the strongly polynomial Minimum Mean Cycle-Cancelling algorithm for networks, and the Improved Primal Simplex. VSD allows identifying a family of algorithms that totally avoid degenerate pivots and also finds interior directions.

### 4. Multi-level Routing-Facility Location Problems

Dimitri Papadimitriou, Copernicuslaan 50, Antwerp, Belgium, 2018

The combined routing-facility location problem (FLP) removes the demand allocation independence property of the standard FLP. With multiple levels, routing also may follow different strategies. In this paper, we propose a generic mixed-integer formulation and present results obtained when applying different routing strategies between levels.

## ■ FC03

Bowl 2

## Tutorial Session: Robust Optimization: Theory and Applications

Contributed Session

Chair: Chaithanya Bandi, Managerial Economics and Decision Sciences, Evanston, IL, United States

### 1. Tutorial: Tractable Simulation-Optimization: A Robust Optimization Approach

Chaithanya Bandi, Managerial Economics and Decision Sciences, Evanston, IL, United States,
Co-Authors: Nataly Youssef, 20 Palermo St., Cambridge, MA, United States, 02141, Dimitris Bertsimas, Operations Research Center, Cambridge, MA, United States

To understand and optimize the performance of a system under uncertainty, the traditional approach consists of positing a probability distribution over the uncertain parameters, and derive information on the system performance and optimize it. When the system's dynamics are complex, computing the above expression analytically becomes challenging, and simulation becomes the only resort. However, simulation models may take a considerable amount of time for the results to be statistically significant and can often be complex, making it difficult to understand and further optimize key qualitative insights. Motivated by these challenges, we propose a robust optimization approach to analyzing and optimizing the expected performance of stochastic systems characterized by linear dynamics, based on a robust optimization framework. Our approach consists of the following steps: (1) We model uncertainty via polyhedral sets inspired by the limit laws of probability. The size of these sets is characterized by few variability parameters, which controls the degree of conservatism of the model. (2) We then cast the performance analysis problem as a optimization problem, for instance maximizing a given performance measure subject to the system's dynamics and the parametric uncertainty sets. The output of this optimization can be thought of as a worst case performance measure function of the variability parameters. (3) Instead of computing the expected performance of the system via averaging simulation outputs, we assume the variability parameters follow limiting distributions (derivatives of the normal distribution for light-tailed systems and the stable distribution for heavy-tailed systems) and average the worst case outputs from the optimization over a discretized space over the variability parameters. (4) To optimize the stochastic system, we propose to cast the problem of finding the optimal design input via a robust optimization problem, e.g., minimizing the average or conditional value-at-risk of the worst case performance outputs. This framework allows a significant reduction in the dimension of uncertainty which provides grounds for tractable analyses. We illustrate the tractability and accuracy of our approach by (a) simulating the transient behavior of multi-server feedforward queueing networks, and (b) determining optimal base stock policies of generalized supply chain networks.

### 2. Empirical Divergence-based Distributionally Robust Optimization

Henry Lam, Ann Arbor, MI, United States

We present an empirical likelihood framework to calibrate the uncertainty set of DRO with divergence ball centered at an empirical distribution, via generalized chi-square quantiles. This is part of "Robust Optimization: Theory and Applications" session.

### 3. Static Vs Adjustable Solutions in Dynamic Robust Optimization

Vineet Goyal, 500W 120th St, New York, NY, United States, 10027
Co-Author: Pranjal Awasthi, Piscataway, NJ, United States
Co-Author: Brian Yin Lu, 539 W. 112th St., Apt. 2F, New York, NY, United States, 10025

We consider two stage adjustable robust optimization problem with uncertain constraint coefficients that models many important applications including resource allocation with uncertain requirements. We show that a static solution gives a near-optimal approximation for the adjustable robust problem for fairly general class of uncertainty sets.

## ■ FC04

Bowl 16

## Multilevel Optimization

Contributed Session

Chair: Cole Smith, Industrial Engineering, Clemson, SC, United States, 29634

### 1. Accurate Optimization for Multiscale Metabolic Networks

Ambros Gleixner, Optimization, Berlin, Germany, 10551

Linear programming has become a useful tool to analyze metabolic networks in computational biology. So-called ME-models yield multiscale LPs that easily pose numerical difficulties to double-precision solvers. We use the exact LP solver SoPlex to measure the accuracy of floating-point solvers and compute high-precision solutions.

### 2. Bilevel Risk Averse formulations of Stochastic Programming Problems

Deniz Seyed Eskandani, PhD Candidate, Rutgers Business School, 100 Rockafeller Rd., Piscataway, NJ, United States, 08854
Co-Author: Jonathan Eckstein, 100 Rockafeller Road, Room 5145, Piscataway, NJ, United States, 08854

We describe a bilevel programming technique to time-consistently formulate 3-stage stochastic programs without using nested risk measures. For some classes of applications, we empirically demonstrate that its behavior can be dramatically different from standard formulations. 3.

### 3. Exact Algorithms and Bounds for the Dynamic Assignment Interdiction Problem

Jorge A Sefair, School of Computing, Informatics, and Decision Systems Engineering, Tempe, AZ, United States
Co-Author: J. Cole Smith, Professor & Department Chair, Department of Industrial Engineering, Clemson, SC, United States

We study a dynamic assignment interdiction game in which two agents, a user and an attacker, compete in an assignment graph. The user assigns tasks one at a time and the attacker interdicts a subset of assignments to maximize the user's minimum cost. The goal is to find an optimal sequence of assignments, and the attacker's optimal interdiction strategy

### 4. A Sampling-based Exact Approach for the Bilevel Mixed Integer Programming Problem

Cole Smith, Industrial Engineering, Clemson, SC, United States, 29634
Co-Author: Leonardo Lozano, Industrial Engineering, Clemson, SC, United States

We examine bilevel MIPs in which each agent optimizes different objectives, and in which each agent's variables affect the other's feasible region. We propose an exact algorithm based on an adaptive sampling scheme. The algorithm can accommodate optimistic or pessimistic follower assumptions.

## ■ FC05

Room 023

### Stochastic Vehicle Routing

Contributed Session

Chair: Lei Zhao, Industrial Engineering, Beijing,China

**1. Offline-online Approximate Dynamic Programming for Dynamic Vehicle Routing**

Marlin W Ulmer, Braunschweig, Germany
Co-Author: Justin Goodson, John Cook School of Business, Davis-Shaughnessy Hall, St. Louis, MO, United States, 63108, Dirk Mattfeld, Braunschweig, Germany, Marco Hennig, Braunschweig, Germany

For a dynamic vehicle routing problem with stochastic requests, we combine (offline) value function approximation (VFA) with rollout algorithms as a means of enhancing the anticipation of the VFA policy, resulting in offline-online approximate dynamic programming.

**2. The Information-collecting Vehicle Routing Problem: Stochastic Optimization for Emergency Storm Response**

Lina Al-Kanj, ORFE, Sherrerd Hall, ORFE Department, Room 115, Princeton University, Princeton, NJ, United States, 08544
Co-Author: Warren B Powell, 230 Sherrerd Hall, Dept of Operations Research and Financial Eng, Princeton, NJ, United States, 08544

We address the problem of managing a single utility truck, which we start by formulating as a sequential stochastic optimization model which captures our belief about the state of the grid. We propose a stochastic lookahead policy, and use Monte Carlo tree search (MCTS) to produce a practical policy that is asymptotically optimal.

**3. The Vehicle Routing Problem with Drones: A Worst-case Analysis**

Xingyin Wang, Mathematics, 8508 16th street, Apt 705, Silver Spring, MD, United States, 20910
Co-Authors : Bruce L Golden, The France-Merrick Chair in Management Science, R. H. Smith School of Business, College Park, MD, United States, 20742, Stefan Poikonen, 3406 Tulane Dr., Apt 22, Hyattsville, MD, United States, 20783

In the Vehicle Routing Problem with Drones (VRPD), a fleet of trucks equipped with drones delivers packages to customers. Drones are launched from and picked up by the trucks. The objective is to minimize the maximum duration of all routes. We compare VRPD to the min-max VRP in a worst-case study to show the large potential savings from using the drones.

**4. Optimal Information Collection in the Vehicle Allocation Problem**

Lei Zhao, Industrial Engineering, Beijing, China
Co-Authors: Ilya O Ryzhov, 4322 Van Munching Hall, Robert H. Smith School of Business, College Park, MD, United States, 20742, Warren B Powell, 230 Sherrerd Hall, Dept of Operations Research and Financial Eng, Princeton, NJ, United States, 08544, Yixiao Huang, Industrial Engineering, 530 Shunde Building, Beijing, China, 100084

We study the vehicle allocation problem where an urban logistics company divides the urban area into several regions and allocates delivery vehicles to these regions. As the cost evaluation of a vehicle allocation decision is normally expensive, we focus on how to efficiently collect the cost evaluation information to allocate the vehicles optimally.

## ■ FC06

Sherrerd

### Low-rank Approximations in Machine Learning

Contributed Session

Chair: Anil Aswani, Berkeley, CA, United States

**1. Low-RankApproximations to Markov Decision Processes**

Tsvetan Asamov, Sherrerd Hall, Charlton Street, Princeton, NJ, United States, 08544

We develop new methods to approximate the optimal policies of Markov decision processes using low-rank approximation techniques applied to sampled lookup tables. Our computational results indicate that the proposed approaches can reduce computational times by several orders of magnitude.

**2. On Low-rank Entropy Maximization of Covariance Matrices**

Fu Lin, Lemont, IL, United States, 60439
Co-Author: Jie Chen, Research Staff Member, Yorktown Heights, NY, United States, 10598

We consider entropy maximization of covariance matrices regularized with nuclear norm. We solve this problem via solving a sequence of linear equations. Based on a nullspace parameterization, we exploit special structure of Hessian matrices that makes matrix-vector multiplication efficient.

**3. Sharp Computational and Statistical Phase Transitions via Oracle Computational Model**

Zhaoran Wang, Princeton, NJ, United States
Co-Authors: Han Liu, 35 Olden Street, Princeton, NJ, United States, 08544, Quanquan Gu, Princeton, NJ, United States

We study the fundamental tradeoffs between computational tractability and statistical accuracy for a general family of hypothesis testing problems with combinatorial structures. For an extended abstract, see http: //arxiv.org/abs/1512.08861.

**4. Low-rank Approximation and Completion of Positive Tensors**

Anil Aswani, Berkeley, CA, United States

We use convex optimization to develop polynomial-time algorithms for low-rank approximation and completion of positive tensors. Our approach uses algebraic topology to define a new decomposition for positive tensors. For tensor completion, we require a polynomial number of measurements, while existing approaches need exponential measurements.

# Friday, 2:40pm – 3:40pm

## ■ FD01

Dodds Auditorium

### Plenary: Optimization in the Emerging World of Robotics – Russ Tedrake

**1. Exploiting Structure for Robust Model-based Optimization in Robotics**

I find that the optimization community often assumes that the algorithms powering the robots seen on youtube are already using all of the best optimization tools available. In almost every case, they are not (yet!). For many of the most impressive robots, simple, intuitive, hand-tuned control systems still rule the day, despite requiring massive amounts of man-power and testing to get right. Optimization-based controllers still fail to match their simplicity and robustness (e.g. to sensing and model errors) in the real world.

This trend is starting to change. For example, the last few years have seen optimization methods emerge as a popular tool for whole-body planning and control on humanoid robots. Quadratic-programming approaches to whole-body control generalize traditional Jacobian-based approaches by supporting hard inequality constraints due to control and friction limits. Online center of mass / zero-moment point planning algorithms can be similarly generalized. Thanks to fast computers and advanced solvers, these optimizations can now be solved inside a high-rate feedback loop. More than technological, this represents a cultural change with more of the field looking towards optimization for more general solutions.

We need your help. In some cases, we are still not posing the right questions. In other cases, we need more work to exploit the structure in the problems to make the problems we want to formulate tractable. This structure includes the famous structure of the floating-base Lagrangian, sparsity in the (inverse) dynamics, the algebraic structure of the kinematics, dynamics, and constraints, and the inherent combinatorial structure imposed by interacting with the environment by making and breaking contact. I will describe my group's initial work towards exploiting this structure in our algorithms, and our detailed approach to applying these tools to humanoid robots and fast-flying UAVs.

# Friday, 4:10pm – 5:40pm

## ■ FE01

Dodds Auditorium

### Decomposition Methods for Learning

Contributed Session

Chair: Darinka Dentcheva, Castle Point on Hudson, Hoboken, NJ, United States, 07030

**1. Learning In Neuromorphic Systems as an Adaptive Optimization Process**

Gary Engler, Mathematical Sciences, Hoboken, NJ, United States
Co-Author: Michael Zabarankin, Dept of Mathematical Sciences, Castle Point on Hudson, Hoboken, NJ, United States, 07030

In a live environment an agent adapts the structure of its neuronal network to maximize the reward, which is an online process. This work develops a class of policies for optimal guiding of the evolution of the neuromorphic system controlling the agent, inspired by brain architecture and governed by a system of differential equations.

**2. Risk Sharing in Classification Problems**

Constantine A Vitt, PhD Candidate, MSIS, 1 Washington Park, Newark, NJ, United States, 07102
Co-Authors: Darinka Dentcheva, Castle Point on Hudson, Hoboken, NJ, United States, 07030, Hui Xiong, Professor, MSIS, Newark, NJ, United States, 07102

We consider a classification problem, in which we take into account the decision-maker's attitude to misclassification errors. Our approach is based on the theory of coherent measures of risk and risk-sharing. We develop risk-sharing ideas for distance based classification. We validate our results via comparison to other robust classifier methods.

**3. Selective Linearization for Multi-block Convex Optimization**

Yu Du, MSIS-RUTCOR, 100 Rockafeller Road, Room 5111, Piscataway, NJ, United States, 08854
Co-Authors: Xiaodong Lin, Associate Professor, MSIS, 100 Rockafeller Road, Room 5143, Piscataway, NJ, United States, 08854,
Andrzej Ruszczynski, 5 Cedar Court, Princeton Junction, NJ, United States, 08550

We introduce the Selective Linearization(SLIN) Algorithm to solve the multi-block structured convex non-smooth functions. Global convergence is achieved and SLIN algorithm is proven to converge at rate of O(1/k). We apply the SLIN algorithm on structured regularization applications, showing fast running time for large scale problems.

**4. An Augmented Lagrangian Method for Distributed Optimization**

Darinka Dentcheva, Professor, Mathematical Sciences, Castle Point on Hudson, Hoboken, NJ, United States, 07030

We present a novel distributed numerical method for convex optimization problems and analyse its convergence. The method is based on sequential local augmented Lagrangian functions. It is applied to two network models, as well as to a two-stage stochastic optimization problem. The method compares favorably to other known distributed methods.

## ■ FE02

Bowl 1

### Optimization Challenges in Robotics

Contributed Session

Chair: Anirudha Majumdar, Cambridge, MA, United States, 02139

**1. Controlling Robot Motions through Contact with Optimization**

Scott Kuindersma, Assistant Professor, Computer Science, Cambridge, MA, United States
Co-Authors: Russell L. Tedrake, 32 Vassar St, Cambridge, MA, United States, 02139, Michael Posa, Cambridge, MA, United States

We will summarize our recent work designing optimization algorithms to plan and stabilize dynamic motions for complex robots including 1) the development and implementation of convex optimization-based controllers for stabilizing legged locomotion on a humanoid and 2) efficient and accurate planning for constrained hybrid dynamical systems.

**2. Solving Conic Optimization Problems via Self-dual Embedding and Facial Reduction: A Unified Approach**

Frank Permenter fperment@mit.edu, Henrik Friberg, and Erling Andersen

Conic optimization problems (e.g., semidefinite programs) arising in robot-controller design can be extremely difficult to solve and frequently cause solver failure. This talk addresses one source of failure. In particular, this talk addresses inevitable failure of the self-dual embedding method (implemented in solvers such as SeDuMi) on problem instances degenerate in a particular sense (e.g., instances with duality gaps or instances with unattained optimal values). As we show, this method can succeed in all cases, even on these degenerate instances, if modified in a conceptually simple way. Specifically, we identify and exploit a connection between the method and facial reduction, a regularization technique due to Borwein and Wolkowicz. This is joint work with Henrik A. Friberg and Erling D. Andersen.

**3. Optimization Techniques for Controlling Highly Agile Robots with Formal Safety Guarantees**

Anirudha Majumdar, Cambridge, MA, United States, 02139
Co-Author: Russ Tedrake, Cambridge, MA, United States

This talk will present work on using tools from convex optimization such as sums-of-squares (SOS) programming for pushing highly agile robotic systems to the brink of their hardware limits while ensuring that they operate in a safe manner despite uncertainty in the environment and dynamics.

## ■ FE03

Bowl 2

### Risk-Averse Optimization of Markov Systems

Contributed Session

Chair: Andrzej Ruszczynski, Man. Sci. Inf. Sys., Piscataway, NJ, United States

**1. Process-based Risk Measures for Partially Observable Markov Decision Problems**

Jingnan Fan, Rutcor, 100 Rockafeller Rd, Piscataway, NJ, United States, 08854
Co-Author: Andrzej Ruszczynski, 100 Rockafeller Rd, Piscataway, NJ, United States, 08854

We provide theoretical foundations of the theory of process-based dynamic risk measures for controlled stochastic processes by introducing a new concept of conditional stochastic time consistency. In this talk, we focus on the applications to the Partially Observable Markov Decision Problems.

**2. Risk-averse Approximate Dynamic Programming for Optimal Learning with Application to Clinical Trial Design**

Curtis McGinity, Management Science and Information Systems, Piscataway, NJ, United States

We formulate the risk-averse optimal learning problem for logistic regression in multistage clinical trials. We characterize the Markov problem in terms of risk composition and log-concavity, and introduce an approximation architecture for its solution via ADP. Finally, we compare prominent policies via simulation study.

**3. Risk-averse Control of Diffusion Process**

Jianing Yao, PhD Student, Operations Research, 100 Rockafellar Ave, Piscataway, NJ, United States, 08854

We consider the risk-averse control of diffusion process, where decoupled FBSDE system is used to evaluate certain policy. Risk-averse DPE and HJB is derived by the by the nice property of generator. We also use a perturbation method to approximate the original formulation by using piece-wise constant control.

**4. Risk-averse Control of Markov Chains**

Andrzej Ruszczynski, Man. Sci. Inf. Sys., Piscataway, NJ, United States, 08854

We introduce the concepts of conditional stochastic time consistency and Markovian risk measures and we derive their structure. This allows us to derive a risk-averse counterpart of dynamic programming equations. Then we extend these ideas to continuous-time Markov chains and derive the structure of risk measures in this case as well.

## ■ FE04

Bowl 16

### Tutorial Session: Online Convex Optimization

Contributed Session

**1. Tutorial: Online Convex Optimization**

Elad Hazan

We describe the recent framework of online convex optimization which naturally merges optimization and regret minimization in games. We survey the basic algorithms and tools at the heart of this framework, which have led to the resolution of fundamental questions in online learning.

**2. Some Topics in Stochastic First-Order Methods**

Mengdi Wang, Princeton University, 226 Sherrerd Hall, Princeton, NJ 08544 United States, mengdiw@princeton.edu

We discuss several research questions about stochastic gradient descent method and its variants, with an emphasis on data-intensive applications. More specifically, we discuss issues such as sample complexity, multi-level randomness, constraint randomization, and non-convexity. We illustrate applications in statistical learning and online solution of Markov decision process.

**3. Near-optimal Stochastic Approximation for Online Principal Component Estimation**

Chris Junchi Li , Princeton Universtiy, Princeton, NJ, United States, junchil@princeton.edu

Principal component analysis has been a prominent tool for high dimensional data analysis. Online algorithms that estimate the principal component by processing streaming data are of tremendous practical and theoretical interests. We prove for the first time a nearly optimal convergence rate result for the online PCA algorithm.

## ■ FE05

Room 023

### Stochastic Optimization and Applications

Contributed Session

Chair: Mark S Squillante, Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY, United States, 10598

**1. Stochastic Power Generation Scheduling Using Temoa**

Anderson R. de Queiroz, Postdoctoral Research Scholar, Civil, Construction and Environmental Engineering, Raleigh, NC, United States, 27609
Co-Authors: Joseph F DeCarolis, Assistant Professor, Civil, Construction and Environmental Engineering, Raleigh, NC, United States, Binghui Li, Ph.D. Candidate, Civil, Construction and Environmental Engineering, Raleigh, NC, United States,

We present the energy economy optimization model Temoa applied in the context of power generation scheduling. We consider a system with thermal plants and stochastic hydro-power in order to decide the operational plan to satisfy demand over a short-term horizon. A case study is presented using PySP.

**2. Distributionally Robust Models with Second Order Asymmetry Information and Applications to the Newsvendor Problem**

Joline Uichanco, Ross School of Business, Ann Arbor, MI, United States, 48103
Co-Author: Karthik Natarajan, 12 Amber Gardens, 02-06, Block B, Kings Mansion, Singapore, Singapore, 439959

We introduce asymmetry information in a distributionally robust newsvendor problem using the covariance of the positive part and the negative part of the random demand minus the mean. The problem is NP-hard, but admits a semidefinite program upper bound. Experiments on spare parts data show our model performs better than mean-covariance models.

**3. Optimal Resource Capacity Management In Stochastic Networks**

Mark S Squillante, Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY, United States, 10598
Co-Authors: Ton Dieker, Associate Professor, IEOR, New York City, NY, United States, Soumyadip Ghosh, Math Sciences and Analytics, Yorktown Heights, NY, United States, 10598

We develop a framework for optimizing resource capacities of stations of a stochastic network. Central to it is an iterative approximation scheme that relies only on observing station queue lengths, unlike standard methods that need gradients. We provide an analysis for single-class Brownian tree networks, and present illustrative experiments.

## ■ FE06

Sherrerd

### Decision Diagrams in Optimization

Contributed Session

Chair: John Hooker, Pittsburgh, PA, United States, 15213

**1. Decision Diagram Decompositions**

David Bergman, 1 University Place, Stamford, NY, United States, 06901

This talk explores the use of decision diagrams (DDs) for decomposing a problem so as to ensure that the DD on each part of the decomposition is limited in size. IP is used to link the DDs by transforming the DDs into network flow models.

**2. Decision Diagrams for Sequencing and Scheduling Problems**

Andre A Cire, Dept. of Management UTSC & Rotman School of Management, Toronto, ON, Canada, M1C 1A4
Co-Author: David Bergman, 1 University Place, Stamford, CT, United States, 06901

In this talk we discuss how decision diagrams can be a powerful tool for modeling and solving a variety of sequencing and scheduling problems. We describe applications in routing with complex objective functions and side constraints, and introduce techniques to combine decision diagrams with other relaxation techniques typically used in scheduling.

**3. Overview of Decision Diagrams in Optimization**

John Hooker, Pittsburgh, PA, United States, 15213

This session introduces decision diagrams, which have recently brought a new perspective to optimization. Decision diagrams relax problems without presupposing linearity or convexity and solve recursive formulations by branch and bound, often faster than standard solvers. This first talk provides an overview of this small but fast-moving area.

# Saturday, 8:30am – 10:00 AM

## ■ SA01

Dodds Auditorium

### Stochastic Optimization in Learning

Contributed Session

Chair: Alec Koppel, 200 South 33rd Street, Philadelphia, PA, United States, 19104

**1. Homotopy Method for Learning Sparse Timevarying Signal**

Xue Wang, 915 Southgate Drive, Apt. 18, State College, PA, United States, 16801
Co-Author: Runze Li, University Park, PA, United States, Hongcheng Liu, 1000 Escalon Avenue, Apt. D2031, Sunnyvale, CA, United States, 94085, Tao Yao, Penn State University, 310 Leonhard Building, University Park, PA, United States, 16802

We propose a new Smoothly Clipped Absolute Deviation (SCAD)-minimization homotopy algorithm for sparse signal learning in a time-varying framework. We demonstrate that the proposed method outperforms benchmark approaches in terms of statistical property and computation complexity.

**2. Learning In Neuromorphic Systems as an Adaptive Optimization Process**

Gary Engler, Mathematical Sciences, Hoboken, NJ, United States
Co-Author: Michael Zabarankin, Dept of Mathematical Sciences, Castle Point on Hudson, Hoboken, NJ, United States, 07030

In a live environment an agent adapts the structure of its neuronal network to maximize the reward, which is an online process. This work develops a class of policies for optimal guiding of the evolution of the neuromorphic system controlling the agent, inspired by brain architecture and governed by a system of differential equations.

**3. Optimal Learning for Parametric Estimation for Nonlinear Belief Models**

Xinyu He, Electrical Engineering, Princeton, NJ, United States, 08544
Co-Author: Warren B Powell, 230 Sherrerd Hall, Dept of Operations Research and Financial Eng, Princeton, NJ, United States, 08544

Given a nonlinear function with unknown parameters, we propose an algorithm to estimate the correct parameters and use the optimal learning methods to find the alternative that maximizes the function through a series of noisy and expensive measurements.

**4. Doubly Random Parallel Stochastic Methods for Large Scale Learning**

Alec Koppel, PhD Student, 200 South 33rd Street, Philadelphia, PA, United States, 19104
Co-Authors: Aryan Mokhtari, Phd Student, Electrical and Systems Engineering, 200 South 33rd Street, Philadelphia, PA, United States, 19104, Alejandro Ribeiro, Electrical and Systems Engineering,
101 Crosshill Rd., Wynnewood, PA, United States, 19096

We propose the random parallel stochastic algorithm (RAPSA), which is the first method that is parallel in both the selection of blocks and the selection of training examples. We propose two extensions of RAPSA: accelerated RAPSA (ARAPSA), as well as an architecture in which parallel processors may operate asynchronously.

## ■ SA02

Bowl 1

### Tutorial Session: Optimization with Equilibrium Constraints

Contributed Session

Chair: Uday Shanbhag, State College, PA, United States

**1. Tutorial: Equilibrium Problems in Power Markets: Models, Analysis and Computation**

Uday Shanbhag, Pennsylvania State University, State College, PA, United States, udaybag@engr.psu.edu

Given the sheer breadth of the field of power markets, there have been several past efforts at reviewing the literature. Yet, many gaps persist in both the reviews as well as in the related literature. Primary amongst these is a lack of a comprehensive treatment of equilibrium models in power markets characterized by the absence of the following: (i) the ability to capture a range of competitive interactions (perfectly competitive, Nash-Cournot, and its variants) via a unified framework; (ii) the capacity to cope with uncertain parameters and risk-based formulations; and (iii) the ability to incorporate hierarchy, a consequence of the multitude of settlements that dene most power markets. Motivated by these lacunae, in this tutorial, via complementarity theory, equilibrium programming, and stochastic optimization, we intend to provide the analytical and algorithmic foundations for contending with challenges arising in (i), (ii), and (iii).

**2. An Exact Solution Method for Binary Equilibrium Problems with Compensation**

Sauleh A Siddiqui, Civil Engineering, 3400 N Charles St, Baltimore, MD, United States, 21218
Co-Author: Daniel Huppmann, Matthaeusgasse 8, Vienna, Austria, 1030

We propose a novel method to find Nash equilibria in games with binary decision variables by including compensation payments and incentive-compatibility constraints from non-cooperative game theory directly into an optimization framework in lieu of using first order conditions of a linearization, or relaxation of integrality conditions.

## ■ SA03

Bowl 2

### Stochastic Optimization

Contributed Session

Chair: Soumyadip Ghosh, Math Sciences and Analytics, Yorktown Heights, NY, United States, 10598

**1. Dynamic Assortment Customization In High Dimensions**

Nathan Kallus, Cambridge, MA, United States,
Co-Author: Madeleine Udell, Pasadena, CA, United States

We study dynamic assortment customization with many heterogenous types and many items, show that natural structure allows us to learn all preferences surprisingly quickly via convex optimization, and devise algorithms that achieve optimal regret in high-dimensional, dynamically-learn-and-optimize settings where existing approaches fail.

**2. Online Allocation with Traffic Spikes: Mixing Adversarial and Stochastic Models**

Hossein Esfandiari, A.V. Wiliams Building, Room #1103, College Park, MD, United States, 20742
Co-Authors: Vahab Mirrokni, 76 Ninth Ave, Google, New York, NY, United States, 10011, Nitish Korula, 76 Ninth Ave, 4th Floor, New York, NY, United States, 10011

While the adversarial arrival models are too pessimistic, many of the stochastic models do not realistically capture uncertainty in predictions. Here, we propose a robust online model that captures the nature of traffic spikes. In our model, in addition to the stochastic input, an unknown number of impressions arrive that are adversarially chosen.

**3. Stochastic Methods for Relaxations of Large-scale Dense Semidefinite Programs**

Yifan Sun, Los Altos, CA, United States, 94022

We give efficient stochastic methods for estimating very large-scale semidefinite programs, where the positive semidefinite (PSD) constraint is relaxed to PSD constraints on all low-dim. principal submatrices. We discuss the benefits and limitations of this estimate for specific applications.

**4. Stochastic Approximation Methods for Computing Worst-case Input Models for Stochastic Simulations**

Soumyadip Ghosh, Math Sciences and Analytics, Yorktown Heights, NY, United States, 10598
Co-Author: Henry Lam, 1205 Beal Avenue, Industrial & Operations Engineering, Ann Arbor, MI, United States, 48109

Simulation models are subject to misspecification of input models when data is limited. We formulate bounds on worst-case performance as a stochastic program on the space of distributions subject to constraints imputed from the data. These are solved efficiently as stochastic approximations. A rate analysis and illustrative examples are presented.

## ■ SA04

Bowl 16

### Integer Programming and Discrete Optimization

Contributed Session

Chair: Jose L Walteros, Industrial and Systems Engineering, 413 Bell Hall, University at Buffalo, Buffalo, NY, United States, 142601.

**1. Derivative Free AUC Optimization**

Hiva Ghanbari, ISE, 11 Duh Drive, Apt. 322, Bethlehem, PA, United States, 18015
Co-Author: Katya Scheinberg, Professor, ISE, Bethlehem, PA, United States

Conventional classification problems in machine learning suffer from the existence of imbalanced data sets. In order to prevent the dominating effect of majority class, AUC optimization rather than error rate minimization is preferred. In this work, we have used trust region based derivative free optimization to optimize AUC function.

**2. Misoco Models and Disjunctive Conic Cuts for Truss Topology Design Problems**

Mohammad Shahabsafa, PhD Student, Industrial and Systems Enginerring, Bethlehem, PA, United States, 18015
Co-Authors: Ali Mohammad-Nezhad, Industrial and Systems Enginerring, Bethlehem, PA, United States, Tamas Terlaky, 200 West Packer Avenue, Dept ISE, Bethlehem, PA, United States, 18015

We present novel models, including Mixed Integer Second Order Cone Optimization (MISOCO), for Truss Topology Design Problems. Preliminary results indicate the effectiveness of Disjunctive Conic Cuts (DCCs) when solving the MISOCO models in a Branch & Conic Cut framework. Application of DCCs reduces number of nodes and in most cases the solution time.

**3. Mixed-Integer Linearly-Constrained Convex Programming Reduced to Unconstrained Convex Programming**

Elmor L Peterson, Retired, 3717 Williamsborough Court, Raleigh, NC, United States, 27609

Each “Integer-relaxed” problem is reformulated as a “GGP problem” whose “conjugate dual” is unconstrained and hence solvable via UCP — after which any “cutting planes” resulting from the “primal-optimal” solutions produce “dual-updates” that are also unconstrained — and hence solvable without “branching”, “bounding”, or “cold re-starting”.

**4. Integer Programming Formulations for Minimum Spanning Tree Interdiction**

Jose L Walteros, Assistant Professor, Industrial and Systems Engineering, 413 Bell Hall, University at Buffalo, Buffalo, NY, United States, 14260
Co-Authors: Ningji Wei, Industrial and Systems Engineering, Buffalo, NY, United States, 14213, Foad Mahdavi Pajouh, Boston, MA, United States

We solve the problem of removing a minimum number of edges in a weighted graph so that the weight of any spanning tree in the remaining graph is bounded below by a value r. We give two formulations and compare their strength analytically. We also study the convex hull of feasible solutions and identify several facets, as well as other polyhedral properties.

## ■ SA05

Room 023

### Optimization on Networks II

Contributed Session

Chair: Halit Uster, Engineering Management, Information and Systems, Dallas, TX, United States

**1. Centralized Optimization of Traffic Engineering (te) Tunnels in a Large ISPBackbone using an SDN Controller**

Gagan L Choudhury, Advanced Technology, Room D5-3C21, 200 Laurel Avenue, Middletown, NJ, United States, 07748
Co-Authors: Aswatnarayan Raghuram, Simon Tse, Narayan Padi, Advanced Technology, Kang Xi, Bruce Cortez, Al Goddard, Andrew Wallace, Advanced Technology, Middletown, NJ, United States,

We develop a novel heuristic optimization algorithm (99% efficiency compared to a globally optimal algorithm but much faster run time) by combining centralized and distributed Traffic Engineering protocols to efficiently run a large ISP backbone network carrying multiple traffic classes with unpredictable traffic surges and failure conditions.

**2. Managing Fluctuations in Network Level Pavement Condition an Optimization Approach**

MD Al Amin, Civil, Architectural, and Environmental Engineering, 2501 Lake Austen Blvd., Apt. M203, Austin, TX, United States, 78703

A multi period multi objective integer programming model is proposed. Two objectives maximization of average network condition and minimization of fluctuation of network condition from a smooth condition score linear function are considered. The proposed model provides a M&R for the pavement network that results in stable network conditions.

**3. New Algorithms for Fairness and Efficiency in Course Allocation**

Hoda Atef Yekta, PhD Student, Operations and Information Management, Storrs, CT, United States, 06268

This research formulates the course allocation problem as a multi objective mathematical model considering both efficiency and measures of fairness. Results of five proposed heuristic algorithms are compared with existing mechanisms and we show that our new algorithms can improve both efficiency and fairness of the results.

**4. Exact and Heuristic Algorithms for Data-gathering Wireless Sensor Network (WSN) Design Problem**

Halit Uster, Engineering Management, Information and Systems, Dallas, TX, United States
Co-Author: Hui Lin, Houston, TX, United States

We consider an integrated topology control and routing problem in WSNs to prolong time until next deployment. We suggest an MIP-based policy to be employed in each period of a deployment cycle, present efficient algorithms for its solution, and computationally examine its effectiveness in prolonging network lifetime against two other approaches.

## ■ SA06

Sherrerd

### Optimal Learning

Contributed Session

Chair: Rajarathnam Chandramouli, ECE, Hoboken, NJ, United States

**1. Accelerating Optimization via Adaptive Prediction**

Scott Yang, Mathematics, 251 Mercer Street, Room 708, New York, NY, United States, 10012
Co-Author: Mehryar Mohri, New York, NY, United States

We present a general framework for data-dependent optimization, unifying recent techniques in adaptive regularization, predictive gradients, and problem-dependent randomization. We provide flexible regret bounds that can both recover bounds of known algorithms as well as guide the design of new methods with state-of-the-art guarantees.

**2. Folded Concave Penalized Sparse Linear Regression**

Tao Yao, Penn State University, 310 Leonhard Building, University Park, PA, United States, 16802
Co-Authors: Hongcheng Liu, 1000 Escalon Avenue, Apt. D2031, Sunnyvale, CA, United States, 94085, Yinyu Ye, Stanford, CA, United States, Runze Li, University Park, PA, United States,

Concerning folded concave penalized sparse linear regression problem, we show that any local solution well approximates the true parameter. Some local solution satisfying a second-order necessary condition can perform comparably to the exponential-time but optimal estimator given enough sample size.

**3. Optimal Learning of Transient Functions**

Nana Aboagye, ORFE, 10 Lawrence Drive, Apt. 303, Princeton, NJ, United States, 08540
Co-Author: Warren B Powell, 230 Sherrerd Hall, Dept. of Operations Research and Financial Eng, Princeton, NJ, United States, 08544

We present a formerly studied algorithm, known as the knowledge gradient, for a sequential Bayesian ranking and selection problem. However, unlike previous iterations of this algorithm, we assume that the underlying truth is transient and evolves over time. We mathematically characterize its behavior and present some simulation results.

**4. Stochastic Learning in Wireless Network Optimization**

Rajarathnam Chandramouli, Thomas Hattrick Chair Professor, ECE, Hoboken, NJ, United States,

In this talk, we present a variety of wireless network optimization problems, optimization metrics and a few stochastic computational learning algorithms for solving these problems. Theoretical notions such as convergence and optimality of the learning algorithms as well as computer simulation results will be presented to illustrate these ideas.

# Saturday, 10:30am – 12:00pm

## ■ SB01

Dodds Auditorium

### Optimization and Machine Learning

Contributed Session

Chair: Ai Kagawa, 244 Loring Ave., Apt. 1A, Edison, NJ, United States, 08817

**1. Learning to Branch In Mixed Integer Programming**

Elias B Khalil, PhD Student, School of Computational Science and Engineering, 1062 Hemphill Avenue NW, Atlanta, GA, United States, 30318
Co-Authors: Le Song, Assistant Professor, School of Computational Science and Engineering, Atlanta, GA, United States, Pierre Le Bodic, Industrial and Systems Engineering, 199 14th Street NE, Apt 2107, Atlanta, GA, United States, 30309,George L Nemhauser, Industrial and Systems Engineering, Dept of Industrial and Systems Engineering, Atlanta, GA, United States, 30332-0205, Bistra Dilkina, Assistant Professor, School of Computational Science and Engineering, 266 Ferst Dr., Klaus Bldg 1304, Atlanta, GA, United States, 30332-0765

We propose a machine learning framework for branching variable selection in MIP. Our framework is instance-specific, running on-the-fly while executing a branch-and-bound search. Using a learning-to-rank formulation, we learn to imitate Strong Branching. Our method produces significantly smaller search trees than existing heuristics.

**2. Probabilistic Approach to One-class Support Vector Machine**

Vincent Leclere, CERMICS, 6-8 Avenue Blaise Pascal, Champs-sur-Marne, France, 77455
Co-Authors: Edouard Grave, Berkley, CA, United States, Laurent El Ghaoui, 421 Sutardja Dai Hall, Berkeley, CA, United States, 94703

We propose a probabilistic formulation of imbalanced classification. This formulation leads to the classical one-class SVM problem if the negative class is assumed to be scaled and centered. We apply the method to generalize one-class SVM to fast highly imbalanced classification.

**3. Tight Bounds for Approximate Caratheodory and Beyond**

Adrian Vladu, Cambridge, MA, United States, 02139
Co-Authors: Sam Chiu-wai Wong, Berkeley, CA, United States, Vahab Mirrokni, New York, NY, United States, Renato Paes Leme, New York, NY, United States

We give a deterministic nearly-linear time algorithm for approximating any point inside a convex polytope with a sparse convex combination of the polytope's vertices. Our result provides a constructive proof for the Approximate Caratheodory Problem, which states that any point inside a polytope contained in the lp ball of radius D can be approximated to within   in lp norm by a convex combination of only O(D2p/ 2) vertices of the polytope for p ≥ 2. We also show that this bound is tight, using an argument based on anti-concentration for the binomial distribution.
Along the way of establishing the upper bound, we develop a technique for minimizing norms over convex sets with complicated geometry; this is achieved by running Mirror Descent on a dual convex function obtained via Sion's Theorem. As simple extensions of our method, we then provide new algorithms for submodular function minimization and SVM training. For submodular function minimization we obtain a simplification and (provable) speed-up over Wolfe's algorithm, the method commonly found to be the fastest in practice. For SVM training, we obtain O(1/ 2) convergence for arbitrary kernels; each iteration only requires matrix-vector operations involving the kernel matrix, so we overcome the obstacle of having to explicitly store the kernel or compute its Cholesky factorization.

**4. The Rectangular Maximum Agreement Problem: Applications and Parallel Solution**

Ai Kagawa, 100 Rockafeller Road, Building A, Piscataway, NJ, United States, 08854
Co-Authors: Jonathan Eckstein, 100 Rockafeller Road, Room 5145, Piscataway, NJ, United States, 08854, Noam Goldberg, Building 605, Bar-Ilan University, Ramat Gan

The NP-hard rectangular maximum agreement (RMA) problem finds a "box" that best discriminates between two weighted datasets. Its data analysis applications include boosting classification methods and boosted regularized regression. We describe a specialized parallel branch-and-bound method for RMA.

## ■ SB02

Bowl 1

### Networks and Graphs

Contributed Session

Chair: John G. Klincewicz, Statistics and Biostatistics, Wayside, NJ, United States, 07712

**1. A Flow Based Pruning Scheme for Enumerative Equitable Coloring Algorithms**

Martin Tieves, Lehrstuhl 2 fuer Mathematik, Maxstraße 18, Aachen, Germany, 52070
Co-Authors: Robert Scheidweiler, Lehrstuhl 2 fuer Mathematik, Aachen, Germany, Arie Koster, Pontdriesch 14-16, Aachen, Germany, 52056

An equitable graph coloring is a vertex coloring, where the sizes of the color classes differ by at most one. We present a flow-based scheme for generating pruning rules for DSATUR based algorithms. Computational experiments show that significant reductions of the search tree can be achieved.

**2. An Extended Formulation of the Convex Recoloring Problem on a Tree**

Sangho Shim, Engineering, 6001 University Blvd., Moon Twp, PA, United States, 15108
Co-Authors: Sunil Chopra, Professor, Managerial Economics and Decision Sciences, 2001 Sheridan Road Suite 548, Evanston, IL, United States, 60201, Minseok Ryu, Industrial and Operations Engineering, 1205 Beal Avenue, Ann Arbor, MI, United States, 48109-2117, Kangbok Lee, Assistant Professor, Business and Economics, Jamaica, NY, United States, 11451

We introduce a strong extended formulation of the convex recoloring problem on a phylogenetic tree. The formulation has only polynomial number of constraints, but dominates the conventional formulation and previously known exponentially many valid inequalities. For real data, the formulation empirically gives integer solutions.

**3. Stochastic Graduated Graph Approximation Algorithm for MRF Optimization**

Sergei Liubich, Future Technologies, DMC R&D Center, Seoul, Korea, Republic of, 06765

Markov random fields (MRF) have been powerful models in computer vision. In this paper graduated optimization technique is applied in a novel way to develop an efficient algorithm for solving general multi-label MRF optimization problem called Stochastic Graduated Graph Approximation (SGGA) algorithm.

**4. Graph Coloring Problems in Telecommunications and Scheduling**

John G Klincewicz, Statistics and Biostatistics, New Brunswick, NJ, United States, 08901

In the Generalized Graph Coloring Problem, links have weights; each node must be given a color. One must minimize the total weight of monochromatic links, for a fixed set of colors. We describe some applications, a greedy randomized adaptive search procedure (GRASP), and computational experience.

## ■ SB03

Bowl 2

### Risk in Stochastic Optimization

Contributed Session

Chair: Matthew Norton, Gainesville, FL, United States

**1. Buffered Probability of Exceedance: Mathematical Properties and Applications**

Stan Uryasev, Professor, Inudstrial and Systems Engineering, Gainesville, FL, United States,

We investigate buffered probability of exceedance (bPOE). With bPOE, it is possible to count tail outcomes averaging to some specific threshold value. We discuss the Cash Matching problem for a Portfolio Bonds, minimizing bPOE that assets exceed liabilities.

**2. Buffered Probability of Exceedance and Applications to Machine Learning**

Matthew Norton, Gainesville, FL, United States

We present a new characterization of uncertainty called Buffered Probability of Exceedance (bPOE). In the context of machine learning, we discuss its benefits as a tool to tackle optimization problems involving a probabilistic objective function with convex and linear programming.

**3. Conservative Probability Density Estimation**

Alexander Mafusalov, Industrial and Systems Engineering, Weil Hall, 303, Gainesville, FL, United States, 32611

The multivariate density estimation problem was solved with a new multidimensional generalization (M-bPOE) of buffered probability of exceedance (bPOE). Entropy maximization with M-bPOE dominance constraints and directional variance constraints is a convex optimization problem with a maximum-of-Gaussians form of the optimal solution.

# Saturday, 1:30pm – 2:30pm

## ■ SC01

Dodds Auditorium

### Plenary: Stochastic Optimization in the Tech Sector: Yelp and Uber

Peter Frazier, Cornell University

We discuss two stochastic optimization problems from the speaker's work in the tech sector. First, we discuss Bayesian optimization of expensive functions at Yelp, with application to parameter choice in content recommendation algorithms, and to the design of websites and mobile apps. This work uses machine learning to model the objective function, and views the problem of choosing those points to sample next as its own stochastic optimization problem. Second, we discuss dynamic pricing of transportation ("surge pricing") at Uber, focusing on the rationale for dynamic pricing viewed as a problem in stochastic optimization.