How to Navigate the Technical Sessions

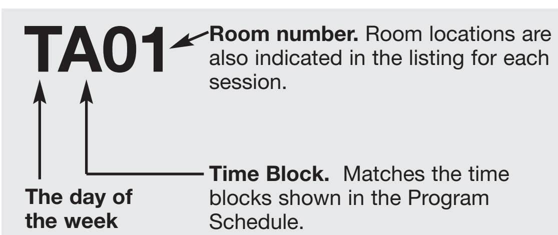
There are four primary resources to help you understand and navigate the Technical Sessions:

- 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.

Quickest Way to Find Your Own Session

The session code for your presentation will be shown along with the room location (sample below). You can also refer to the full session listing for the room location of your session.

The Session Codes



Session Time Blocks

Friday

A - 8:30am - 10:00am
B - 11:30am - 1:00pm
C - 2:30pm - 4:00pm

Saturday

A - 8:30am - 10:00am
B - 2:00pm - 3:30pm

Sunday

A - 11:00am - 12:00pm

Plenary Time Blocks

Friday

10:30am - 11:30am
4:30am - 5:30pm

Saturday

10:30am - 11:20am
11:30pm - 12:20pm
4:00pm - 5:00pm

Sunday

A - 8:30am - 9:20pm

Friday, 8:30AM - 10:00AM

■ FA01

North Classroom 1806

Simulation-Based Optimization

General Session

Chair: Sven Leyffer, Argonne National Laboratory, Argonne, IL, 60439, United States, leyffer@mcs.anl.gov

1 - Optimization of SWAPPS Simulations using Integer Surrogate Model Algorithm

Juliane Mueller, Lawrence Berkeley National Laboratory, Berkeley, CA, United States, juliane.mueller2901@gmail.com, Osman Karslioglu, Mathias Gehlmann, Hendrik Bluhm, Chuck Fadley

We present an optimization algorithm for solving a computationally expensive optimization problem arising in Standing Wave Ambient Pressure Photoelectron Spectroscopy (SWAPPS). SWAPPS has recently been shown to be a powerful tool for the investigation of solid-liquid interfaces in-situ and in operando. Analysis of solid-liquid interfaces can provide important insights into electrochemical devices such as batteries, fuel-cells and electrolyzers, as well as electrochemical processes such as corrosion. Multilayer-mirror (MLM) substrates are necessary for producing the standing-wave, but higher temperature can reduce their effectiveness. Thus, we have to find an optimal multilayer structure that is thermally stable up to a high temperature. We formulate this problem as a pure integer optimization problem that involves the simulation of experimental data using an X-ray optics simulation code. We use a derivative-free optimization algorithm that exploits surrogate models to find the optimal multilayer structure. In numerical experiments, we optimize the structure of different MLM's and we can show that the simulation agrees well with the experimental data. The optimization algorithm allows us to find better solutions more efficiently than with current state of the art methods.

2 - Derivative-free Robust Optimization by Outer Approximations

Matt Menickelly, Argonne National Laboratory, 9700 S. Cass Ave., Argonne, IL, 60439, United States, mmenickelly@anl.gov, Stefan Wild

We develop an algorithm for minimax problems that arise in robust optimization in the absence of objective function derivatives. The algorithm utilizes an extension of methods for inexact outer approximation in sampling a potentially infinite-cardinality uncertainty set. Clarke stationarity of the algorithm output is established alongside desirable features of the model-based trust-region subproblems encountered. We demonstrate the practical benefits of the algorithm on a new class of test problems.

3 - Nonsmooth Optimization via Manifold Sampling

Jeffrey Larson, Argonne National Laboratory, Argonne, IL, 60439, United States, jmlarson@anl.gov, Kamil Khan, Stefan Wild

We develop a manifold sampling algorithm for the unconstrained minimization of a nonsmooth composite function $h \circ F$ when h is a known, possibly nonconvex, piecewise-linear function and F is smooth but expensive to evaluate. This trust-region algorithm classifies evaluated points as belonging to different manifolds of the objective; since h is known, classifying points as belonging to manifolds using only the value of F is simple. This manifold information is then used when computing descent directions. We prove that all cluster points of the sequence of algorithmic iterates are Clarke stationary. This holds even when points evaluated by the algorithm are not assumed to be differentiable and when only approximate derivatives of F are available. Numerical results show that manifold sampling using zeroth-order information of F is competitive with gradient sampling algorithms that are given access to first-order information of F.

4 - Scaling Up Model-based Derivative Free Method

Liyuan Cao, Lehigh University, Bethlehem, PA, United States, lic314@lehigh.edu, Katya Scheinberg

Trust-region derivative-free method is an existing approach to optimizing black-box functions. We focus on a specific classical method that builds a quadratic model at every iteration through interpolating function values. The complexity of constructing interpolation models prevents its application to large-scale problems. We present a more scalable version of this method, where the computational complexity of each iteration is lowered by using improvements to linear algebra steps. Our method is intended to solve simulation-based problems and hyper-parameter tuning problems alike, so we assume the evaluation of objective function value is relatively costly. Hence, the method is designed to take as few sample points as possible and any option that requires mass sampling at each iteration is considered prohibitive. Due to this assumption, we devised a sample management mechanism, which carefully chooses points to evaluate and keep them as long as they are useful. Experimental results are presented to show the effectiveness our method.

FA02

North Classroom 1602

New Convexification Techniques for Nonconvex Optimization

General Session

Chair: Taotao He, Purdue University, West Lafayette, IN, 47906, United States, he135@purdue.edu

1 - Outer-product-free Sets for Polynomial Optimization

Chen Chen, The Ohio State University, 1025 Dennison Ave #406, Columbus, OH, 43201, United States, chen.8018@osu.edu, Daniel Bienstock, Gonzalo Munoz

We present new results on S-free sets for polynomial optimization, which we call outer-product-free sets. We identify families of maximal outer-product-free sets, and discuss how to generate cuts from such sets.

2 - Improved Representations of the Quadratic Linear Ordering Problem

Audrey DeVries, Clemson University, Clemson, SC, United States, adevrie@clemson.edu, Warren P Adams, Boshi Yang

The Quadratic Linear Ordering Problem is a generalization of the Linear Ordering Problem that incurs costs based on products of pairwise orderings. We give new mixed 0-1 linear forms that lead to two key results. First, we obtain a representation that uses only half the number of restrictions in the same variable space as recent work, while maintaining the linear programming relaxation strength. Second, we characterize five classes of inequalities that define the convex hull for the size 4 instance. Inequalities for smaller instances are important because they retain their validity for larger cases.

3 - A Data-Driven Distributionally Robust Bound on the Expected Optimal Value of Uncertain Mixed 0-1 Linear Programming

Guanglin Xu, University of Minnesota, WI, gxu@umn.edu, Samuel Burer

This paper studies the expected optimal value of a mixed 0-1 programming problem with uncertain objective coefficients following a joint distribution. We assume that the true distribution is not known exactly, but a set of independent samples can be observed. Using the Wasserstein metric, we construct an ambiguity set centered at the empirical distribution from the observed samples and containing the true distribution with a high statistical guarantee. The problem of interest is to investigate the bound on the expected optimal value over the Wasserstein ambiguity set. Under standard assumptions, we reformulate the problem into a copositive program, which naturally leads to a tractable semidefinite-based approximation. We compare our approach with a moment-based approach from the literature on three applications. Numerical results illustrate the effectiveness of our approach.

4 - Product Convexification a New Relaxation Framework for Nonconvex Programs

Taotao He, Purdue University, 403 W State St, West Lafayette, IN, 47906, United States, he135@purdue.edu, Mohit Tawarmalani

We develop a new relaxation that exploits function structure while convexifying a product of n functions. The function structure is encapsulated using at most d over and underestimators. We convexify the function product in the space of estimators. The separation procedure generates facet-defining inequalities in time polynomial in d. If the functions are non-negative, the concave envelope can be separated in O(n d log(d)). We extend our construction to infinite families of under and overestimators. Finally, we interpret the relaxation procedure for non-negative functions as expressing the product as a telescoping sum followed by a simple relaxation operator. We finally introduce a relaxation framework providing ways to generate the over and underestimators and propose ways to use the techniques to improve most current relaxation schemes for factorable programs.

FA03

North Classroom 1202

Risk Averse Optimization in Stochastic Programming

General Session

Chair: Saravanan Venkatachalam, Wayne State University, Detroit, MI, 48377, United States, saravanan.v@wayne.edu

1 - Generalized Risk Parity Model for Hazmat Transportation

Aleksandr Vinel, Auburn University, Auburn, AL, United States, azv0019@auburn.edu, Nasrin Mohabbati

We propose to consider a novel modeling approach to the problem of equitable distribution of risk due to hazmat shipment throughout a transportation network. Hazardous material (hazmat) transportation problem aims to ship hazmats from an origin to one or several destinations through a road network. Using a single optimal path repeatedly over time increases the risk of accidents for the population in surrounding area, which leads to the problem of achieving a risk equity over the network. In this research, risk parity concept is combined with modern risk measures in order to obtain a potential solution. The objective of the optimization model is to find multiple routes and fairly distribute the exposure to risk in the transportation network to guarantee the equal risk contribution of each route to the total risk.

2 - Stochastic Decomposition for Risk-averse Multistage Stochastic Linear Programs

Prasad Parab, PhD Candidate, Texas A&M University, 1901 Holleman Drive W, # 403, College Station, TX, 77840, United States, prasaddparab@tamu.edu, Lewis Ntaimo, Bernardo Kulnig Pagnoncelli

Risk-averse multistage stochastic linear programs (MSLPs) can be hard to solve because of their large-scale nature due to having multiple stages each with a large number of possible outcomes. This is further compounded by the risk measure being used. In this talk, we present a study of stochastic decomposition for MSLPs with quantile and deviation risk measures. The stochastic decomposition approach allows for interior sampling whereby scenarios are sampled one-at-a-time in the course of the algorithm until a termination criterion is met. Preliminary computation results will be presented.

3 - Risk Averse Model for Fleet Assignments in Airlines

Saravanan Venkatachalam, Wayne State University, 4815 Fourth St., Detroit, MI, 48377, United States, saravanan.v@wayne.edu, Sujeevraja Sanjeevi

Fleet assignment model (FAM) assigns an aircraft to scheduled flights considering demand, cost, revenue and other operational constraints. We propose a mean-risk two-stage stochastic programming approach under demand uncertainty for FAM. The mean-risk models pose computational challenges. In this talk, we present decomposition algorithms and computational results.

FA04

North Classroom 1204

Recent Advances in Two-stage Stochastic and Distributionally Robust Optimization

General Session

Chair: Manish Bansal, Virginia Tech., Blacksburg, VA, 24060, United States, bansal@vt.edu

1 - Effective Scenarios in Distributionally Robust Stochastic Programs

Guzin Bayraksan, The Ohio State University, 210 Baker Systems, 1971 Neil Avenue, Columbus, OH, 43210-1271, United States, bayraksan.1@osu.edu, Hamed Rahimian, Tito Homem-de-Mello

Traditional stochastic programs assume that the probability distribution of uncertainty is known. However, in practice, the probability distribution oftentimes is not known or cannot be accurately approximated. One way to address such distributional ambiguity is to work with distributionally robust convex stochastic programs (DRSPs), which minimize the worst-case expected cost with respect to a set of probability distributions. In this talk, we illustrate that not all, but only some scenarios might have an effect on the optimal value, and we formally define this notion for DRSPs. We also examine properties of effective scenarios. In particular, we investigate problems where the distributional ambiguity is modeled by the total variation distance with a finite number of scenarios. We propose easy-to-check conditions to identify effective and ineffective scenarios for this class of DRSPs. We also discuss a decomposition algorithm to solve the DRSP. Computational results show that identifying effective scenarios provides useful insight on the underlying uncertainties of the problem.

2 - Reassessing the State-of-the-Art in Stochastic Unit Commitment Solvers

Jean-Paul Watson, Distinguished Member of Technical Staff, Sandia National Laboratories, P.O. Box 5800, MS 1326, Albuquerque, NM, 87185, United States, jwatson@sandia.gov, David L Woodruff

A variety of advances in both mixed-integer formulations and decomposition strategies over the past 5 years have yielded significant improvements in the performance of solvers for stochastic unit commitment, a variant of a core power systems operations problem in which renewables and load are explicitly treated as uncertain quantities. In this talk, we review key advances, and provide an empirical assessment of the state-of-the-art for both benchmarks and solution algorithms, the latter being focused on the scenario-based decomposition progressive hedging algorithm. The talk focuses on the relationship between scenario structure, specifically for renewables, and overall problem difficulty. We further describe performance on a range of new publicly available benchmarks.

3 - Decomposition Algorithms for Two-stage Distributionally Robust Mixed Binary Programs

Manish Bansal, Assistant Professor, Virginia Tech., 227 Durham Hall, 1145 Perry Street, Apt 307, Blacksburg, VA, 24060, United States, bansal@vt.edu

We present first decomposition algorithm to solve two-stage distributionally robust mixed binary problems (TSDR-MBPs) where the random parameters follow the worst-case distribution belonging to an uncertainty set of probability distributions. We investigate conditions and families of uncertainty set for which our algorithm is finitely convergent. In addition, we present a cutting surface algorithm to solve TSDR-MBPs. We computationally evaluate the performance of our algorithms in solving distributionally robust versions of a few instances from the Stochastic Integer Programming Library, in particular stochastic server location and stochastic multiple binary knapsack problem instances.

■ FA05

North Classroom 2001

Optimization on Graphs

General Session

Chair: Eugene Lykhovyd, Texas A&M University, 3131 TAMU, College Station, TX, 77843, United States, lykhovyd@tamu.edu

1 - Solving the Maximum Independent Union of Cliques Problem via Polynomial Binary Optimization

Seyedmohammadhossein Hosseinian, Texas A&M University, 3131 TAMU, College Station, TX, 77843-3131, United States, hosseinian@tamu.edu, Sergiy Butenko

Given a simple and undirected graph, the maximum Independent Union of Cliques (IUC) problem is to find a subset of vertices such that every connected component of the induced subgraph is complete. This problem is equivalent to the maximum -cluster problem without connectivity constraint for =1, which has application in graph clustering and social network analysis. It is known that the maximum clique problem can be formulated as the maximum 1-cluster problem, hence the maximum IUC problem can also be viewed as a relaxation of the maximum clique problem. We consider an unconstrained polynomial formulation of this problem, and present a heuristic approach based on the new formulation. Result of numerical experiments are also presented.

2 - Network Connectivity Assessment and Enhancement

Maggie Cheng, PhD, New Jersey Institute of Technology, Newark, NJ, United States, maggie.cheng@njit.edu

We consider the network connectivity problem in a wireless ad hoc network. Network connectivity is measured by the conductance of the network, also called the Cheeger constant of the graph. A partition algorithm based on this measure is developed that divides the network at the bottleneck area. After the network is bisected, a relay node may be deployed between the two parts to increase the conductance of the network. The relay node deployment problem is formulated as an integer linear program to maximize the number of connections between the nodes on the two sides of the cut, and then a convex optimization algorithm is used to find the precise location of the relay node, which is within the convex hull defined by the radio transmission ranges of all the nodes that can connect to the relay node. The relay node significantly relieves the bottleneck, and the graph connectivity measured by other metrics such as the widely used Fiedler value are also increased.

3 - Independent Unions in Uniform Random Graphs

Eugene Lykhovyd, Texas A&M University, 3131 TAMU, Industrial Engineering, College Station, TX, 77843, United States, lykhovyd@tamu.edu, Sergiy Butenko

The independent union in a simple, undirected graph is a subset of vertices where every connected component satisfies a certain property, e.g., Independent Union of Cliques is a subset of vertices where every connected component is a clique. We study the maximum independent union problem in uniform random graphs. We provide criteria on a property such that the maximum subgraph has a logarithmic upper bound. This result extends the classical result about the size of maximum clique in uniform random graphs.

■ FA06

North Classroom 2002

Optimization on Social Networks

General Session

Chair: Rui Zhang, University of Colorado Boulder, Boulder, CO, 80309, United States, rui.zhang@colorado.edu

1 - An Upper-bounding Technique for the Maximum Quasi-clique Problem

Balabhaskar Balasundaram, Associate Professor, Oklahoma State University, 322 Engineering North, Stillwater, OK, 74078, United States, baski@okstate.edu, Zhuqi Miao

A quasi-clique is a density-based clique relaxation, defined as a subset of vertices that induce a subgraph with edge density at least gamma belonging to the interval [0,1]. If gamma equals one, this definition corresponds to a classical clique. If gamma is strictly less than one, it relaxes the requirement of all possible edges in the clique definition. Quasi-cliques have been used to detect dense clusters in graph-based data mining, especially in large-scale, error-prone data sets in which the clique model can be overly restrictive. The maximum quasi-clique problem, which seeks a quasi-clique of maximum cardinality from a given graph, can be formulated as a mathematical program with a linear objective function and a single quadratic constraint in binary variables. This talk will discuss a Lagrangian dual based upper-bounding technique. The tightness of the bounds that result will be compared to what can be obtained using linear mixed-integer programming reformulations of the aforementioned quadratically constrained formulation. The implications for solving the maximum quasi-clique problem will also be discussed.

2 - Algorithms for the Maximum Edge Weight Clique Problem

Dalila Martins Fontes, Universidade do Porto, Porto, Portugal, fontes@fep.up.pt, Sergiy Butenko, Seyed Mohammadreza Hosseinian

Given an edge-weighted graph, the maximum edge weight clique (MEWC) problem is to find a clique that maximizes the sum of edge weights within the corresponding complete subgraph. This problem generalizes the classical maximum clique problem and finds many real-world applications. We propose exact and heuristic algorithms for the MEWC problem and report the results of numerical experiments with the proposed approaches.

3 - Chance-constrained Combinatorial Optimization with a Probability Oracle and its Application to Probabilistic Set Covering

Simge Kucukyavuz, University of Washington, Box 352650, Industrial & Systems Engineering, Seattle, WA, 98195, United States, simge@uw.edu, Hao-Hsiang Wu

A chance-constrained combinatorial optimization problem (CCOP) aims to find a minimum cost selection of binary decisions, which satisfies a constraint with high probability. Suppose that we have an oracle that can compute the probability of satisfying the constraint exactly. Using this oracle, we propose a general method for solving CCOP chance-constrained problem exactly. In addition, if CCOP is solved by a sampling-based approach, the oracle can be used as a tool for checking and fixing the feasibility of the resulting solution. To demonstrate the effectiveness of the proposed methods, we apply them to an NP-hard probabilistic set covering problem motivated by a problem in social networks, which admits a polynomial-time exact probability oracle.

4 - Generalizations of the Dominating Set Problem on Social Networks

Rui Zhang, University of Colorado Boulder, Boulder, CO, 80309, United States, rui.zhang@colorado.edu

The positive influence dominating set problem is a generalization of the dominating set problem that arises on social networks. First, we show that it can be solved in linear time on trees. Next, we provide a tight and compact extended formulation, and derive a complete description of its polytope on trees. The formulation is also valid on general graphs, thus providing a new and stronger one. Facet defining conditions for the new inequalities are provided. A computational study is conducted.

FA07

North Classroom 3004

Underground Mining Applications

General Session

Chair: Alexandra M. Newman, Colorado School of Mines, Golden, CO, 80401, United States, anewman@mines.edu

1 - Production Scheduling in Underground Mining Operations Incorporating Heat Loads

Oluwaseun Babatunde Ogunmodede, Colorado School of Mines, 7216 Winter Ridge Drive, Castle Pines, CO, 80108, United States, setotravel@gmail.com

Mine production scheduling determines when, if ever, notional three-dimensional blocks of ore should be extracted. The accumulation of heat in the tunnels where operators are extracting ore is a major consideration when designing a ventilation system and, often, the production scheduling and ventilation decisions are not made in concert. Rather, heat limitations are largely ignored despite contributors to heat output such as: (i) the equipment used for underground activities, e.g., development, extraction, and backfilling; (ii) auto compression; (iii) broken rock; and (iv) strata rock. We incorporate heat considerations into a resource-constrained production scheduling model using knapsack constraints. The model maximizes net present value subject to additional constraints on precedence and other limiting factors — mill and extraction capacities. The model produces more realistic schedules that could increase revenue by lowering ventilation costs for the mine— specifically, refrigeration costs influenced by the location of fans throughout the mine.

2 - Optimizing Truck Dispatching Decisions in Open-pit Mining using Integer Programming

Amanda Smith, University of Wisconsin-Madison, 1513 University Ave, Madison, WI, 53706, United States, amanda.smith@wisc.edu, Jeff T Linderoth, James Luedtke

We propose a novel approach to the open-pit mining truck dispatching problem that employs mixed-integer programming (MIP). The truck dispatching problem seeks to determine how trucks should be routed through the mine as they become available. Among the challenges of the dispatching problem is the need to make decisions in real-time for the constantly changing system. In addition, the dispatching problem attempts to balance the distinct (and potentially competing) objectives of meeting processing and mining targets and maintaining grade targets at the processing sites. We propose an optimization-driven approach to solving the dispatching problem in the form of a MIP model. The model is difficult to solve directly within time constraints due to its large size. Therefore, we propose a heuristic algorithm to quickly produce high quality feasible solutions to the model. We also propose two competing dispatching policies that match dispatching decisions to average flow rate targets obtained from a nonlinear flow-rate model. To evaluate the dispatching policies, we embed them in a discrete-event simulation of an open-pit mine. We conclude with computational results demonstrating how each policy performs on open-pit mines with different characteristics.

3 - Optimal Transition From Top-down to Bottom-up Mining

Peter Nesbitt, Major, United Stated Army, Golden, CO, United States, nesbitt@mymail.mines.edu, Levente Sipeki

Underground mining requires complex sequencing of activities that depend on the depth of the deposit, distribution of the ore, and the mining method(s). For deposits that extend deep underground, the mining method most profitable at the highest vertical level may not be the most profitable at the deepest level. We consider two mining methods- top-down open stoping (in which mining starts at the highest vertical level of extraction, but the percentage of extraction decreases with each vertical level) and bottom-up stoping with backfilling (in which the percentage of extraction remains favorably high, but excavation begins later at the deepest level first). These methods offer competing advantages regarding time to extraction and overall extraction rate. However, there is no industry standard regarding which method maximizes profitability over the life of the mine. We present an optimization-based heuristic that determines a schedule, including method of extraction, while incorporating the rules on viable extraction sequences. Our methodology consists of (i) preprocessing, (ii) integer programming, and (iii) a heuristic. This hybrid mining approach benefits mining companies by better informing strategic plans with regard to higher expected value.

FA08

North Classroom 3210

Health Care

Contributed Session

Chair: Nalan Gulpinar, Warwick Business School, Scarman Road, Coventry, CV4 7AL, United Kingdom, nalan.gulpinar@ubs.ac.uk

1 - A Methodology to Schedule Combination Appointments for a System of Facilities: A Case Study of a Radiology Department

Papiya Bhattacharjee, Xavier Institute of Management Bhubaneswar, Studio Apartment 112, Xavier Institute of Management, Xavier Square, Jayadev Vihar, Bhubaneswar, 751013, India, papiya2506@gmail.com

In service systems with multiple facilities, such as hospitals, customers/patients often seek appointments for more than one facility on the same day. However, the scheduling of appointments in each of these facilities is usually performed in isolation, without considering interactions among the facilities and the overall performance of the system and the satisfaction of patients. The research work presented in this paper makes an attempt to identify the appropriate scheduling rules while assigning appointment times to patients seeking appointments for more than one facility (in this paper, such appointments are referred to as combination appointments). Based on the evaluation of a number of scheduling alternatives through simulation experiments, this paper proposes a methodology for scheduling combination appointments. The methodology is illustrated with the help of a case study of a Radiology department where the combination appointments for some of the facilities is very common.

2 - Paraspinal Muscle Segmentation in CT and MR Images using Interval Methods

Mengyi Ying, Assistant Professor, University of North Georgia, Newton Oakes Ctr.212D,82 College Circle, Dahlonega, GA, 30597, United States, mying@ung.edu

Paraspinal muscles support the spine and are the source of movement force. The size, shape, density and volume of the paraspinal muscles cross section area (CSA) are affected by many factors, such as age, health condition, exercise, and low back pain. It is invaluable to segment the paraspinal muscle regions in images in order to measure and study them. Manual measurements of paraspinal muscle CSA and volume in CT and MR images are inaccurate and time consuming. In this proposed project, the Interval branch and bound methods (Interval methods) will be implemented by a mutual information-driven atlas-based image registration algorithm in order to segment the muscle regions in CT and MR images automatically. The Interval methods will be the add-on acceleration strategies to target a better solution, reduce the time consumed and the use of computer CPU memory. The optimal results can be used to evaluate tissue injury and post-operative back muscle atrophy of patients. Further applications of the algorithm obtained in the study can facilitate diagnosis of spine related problems of old patients and/or patients with low back pain.

3 - Restructuring Existing Ward Structures at a Public Hospital

Nalan Gulpinar, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, United Kingdom, nalan.gulpinar@wbs.ac.uk, Sebastian Rachuba, Elvan Gokalp

Ward structures and sizes at hospitals have grown historically and only few changes are made to the allocation of beds to wards and departments. Although restructuring of existing allocations appears to be promising towards a more efficient bed utilisation, the resulting changes cause disruption and would affect daily working practice. With more changes to an existing allocation, an increasing amount of organizational changes will become necessary which will likely affect the effectiveness of service delivery, although cost savings might be significant. In this paper, we present a two-stage stochastic optimisation model for the ward restructuring problem at a public hospital. The effects of restructuring are demonstrated in a case study using real world data. We analyse the trade-offs between restructuring effort, the resulting costs and the quality of new structures. Finally, both operational and managerial implications of our research are discussed.

■ FA09

Student Commons 1500

Statistical Optimization: Computational and Statistical Tradeoffs

General Session

Chair: Hongcheng Liu, PhD, University of Florida, Gainesville, FL, 32611, United States, liu.h@ufl.edu

1 - High Dimensional Statistical Learning Problem with Grouped Folded Concave Penalty

Xue Wang, Pennsylvania State University, 310 S Barnard Street, 240 Leonhard, University Park, PA, 16802, United States, wxie91@gmail.com, Yao Tao, Hongcheng Liu

We consider the problem of selecting grouped variables (factors) for accurate prediction in the high-dimensional statistical learning problem. Such a problem arises naturally in many practical situations. The traditional approach involves Lasso type penalty and may yield a solution with the sub-optimal statistical performance. In this paper, we replace the Lasso type penalty with grouped folded concave penalty. We show that under mild conditions our model can guarantee the solution with the optimal statistical performance theoretically. We also develop an efficient coordinate descent algorithm as the solution scheme. Numerical tests validate the effectiveness of our approach.

2 - Exponential Error Rates of Semidefinite Programming for Block Models

Yingjie Fei, Cornell University, Ithaca, NY, United States, yf275@cornell.edu, Yudong Chen

We consider the community detection problem under the Stochastic and Censored Block Models. We show that the semidefinite programming (SDP) formulation for this problem achieves an error rate that decays exponentially in the signal-to-noise ratio. Significantly, even though we are estimating a combinatorial structure by solving a continuous optimization problem, this error rate is achieved by the SDP itself without any further pre/post-processing or rounding. Our results improve upon existing polynomially-decaying error bounds obtained via the Grothendieck's inequality. The analysis highlights the implicit regularization effect of the SDP, and its robustness in the sparse graph regime. If time permits, we will also discuss our recent results on the Gaussian Mixture Model.

3 - Accelerated Cubic Regularization of Sub-Sampled Newton Method for Sum-of-Nonconvex Optimization

Tianyi Lin, University of California, Berkeley, 407, Cornell Avenue, Unit 11, Albany, CA, 94706, United States, darren.august.ty@gmail.com, Xi Chen, Bo Jiang, Shuzhong Zhang

We consider variants of adaptive cubic regularization methods for sum-of-non-convex optimization, in which the Hessian matrix is approximated via random sampling strategies. Under the assumption that the sub-problems can be solved approximately, we establish overall iteration complexity bounds for two newly proposed algorithms to obtain an ϵ-optimal solution. In specific, we show that the standard and accelerated adaptive cubic regularization method achieves an iteration complexity in the order of $O(1/\epsilon^2)$ and $O(1/\epsilon^3)$ which matches that of the original corresponding cubic regularization method assuming the availability of the exact Hessian information and the Lipschitz constants, and the global solution of the sub-problems. The statistical sampling complexity is also provided under both uniform and non-uniform sub-sampling strategies. Our numerical experiment results show a clear effect of acceleration displayed in the adaptive Newton's method with cubic regularization on a set of regularized logistic regression instances.

4 - Second-order Necessary Optimality Condition and Its Implications in High-dimensional Statistical Learning with Linear Constraints

Hongcheng Liu, University of Florida, Gainesville, FL, 32611, United States, liu.h@ufl.edu

This presentation is concerned with the computational and statistical complexities involved in the folded concave penalized high-dimensional M-estimation problems with linear constraints. We show that certain local solutions which satisfy a certain second-order necessary optimality condition yield desirable statistical performance; exponential growth of dimensionality can be compensated by polynomially increasing the sample size. The desired local solutions admit a pseudo-polynomial-time algorithm that only employs partial information on the Hessian matrix.

■ FA10

Student Commons 1600

Optimization for Statistical Learning

General Session

Chair: Sam Davanloo Tajbakhsh, Ohio State University, Columbus, OH, 43210, United States, sam.tajbakhsh@gmail.com

1 - Factor Analysis via Modern Optimization

Rahul Mazumder, Massachusetts Institute of Technology, Sloan School of Management, 100 Main Street, Cambridge, MA, 02139, United States, rahulmaz@mit.edu

Factor analysis, a classical multivariate statistical technique is popularly used as a fundamental tool for dimensionality reduction in statistics, econometrics and related disciplines. Estimation is often carried out via the Maximum Likelihood (ML) method, which seeks to estimate the best positive definite covariance matrix that can be decomposed as the sum of a low-rank positive semidefinite matrix and a diagonal matrix with positive entries. The associated optimization problem, central to ML is a difficult nonconvex optimization problem with rank constraints; and little is known about its computational propertie.This motivates our investigation. We reformulate the low-rank Factor Analysis problem as a nonlinear nonsmooth Semidefinite Optimization problem. We study various structural properties of this problem and propose fast and scalable algorithms employing techniques in modern nonlinear optimization to obtain high quality solutions for the task of ML estimation. Our approach has computational guarantees, gracefully scales to large problems, is applicable to situations where the sample covariance matrix is rank deficient and adapts to variants of the ML problem with additional constraints on the problem parameters. Our numerical experiments validate the usefulness of our approach over existing state-of-the-art approaches for low-rank ML.

2 - Taming Nonconvexity with Data

Zhaoran Wang, Northwestern University, Evanston, IL, United States, zhaoranwang@gmail.com

Nonconvex optimization is becoming one of the most powerful workhorses of data science and artificial intelligence. Compared with convex optimization, it enjoys superior statistical accuracy, computational efficiency, and modeling flexibility in numerous modern settings. However, the empirical success of nonconvex optimization largely eludes the reach of classical statistical and optimization theory, which prohibits us from designing more efficient algorithms in a principled manner. In this talk, I will illustrate how statistical thinking enables us to harness the power of nonconvex optimization. In specific, I will present an algorithmic framework for exploiting the latent geometry induced by the randomness of data. By integrating three new global exploration meta-algorithms — namely, homotopy continuation, tightening after relaxation, and noise regularization — with local search heuristics — such as the variants of gradient descent — this unified framework leads to new nonconvex optimization algorithms for a broad variety of challenging learning problems. In particular, these algorithms enjoy provably optimal statistical accuracy and computational efficiency, and moreover, lead to new scientific discoveries. Time permitting, I will discuss an interesting "more data, less computation" phenomenon, which arises from nonconvex optimization, but generalizes to even more algorithms.

3 - Solving Large Scale Linear Programs in Adaptive Enrichment Clinical Trial Design

Hongbo Dong, Washington State University, Pullman, WA, 99163-5142, United States,

We consider the approach of adaptive enrichment design for clinical trial design, an especially flexible trial design framework that allows trial administrators to change patient enrollment rules during the trials. A recent work (Rosenblum, Fang and Liu, 2017) shows substantial gain (in statistical as well as economic aspects) by applying optimization techniques in adaptive enrichment design. A large scale linear programming (LP) model is formulated to solve the design problem with two planning stages and two subpopulations. We develop specialized algorithms for this LP model exploiting its dual formulation. Preliminary computational results showing significant speed-up over off-the-shelf standard solvers. We will discuss approaches for solving problems in even larger scales with more planning stages and subpopulations. "

■ FA11

Student Commons 2500

Advances in Mixed-Integer Formulations

General Session

Chair: Joey Huchette, MIT, Cambridge, MA, 02139, United States, huchette@mit.edu

1 - Extended Formulations and Outer Approximation in Mixed-integer Nonlinear Optimization

Christopher D. Coey, MIT, 10 Magazine Street, Cambridge, MA, 02139, United States, coey@mit.edu, Juan Pablo Vielma

We present some developments in mixed-integer convex conic optimization and showcase our open-source solver Pajarito (joint work with Miles Lubin). Outside the conic regime, we present "symmetric" extended formulations for factorable nonlinear problems, and demonstrate how these may be useful in nonconvex global optimization.

2 - BDD-based Cutting-plane Algorithms for Solving Integer Programs

Cole Smith, Clemson University, 110 Freeman Hall, Clemson, SC, 29634, United States, jcsmith@clemson.edu, Leonardo Lozano, David Bergman

In this talk, we examine a range of modeling approaches that decompose an optimization problem into several interdependent binary decision diagrams (BDDs). In the applications we consider, solving the original optimization problem is tantamount to solving a so-called consistent shortest path problem (CSSP) over the BDDs. The CSSP seeks a set of paths across the BDDs whose total cost is minimized, and which satisfy certain "consistency" side constraints that model consistent variable choices in the original variable space. We first discuss the complexity of the CSSP problem itself, showing that in most cases the problem is strongly NP-hard, but that the problem becomes polynomially solvable in an important special case. For the polynomially solvable case, we derive the convex hull representation of the problem in the original variable space, and use the inequalities from that case to derive an effective cutting-plane approach to solving integer programs amenable to decomposition by BDDs. We close by briefly discussing the efficacy of this approach on two problem classes.

3 - Nonconvex Piecewise Linear Functions: Advanced Formulations and Simple Modeling Tools

Joey Huchette, MIT, 77 Massachusetts Avenue, Cambridge, MA, 02139, United States, huchette@mit.edu

Piecewise linear functions are a central modeling primitive throughout optimization. In this work, we present novel mixed-integer programming (MIP) formulations for (nonconvex) piecewise linear functions. Leveraging recent advances in the systematic construction of MIP formulations for disjunctive sets, we derive new formulations for univariate functions using a geometric approach, and for bivariate functions using a combinatorial approach. All formulations derived are small (logarithmic in the number of piecewise segments of the function domain) and strong, and we present extensive computational experiments in which they offer substantial computational performance gains over existing approaches.

■ FA12

Student Commons 2504

Integer Programming

General Session

Chair: Akshay Gupte, Clemson University, Clemson, SC, 29634-0975, United States, agupte@clemson.edu

1 - Time-indexed Relaxations for the Online Bipartite Matching Problem

Alejandro Toriello, Georgia Institute of Technology, 765 Ferst Drive NW, Atlanta, GA, 30332, United States, atoriello3@isye.gatech.edu, Alfredo Torrico

We study the i.i.d. online bipartite matching problem where one side of the bipartition is fixed and known in advance, while nodes from the other side appear one at a time as i.i.d. realizations of an underlying distribution, and must immediately be matched or discarded. We consider time-indexed relaxations of the set of achievable matching probabilities, introduce complete subgraph inequalities, show how they theoretically dominate inequalities from a lower-dimensional relaxation presented in previous work, and discuss when they are facet-defining. We finally present a computational study to demonstrate the empirical quality of the new relaxations and the heuristic policies they imply.

2 - Scenario Aggregation using Binary Decision Diagrams for Stochastic Programs with Endogenous Uncertainty

Carla Michini, University of Wisconsin Madison, 330 North Orchard Street, Madison, WI, 53715, United States, michini@wisc.edu, Utz-Uwe Haus, Marco Laumanns

Modeling decision-dependent scenario probabilities in stochastic programs is difficult and typically leads to large and highly non-linear models that are difficult to solve to optimality. We develop a new approach to obtain a compact representation of the recourse function using binary decision diagrams (BDDs). The resulting BDDs are then used to efficiently characterize the decision-dependent scenario probabilities by a set of linear inequalities. This construction corresponds to a factorization of the probability distribution and yields a reformulation of the stochastic program as a mixed 0-1 linear program, whose size (number of variables and constraints) is proportional to the size of the BDD (number of nodes). The approach is applicable to a large class of stochastic programs with multivariate binary scenario sets, such as stochastic network design, network reliability, and stochastic network interdiction problems. Computational results show that the BDD-based scenario representation reduces the size of the MILP reformulation significantly compared to previous approaches.

3 - Approximate Positive Correlated Distributions and Approximation Algorithms for D-optimal Design

Weijun Xie, Virginia Tech, Blacksburg, VA, 24060, United States, wxie@vt.edu, Mohit Singh

Experimental design is a classical problem in statistics and has also found new applications in machine learning. In the experimental design problem, the aim is to estimate an unknown vector x in m-dimensions from linear measurements where a Gaussian noise is introduced in each measurement. The goal is to pick k out of the given n experiments so as to make the most accurate estimate of the unknown parameter x. Given a set S of chosen experiments, the most likelihood estimate x' can be obtained by a least squares computation. One of the robust measures of error estimation is the D-optimality criterion which aims to minimize the generalized variance of the estimator. This corresponds to minimizing the volume of the standard confidence ellipsoid for the estimation error x-x'. The problem gives rise to two natural variants depending on whether repetitions are allowed or not. The latter variant, while being more general, has also found applications in the geographical location of sensors. In this work, we first show that a 1/e-approximation for the D-optimal design problem with and without repetitions giving us the first constant factor approximation for the problem. We also consider the case when the number of experiments chosen is much larger than the dimension of the measurements and provide an asymptotically optimal approximation algorithm.

4 - Optimization Bounds from the Branching Dual

Gerdus Benade, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213, United States, jbenade@andrew.cmu.edu

We present a general method for obtaining strong optimization bounds based on a concept of branching duality. It can be applied to combinatorial problems for which no useful integer programming model is available, and we illustrate this with the minimum bandwidth problem. The method strengthens a known bound for a given problem by formulating a dual problem whose feasible solutions are partial branching trees. We solve the dual problem with a ``worst-bound" local search heuristic that explores neighboring partial trees. After establishing general properties of the heuristic, we show that it substantially improves known combinatorial bounds for the minimum bandwidth problem with a modest amount of computation.

■ FA13

Student Commons 2600

Distributed Optimization and Large Scale Machine Learning

General Session

Chair: Peter Richtarik, King Abdullah University of Science and Technology, 6317 James Clerk Maxwell Building, Jeddah, Saudi Arabia, Peter.Richtarik@kaust.edu.sa,

1 - Coordinate Descent Methods

Mert Gurbuzbalaban, Rutgers University, 100 Rockafeller Rd, Piscataway, NJ, 08854, United States, mg1366@business.rutgers.edu

The coordinate descent (CD) methods have seen a resurgence of recent interest because of their applicability in machine learning as well as large scale data analysis and superior empirical performance. CD methods have two variants, cyclic coordinate descent (CCD) and randomized coordinate descent (RCD) which are deterministic and randomized versions of the CD methods. In light of the recent results in the literature, there is a large gap between the theory and practice of CCD methods, in particular existing theoretical guarantees for CCD are far worse than RCD despite the fact that CCD works often well in practice. In the first part of the talk, we provide problem classes in quadratic optimization for which CCD (or CD with any deterministic order) is provably faster than RCD in terms of asymptotic worst-case convergence and we quantify the improvement compared to RCD depending on the order chosen to update the coordinates in CCD. We also provide some lower bounds on the performance of CCD by characterizing the best order and discuss applications to the Kaczmarz method for solving Laplacian linear systems and spectral clustering problems over graphs. In the second part of the talk, we develop a distributed algorithm for computing effective resistances in an undirected graph based on the Kaczmarz method. The effective resistance between a pair of nodes in a weighted undirected graph is a key quantity of interest in many applications, defined as the potential difference induced between them when a unit current is injected at the first node and extracted at the second node, treating edge weights as the conductance values of edges. We show that our algorithm is efficient in terms of local communications with a global linear rate and discuss applications to continuous-time averaging over networks and solving Laplacian linear systems.

2 - More Iterations Per Second Same Quality Why Asynchronous Algorithms May Drastically Outperform Traditional Ones

Robert Hannah, UCLA, Los Angeles, CA, United States, RobertHannah89@math.ucla.edu, Wotao Yin

In this talk, we present the first strong theoretical evidence that asynchronous-parallel algorithms may drastically outperform traditional ones. Asynchronous algorithms overcome costly synchronization penalty by having nodes compute updates with the most recent data available, rather than waiting for all nodes to complete their computation. However, this means that the solution vector is updated with outdated information. Until now, how many additional iterations that asynchronous algorithms need to compensate for this outdatedness has been an open question. We prove our thesis with a series of results. We first use renewal theory to show how many factors cause asynchronous algorithms to complete “more iterations per second”. We then prove the first sharp iteration complexity results for a variety of synchronous algorithms (including randomized block gradient descent, block proximal gradient, etc.) so that we can make a fair comparison to asynchronous algorithms. Finally, we prove that the iteration complexity of the asynchronous counterparts of these algorithms is only negligibly higher, meaning asynchronous iterations are of the “same quality”. Taking these factors together, we can conclude that many asynchronous algorithms may drastically outperform traditional ones. We also present a similar result for accelerated block gradient descent.

3 - Asynchronous Distributed Optimization with Unbounded Delays

Konstantin Mishchenko, King Abdullah University of Science and Technology, Thuwal, Jeddah, 23955, Saudi Arabia, konstantin.mishchenko@kaust.edu.sa, Jerome Malick, Franck Iutzeler

We develop and analyze a distributed algorithm for solving convex learning optimization problems. This algorithm is particularly well-suited for large-scale machine learning as it relies on shared-nothing memories and asynchronous communications with a master machine. Unlike many existing methods, the algorithm is adjustable to various levels of communication cost, machines computational power, and data distribution evenness. A unique feature is that the learning rate does not depend on communication delays nor number of machines, which is highly desirable for scalability. We prove that the algorithm converges in general convex case with linear speedup under strong convexity assumption.

Friday, 10:30AM - 11:20AM

■ Plenary I

Student Commons 2600

Non-convex Non-optimization

Plenary Session

1 - Non-convex Non-optimization

Moritz Hardt, Google and University of California, Berkeley, Berkeley, CA, 9, United States, hardt@berkeley.edu

Recent progress in machine learning highlights at least two conceptual and technical challenges for optimization research. One is to understand how different model architectures—-equivalent in expressivity—-affect the difficulty of non-convex optimization. The other is how external criteria, such as out-of-sample generalization, interact with the choice of optimization algorithm. We discuss some emerging theory addressing these intriguing problems, while emphasizing the many challenges that remain.

Friday, 11:30AM - 1:00PM

■ FB01

North Classroom 1806

New Advances in Stochastic and Nonlinear Optimization

General Session

Chair: Uday Shanbhag, Pennsylvania State University, University Park, PA, 16802, United States, udaybag@psu.edu

1 - Dynamic Stochastic Approximation for Multi-stage Stochastic Optimization

Zhiqiang Zhou, ISyE GA Tech, 755 Ferst Drive, NW, Atlanta, GA, 30332, United States, zzhou71@gatech.edu, Guanghui Lan

In this talk, we consider multi-stage stochastic optimization problems with convex objectives and conic constraints at each stage. We present a new stochastic first-order method, namely the dynamic stochastic approximation (DSA) algorithm, for solving these types of stochastic optimization problems. We show that DSA can achieve an optimal O (1 /\epsilon^4) rate of convergence in terms of the total number of required scenarios when applied to a three-stage stochastic optimization problem. We further show that this rate of convergence can be improved to O (1 /\epsilon^2) when the objective function is strongly convex. We also discuss variants of DSA for solving more general multi-stage stochastic optimization problems with the number of stages T > 3. The developed DSA algorithms only need to go through the scenario tree once in order to compute an \epsilon-solution of the multi-stage stochastic optimization problem. As a result, the memory required by DSA only grows linearly with respect to the number of stages. To the best of our knowledge, this is the first time that stochastic approximation type methods are generalized for multi-stage stochastic optimization with T ≥ 3.

2 - A Randomized Block Proximal Stochastic Gradient Method for Nonconvex Nonsmooth Stochastic Optimization

Jinlong Lei, Pennsylvania State University, State College, PA, United States, jxl800@psu.edu, Uday V. Shanbhag

This work considers the minimization a nonconvex nonsmooth function, which is the sum of a expectation-valued smooth possibly nonconvex function and a separable convex nonsmooth regularizer. We design a randomized block proximal stochastic gradient algorithm, in which at each iteration, a single block is randomly chosen toinitiate an update while the other blocks keep invariant. The selected block updates its estimates bya novel proximal stochastic gradient method, in which a variable number of sampled gradients is used. By appropriately chosen sample sizes, we prove that the iterates converge almost surelyto the stationary points. Besides, we show that the mean-squared error of the gradient mapping metric is $\mathcal{O}(1/K) $ when the scheme is terminated in K steps and is $\mathcal{O}(1/\sqrt{M})$ when thesimulation budget of the sampled gradients is M. Furthermore, under the proximal Polyak-Lojasiewicz condition, when the variable sample size increases at a suitable selected geometric rate, we prove that the expectation of the optimality gap function diminishes at a geometric rate, which is the optimal deterministic rate.Besides, we show that the iteration complexities of the scheme in terms of the number of proximal operators and the sampled gradients for computing an ϵ-accurate optimal solution are $\mathcal{O}(\log(1/\epsilon))$ and$\mathcal{O}\left((1/\epsilon)^{1+c} \right)$, respectively, in which $c=0$ when $n=1$ leads to the {\em optimal bound $\mathcal{O} (1/\epsilon) $}, and $c>0$ when $n\geq 2$represents the positive cost of multiple blocks. Concerning the special case with a single block, the expectation-valued optimality gap is shown todiminish at a geometric rate with the number of proximal operator and achievethe optimal rate $\mathcal{O}({ 1/ M})$ in terms of the total simulation budget $M.$In addition, we consider the cyclic block-coordinate proximal stochastic gradient method and prove its almost sure convergencel.

3 - Obtaining Deterministic Rates in Stochastic ADMM Schemes via a Variable Sample-size Technique

Yue Xie, Pennsylvania State University, 445 Waupelani Dr, State College, PA, 16801, United States, yux111@psu.edu, Uday Shanbhag

We propose a stochastic variable sample-size ADMM scheme to resolve min{E[f(x,\xi)] + g(y) | Ax + By = b}. In this framework, at every step, the Lagrangian subproblem is constructed using an increasing sample size Nk. This allows for showing that the gap of convergence rate is closed between stochastic ADMM schemes and its deterministic counterpart. Furthermore, there is little or no loss in statistical accuracy. Notably, such avenues provide significant computational benefit over their single sample counterparts (Nk = 1, for all k).

4 - Distributed Asynchronous Nonconvex Multiagent Optimization

Gesualdo Scutari, Purdue University, 384 Grissom Hall, 315 N. Grant Street, West Lafayette, IN, 47907, United States, gscutari@purdue.edu, Ying Sun, Ye Tian, Bin Du

We study distributed (constrained) nonconvex optimization over networks, modeled as time-invariant directed graphs. We introduced a novel asynchronous algorithmic framework whereby agents can compute and communicate independently at different times while using possibly outdated information. This reduces the waiting time for the slowest agent or longest communication delay and also eliminates the need for a global clock. The algorithm leverages a novel tracking mechanism aiming at estimating locally the average of agents' gradients, which is robust against asynchrony, outdated information, and packer losses. Convergence is proved under more realistic assumptions than those in the literature (available only for convex problems), in particular, always without the independence assumption between the agent's index updating and the age of the shared optimization variables being used in the update. Numerical results show the the proposed algorithm compares favorably with state-of-the-art methods on convex and nonconvex problems.

■ FB02

North Classroom 1602

Theory and Applications of Global Optimization

General Session

Chair: Emily Speakman, Otto von Guericke University, Magdeburg, 39106, Germany, emily.speakman@ovgu.de

1 - Sparse Pseudoinverses via Relaxations of Moore-penrose

Victor Fuentes, University of Michigan, 1891 IOE Building, 1205 Beal Avenue, Ann Arbor, MI, 48109-2117, United States, vicfuen@umich.edu, Marcia Fampa, Jon Lee

Pseudoinverses are ubiquitous tools for handling over- and under-determined systems of equations. They are used for calculating least-squares solutions for over-determined systems and minimum-norm solutions for under-determined systems. For computational efficiency in the setting of large matrices, sparse pseudoinverses are highly desirable. Recently, sparse left- and right-pseudoinverses were introduced, using 1-norm minimization and linear programming. But a very classical pseudoinverse is the Moore-Penrose pseudoinverse. It is well known to be the unique solution of the four Moore-Penrose properties, and it enjoys other very nice attributes as well. Three of the Moore-Penrose properties are linear (in the input matrix), and one is a quadratic matrix equation. By enforcing or relaxing various combinations of the properties and using 1-norm minimization, we recently introduced several new sparse pseudo inverses. Omitting the quadratic Moore-Penrose property and enforcing various combinations of the linear Moore-Penrose properties, we are led to various linear-programming based sparse pseudo inverses with nice properties. We are developing and experimenting with many strategies for incorporating relaxations of the quadratic Moore-Penrose property; in particular, McCormick inequalities and semi-definite programming relaxations. Because of the heaviness of the semi-definite programming relaxation, we concentrate on several linearization schemes for further relaxing the semi-definiteness constraint. Initial results are promising.

2 - More Virtuous Smoothing

Luze Xu, University of Michigan, 1205 Beal Avenue, Ann Arbor, MI, 48109-2117, United States, xuluze@umich.edu, Jon Lee, Daphne Skipper

In the context of global optimization of mixed-integer nonlinear optimization formulations, we consider smoothing univariate functions f that satisfy f(0) = 0, f is increasing and concave on [0,+inf), f is twice differentiable on all of (0,+inf), but f'(0) is undefined or intolerably large. The canonical examples are root functions f(w) := wp, for 0 < p < 1. We consider the earlier approach of defining a smoothing function g that is identical with f on (d,+inf), for some chosen d > 0, then replacing the part of f on [0,d] with a homogeneous cubic, matching f, f' and f" at d. The parameter d is used to control the derivative at 0 (which controls it on all of [0,+inf) when g is concave). Our main results: (i) we weaken an earlier sufficient condition to give a necessary and sufficient condition for the piecewise function g to be increasing and concave; (ii) we give a general sufficient condition for g'(0) to be nondecreasing in the smoothing parameter d; (iii) we give a general sufficient condition for g to underestimate f; (iv) we give a general sufficient condition for g to dominate the simple `shift smoothing' h(w) := f(w + lambda) - f(lambda) (lambda > 0), when the parameters d and lambda are chosen ``fairly" — i.e., so that g'(0) = h'(0). In doing so, we solve two natural open problems of Lee and Skipper (2016), concerning (ii) and (iii) for root functions.

3 - The Discrete Moment Problem

Christopher Ryan, University of Chicago, 5807 S Woodlawn Ave, Chicago, IL, 60637, United States, chris.ryan@chicagobooth.edu, Xi Chen, Simai He, Bo Jiang, Teng Zhang

The discrete moment problem is a foundational problem in distribution-free robust optimization, where the goal is to find a worst-case distribution that satisfies a given set of moments. This paper studies the discrete moment problems with additional ``shape constraints" that guarantee the worst case distribution is either log-concave or has an increasing failure rate. These classes of shape constraints have not previously been studied in the literature, in part due to their inherent nonconvexities. Nonetheless, these classes of distributions are useful in practice. We characterize the structure of optimal extreme point distributions by developing new results in reverse convex optimization, a lesser-known tool previously employed in designing global optimization algorithms.We are able to show, for example, that an optimal extreme point solution to a moment problem with m moments and log-concave shape constraints is piecewise geometric with at most m pieces.Moreover, this structure allows us to design an exact algorithm for computing optimal solutions in a low-dimensional space of parameters. Moreover, we describe a computational approach to solving these low-dimensional problems, including numerical results for a representative set of instances.

4 - On Branching-point Selection for Triple Products in Spatial Branch-and-bound: Using Mccormick Relaxations

Emily Speakman, Otto von Guericke University, Universitatplatz 2, Magdeburg, 39106, United States, eespeakm@umich.edu, Jon Lee

Muliplying together three or more expressions occurs frequently in factorable global-optimization models. For these triple products, we present some analytic results regarding the choice of branching variable and branching point in the context of sBB for factorable formulations. Following the approach we previously introduced to compare various common relaxations associated with trilinear monomials, we use 4-dimensional volume as a comparison measure to evaluate the sBB choice of branching variable and branching point in this context. Here, we extend our earlier work in this area (which focused only on the Meyer and Floudas hull) to relaxations based on iterating McCormick inequalities.

■ FB03

North Classroom 1202

Dynamic Optimization under Uncertainty

General Session

Chair: Vineet Goyal, Columbia University, 304 SW Mudd Building, 500w 120th Street, New York, NY, 10027, United States, vgoyal@ieor.columbia.edu

1 - A Stochastic Integer Programming Perspective on Multi-Product Newsvendor Model with Substitutions

Weijun Xie, Assistant Professor, Virginia Tech, 205 Durham Hall (MC 0118), 11, Blacksburg, VA, 24061, United States, wxie@vt.edu, Jie Zhang, Subhash Sarin

This paper studies a multi-product newsvendor model with customer-driven demand substitutions, where each product once run out of stock can be proportionally substituted by the others. This model has been widely studied in the literature, however, only very limited analytical properties have been discovered nor efficient solution approaches. This paper first completely characterizes the optimal order policy when the demand is known and reformulates the model as an integer program. When the demand is stochastic, we formulate the model as a two-stage stochastic integer program, derive several necessary optimality conditions, prove its submodularity property, and also develop a polynomial approximation algorithm with the performance guarantee. We further propose a tight upper bound via nonanticipativity dual, which is proven to be quite close to the true optimal and yields a good-quality feasible solution. Numerical studies demonstrate the effectiveness of the algorithms and interesting findings and managerial insights are obtained from a series of sensitivity analyses.

2 - Beyond Worst-case: A Probabilistic Analysis of Affine Policies in Dynamic Robust Optimization

Omar El Housni, Columbia University, New York, NY, 10027, United States, oe2148@columbia.edu, Vineet Goyal

Affine policies are widely used in dynamic optimization where computing an optimal adjustable solution is usually intractable. While the worst case performance of affine policies can be significantly bad, the empirical performance is observed to be near-optimal for a large class of problem instances. In this paper, we aim to address this stark-contrast. In particular, we consider a two-stage dynamic robust optimization problem with linear covering constraints and uncertain right hand side and show that affine policies give a good approximation on random instances generated from a large class of distributions including the commonly used distributions such as uniform and Gaussian.

3 - Optimization under Decision-dependent Uncertainty

Omid Nohadani, Northwestern University, 2145 Sheridan Road, Technological Institute M233, Evanston, IL, 60208-3119, United States, nohadani@northwestern.edu, Kartikey Sharma

The efficacy of robust optimization spans a variety of settings with uncertainties bounded in predetermined sets. In many applications, uncertainties are affected by decisions and cannot be modeled with current frameworks. This paper takes a step towards generalizing robust optimization to problems with decision-dependent uncertainties. In general settings, we show these problems to be NP-complete. To alleviate the computational inefficiencies, we introduce a class of uncertainty sets whose size depends on decisions. We propose reformulations that improve upon alternative standard linearization techniques. To illustrate the advantages of this framework, a shortest path problem is discussed, where the uncertain arc lengths are affected by decisions. Beyond the modeling and performance advantages, the proposed proactive uncertainty control also mitigates over conservatism of current robust optimization approaches.

■ FB04

North Classroom 1204

Robust Optimization: Theory and Applications II

General Session

Chair: Ruiwei Jiang, University of Michigan, Ann Arbor, MI, United States, ruiwei@umich.edu

1 - Distributionally Robust Chance-constrained Optimal Power Flow with Wasserstein-moment Ambiguity Set

Yuanyuan Guo, University of Michigan, 1205 Beal Avenue, Ann Arbor, MI, 48109, United States, yuanyg@umich.edu, Ruiwei Jiang

In this talk, we study a distributionally robust chance-constrained optimal power flow (DRCC-OPF) model, where the distributional information consists of a Wasserstein ball and moments of the renewable energy. We recast the DRCC-OPF model as a conic program and validate its out-of-sample performance (e.g., reliability and cost-effectiveness) via case studies based on real-world wind power data.

2 - Nurse Staffing under Uncertain Demand and Absenteeism

Minseok Ryu, University of Michigan, 1535 Pine Valley Boulevard, Apt 212, Ann Arbor, MI, 48104, United States, msryu@umich.edu, Ruiwei Jiang

This paper describes a data-driven approach for nurse staffing decision under uncertain demand and absenteeism. We propose a distributionally robust nurse staffing (DRNS) model with both exogenous (stemming from demand uncertainty) and endogenous uncertainty (stemming from nurse absenteeism). We provide a separation approach to solve the DRNS model with general nurse pool structures. Also, we identify several classes of nurse pool structures that often arise in practice and show how the DRNS model in each of these structures can be reformulated as a mixed-integer linear program that facilitates off-the-shelf commercial software. Built upon the DRNS model, furthermore, we propose an optimal nurse pool design model, which produces an optimal pool structure that minimizes the number of cross-training while achieving the target staffing costs.

3 - Ambiguous Chance-constrained Bin Packing under Mean-covariance Information

Yiling Zhang, University of Michigan, 2222 Fuller Ct, Ann Arbor, MI, 48105, United States, zyiling@umich.edu Ruiwei Jiang, Siqian Shen

The bin packing structure arises in a wide range of service operational applications, where a set of items are assigned to multiple bins with fixed capacities. With random item weights, a chance-constrained bin packing problem bounds, for each bin, the probability that the total weight of packed items exceeds the bin's capacity. Different from the stochastic programming approaches relying on full distributional information of the random item weights, we assume that only the information of the mean and covariance matrix is available, and consider distributionally robust chance-constrained bin packing (DCBP) models in this paper. Using two types of ambiguity sets, we equivalently reformulate the DCBP models as 0-1 second-order cone (SOC) programs. We further exploit the submodularity of the 0-1 SOC constraints under special and general covariance matrices, and utilize the submodularity as well as lifting and bin-packing structure to derive extended polymatroid inequalities to strengthen the 0-1 SOC formulations. We incorporate the valid inequalities in a branch-and-cut algorithm for efficiently solving the DCBP models. Finally, we demonstrate the computational efficacy of our approaches and performance of DCBP solutions on diverse test instances.

■ FB05

North Classroom 2001

Models and Algorithms for Network Operations

General Session

Chair: Jorge A Sefair, Arizona State Univerity, Tempe, AZ, 85287-8809, United States, jorge.sefair@asu.edu

1 - Conditional Supervalid Inequalities on General Graph Interdiction Problems

Ningji Wei, University at Buffalo, SUNY, Buffalo, NY, 14260, United States, ningjiwe@buffalo.edu, Jose Luis Walteros

We focus on attacker-defender network interdiction problems in which: 1) the objective minimizes the attacker's cost of achieving a fixed disruption level over the defender's problem; 2) the defender's problem is to select an optimal graph structure (e.g., a shortest path, a spanning tree) over the residual graph; 3) the attacker strategies are defined over the same ground set of the graph structures. e.g., if the structures are defined over the edges, the attacker also interdicts edges. We develop a cut-generating framework that produces a general class of supervalid inequalities that remove non-trivial suboptimal solutions. We show how to adapt our framework to tackle a wide variety of interdiction problems.

2 - A Binary Decision Diagram Based Algorithm for Solving a Class of Binary Two-stage Stochastic Programs

Leonardo Lozano, University of Cincinnati, 3031 Eden Avenue, Apt 148, Cincinnati, OH, 45219, United States, leolozano@uc.edu, Cole Smith

We consider a special class of two-stage stochastic integer programming problems with binary variables appearing in both stages. The class of problems we consider constrains the second-stage variables to belong to the intersection of sets corresponding to first-stage binary variables that equal one. Our approach seeks to uncover strong dual formulations to the second-stage problems by transforming them into dynamic programming (DP) problems parameterized by first-stage variables. We demonstrate how these DPs can be formed by use of binary decision diagrams, which then yield traditional Benders inequalities that can be strengthened based on observations regarding the structure of the underlying DPs. We show the efficacy of our approach on a set of stochastic traveling salesman problems.

3 - Sequential Shortest Path Interdiction with Incomplete Information and Limited Feedback

Oleg Prokopyev, University of Pittsburgh, Pittsburgh, PA, United States, prokopyev.engr@pitt.edu

We study a sequential interdiction problem where the interdictor has partial information about the network, while the evader has complete information about it, including the network's structure and arc costs. At each period, she interdicts at most k arcs and then the evader traverses the shortest path from two fixed nodes in the remaining graph. The interdictor seeks to maximize the evader's cumulative costs, however, she does not know the real cost of each arc but only upper and lower bounds. We assume that at each time the interdictor only observes the total cost of the path traversed by the evader, and using this limited information she has to update her knowledge of the arcs' cost so as to make more accurate interdiction decisions. We study in detail greedy and robust interdiction policies and investigate different updating mechanisms, i.e., ways by which the interdictor can exploit the information she observes to update her knowledge of the network's costs. Our results indicate that there is a tradeoff between the computational tractability of the different updating mechanisms and the time each mechanism takes to detect the full information optimal solution. Moreover, we show that even under the strongest possible updating mechanism, convergence to the optimal solution takes a worst-case exponential time periods. These results suggest that the interdictor should have access to less stringent modes of feedback so as to assure polynomial-time convergence to the optimal solution. In this sense, we consider different versions of imperfect and randomized feedback, and prove worst-case polynomial convergence bounds.

4 - An Exact Model and Algorithm for the Multi-vehicle Path Selection and Scheduling Problem with Trajectory Coordination

Jorge A. Sefair, Assistant Professor, Arizona State Univerity, 699 S. Mill Ave., P.O. Box 878809, Tempe, AZ, 85287-8809, United States, jorge.sefair@asu.edu, Navid Matin Moghaddam

We study the problem of finding an optimal set of paths for multiple vehicles between two known nodes in a directed network. Vehicles must coordinate their trajectories in such a way that no two vehicles can be closer to each other than a given distance at any time. For this purpose, we need to find not only the path for each vehicle but also a schedule to traverse the network. Typical examples of this problem include the transportation of hazardous materials and ground operations subject to geographic failures and threats (e.g., natural disasters, air strikes). We discuss the hardness of this problem and present an exact solution method based on a network decomposition. We illustrate the performance of our method over a set of small and medium sized randomly generated networks. We also present an acceleration technique to solve large-scale problems.

■ FB06

North Classroom 2002

Submodular Optimization

General Session

Chair: Aryan Mokhtari, MIT, 77 Massachusetts Avenue, Room 32-D608, Room 32-D608, Cambridge, MA, 02139, United States, aryanm@mit.edu

Co-Chair: Hamed Hassani, University of Pennsylvania, Philadelphia, PA, 19104, United States, hassani@seas.upenn.edu

1 - Fast and Distributed Algorithms for Submodular Optimization

Alina Ene, University of Warwick, Department of Computer Science, Warwick, CV4 7AL, United Kingdom, A.Ene@dcs.warwick.ac.uk

Constrained submodular maximization problems capture many problems of interest both in theory and in practice, including facility location, sensor placement, clustering, influence maximization in social networks, and welfare maximization problems. These problems have received considerable attention over the years, leading to the development of approximation algorithms for a rich class of constraints. Despite this remarkable progress, several fundamental challenges remain, both in terms of understanding the approximability of central problems and improving the running times. In this talk, we describe recent progress on both fronts.

2 - Robust Budget Allocation via Continuous Submodular Functions

Matthew Staib, MIT, Cambridge, MA, United States, mstaib@mit.edu, Stefanie Jegelka

The optimal allocation of resources for maximizing influence, spread of information or coverage, has gained attention in the past years, in particular in machine learning and data mining. But in applications, the parameters of the problem are rarely known exactly, and using wrong parameters can lead to undesirable outcomes. We hence revisit a continuous version of the Budget Allocation or Bipartite Influence Maximization problem introduced by Alon et al. (2012) from a robust optimization perspective, where an adversary may choose the least favorable parameters within a confidence set. The resulting problem is a nonconvex-concave saddle point problem (or game). We show that this nonconvex problem can be solved exactly by leveraging connections to continuous submodular functions, and by solving a constrained submodular minimization problem. Although constrained submodular minimization is hard in general, here, we establish conditions under which such a problem can be solved to arbitrary precision.

3 - Gradient Methods for Submodular Maximization

Amin Karbasi, Yale University, New Haven, CT, United States, amin.karbasi@yale.edu

In this talk, we study the problem of maximizing continuous submodular functions that naturally arise in many learning applications such as those involving utility functions in active learning and sensing, matrix approximations and network inference. Despite the apparent lack of convexity in such functions, we prove that stochastic projected gradient methods can provide strong approximation guarantees for maximizing continuous submodular functions with convex constraints.

4 - Stochastic Submodular Maximization

Hamed Hassani, University of Pennsylvania, 3330 Walnut st., Philadelphia, PA, 19104, United States, hassani@seas.upenn.edu

Stochastic optimization of continuous objectives is at the heart of modern machine learning. However, many important problems are of discrete nature and often involve submodular objectives. We seek to unleash the power of stochastic continuous optimization, namely stochastic gradient descent and its variants, to such discrete problems. We first introduce the problem of stochastic submodular optimization, where one needs to optimize a submodular objective which is given as an expectation. Our model captures situations where the discrete objective arises as an empirical risk (e.g., in the case of exemplar-based clustering), or is given as an explicit stochastic model (e.g., in the case of influence maximization in social networks). By exploiting that common extensions act linearly on the class of submodular functions, we employ projected stochastic gradient ascent and its variants in the continuous domain, and perform rounding to obtain discrete solutions. We focus on the rich and widely used family of weighted coverage functions. We show that our approach yields solutions that are guaranteed to match the optimal approximation guarantees, while reducing the computational cost by several orders of magnitude, as we demonstrate empirically.

■ FB07

North Classroom 3004

Energy

Contributed Session

1 - Successive Convex Approximation for Power Economic Dispatch with Valve-point Effect

Yu Zhang, Assistant Professor, University of California, Santa Cruz, 1156 High St SOE2, Santa Cruz, CA, 95064, United States, zhangy@ucsc.edu

By determining the optimal power output of a set of electricity generators, economic dispatch problem aims to satisfy a given system load at the lowest possible cost, which plays a critical role in the operation and control of power systems. The valve-point effect that results in a rippled generating cost curve makes the problem highly nonconvex and hard to solve in general. By capitalizing on convex approximation techniques, this paper presents a novel and efficient approach, which features provable guarantees of convergence to a local optimal solution. Numerical results on several benchmark systems are reported to corroborate the merits of the proposed method.

2 - Sub-hourly Dispatch Optimization of Photovoltaic and Concentrating Solar Power Hybrid Systems

William Hamilton, Colorado School of Mines, 90 Corona St., Denver, CO, 80218, United States, whamilton@mines.edu

Photovoltaic (PV) systems convert solar light directly into electricity. The cost of PV systems has dropped dramatically in the last few decades, allowing greater penetration into the energy market. However, studies have shown that because PV systems have no cost-effective way to store large amounts of electric energy, their market penetration is limited. PV systems also suffer from variable generation that requires conventional electricity sources to ensure fulfillment of market demands.

Concentrated Solar Power (CSP) technologies capture thermal radiation from the sun utilizing a field of solar tracking heliostats. Coupling CSP systems with inexpensive, dispatchable thermal energy storage (TES) provides CSP technologies with an advantage over traditional solar photovoltaics for utility-scale electricity generation. In a CSP system with TES, the heat transfer fluid used to collect thermal energy can either be sent to the power generation cycle or stored for later use. The coupling, or hybridization, of PV and CSP with TES systems could theoretically generate electricity throughout the year. Ideally, these systems would operate with PV system generation during the day, while the CSP system fills the TES tanks and the power cycle generates at the minimum turndown rate. As generation from PV decreases towards the end of the solar day, the CSP power cycle ramps to full capacity and runs throughout the night. PV-CSP hybrid systems have high capacity factors and stable electricity generation. This coupling also allows for increased PV market penetration.

We have developed a profit-maximizing mixed-integer linear program that determines the operation schedule and allows for evaluation of critical system design decisions. Here, we present the preliminary results from such model with an emphasis on parametric analysis of system design decisions.

■ FB08

North Classroom 3210

Security

Contributed Session

Chair: Ankit Shah, George Mason University, 4400 University Drive, Mail Stop 5B5, Fairfax, VA, 22030-4422, United States, ashah20@gmu.edu

1 - A Study on Generalized Security Games

Wei Wang, University of Pittsburgh, 1025 Benedum Hall, 3700 O'Hara Street, Pittsburgh, PA, 15261, United States, w.wei@pitt.edu, Bo Zeng

We build a generalized defender-attacker-defender (DAD) model, in which the attack decision is made by solving a distinct attacker-defender (AD) model rather than by choosing from a predetermined set as in traditional DAD model. In this model, both upper level defender and lower level attacker can have their own considerations and asymmetric information in their decision making procedures. We demonstrate those generalized DAD games in power grid protection with novel formulations and solution methods. Computational results will be presented, along with insights for practice.

2 - Identifying Critical Nodes in Interdependent Networks with Cascading Failures

Vladimir Boginski, Associate Professor, University of Central Florida, 12800 Pegasus Dr., Orlando, FL, 32541, United States, Vladimir.Boginski@ucf.edu

Interdependent networked systems arise in many application domains, including critical infrastructure networks. One of the well-known examples of coupled interdependent networks are electric power and SCADA networks, which exhibit interdependence in the sense that a failure in one of the network layers may propagate to another layer and then spread through the entire network via a sequence of steps in a cascading fashion. We consider optimization problems of identifying “critical” nodes in coupled interdependent networks, that is, a subset of nodes whose deletion causes maximum network fragmentation (as quantified by an appropriate metric) in the presence of cascading failures. Theoretical and computational aspects of these problems will be discussed.

3 - The Power of ‘N’ Choices in Redundancy with Balking

Maryam Sadeghi, Student, University of British Columbia-Okanagan, 207, 1960 Pacific Court, Kelowna, BC, V1Y8B3, Canada, m.sadeghi@alumni.ubc.ca, Javad Tavakoli

In cloud-storage systems with a large number of servers, reducing latency has become increasingly important since almost every one spends time downloading files or pictures from websites such as google and facebook. Various strategies have been introduced to reduce latency. In particular, redundancy can play a critical role in reducing the waiting time in queues. In a system with L servers, each incoming job can be sent to a constant number of servers, N, chosen at random. Arriving jobs are modeled as a Poisson process with the arrival rate depending on the number of jobs in the queue, and service times are exponentially distributed with the mean 1. Under the assumption that job submissions are discouraged when the system is used frequently, arriving jobs join the system with probability depending on the number of existing jobs in the system. In this research, we propose a theoretical model for discouraged arrival jobs to evaluate the steady-state queue length distribution.

4 - Dynamic Optimization of the Level of Operational Effectiveness of a Cybersecurity Operations Center

Ankit Shah, George Mason University, 4400 University Drive, Mail Stop 5B5, Fairfax, VA, 22030-4422, United States, ashah20@gmu.edu, Rajesh Ganesan

The analysts at a cybersecurity operations center (CSOC) analyze the alerts that are generated by intrusion detection systems (IDSs). Under normal operating conditions, sufficient numbers of analysts are available to analyze the alert workload in a reasonable amount of time. However, there are many disruptive factors (organizational and adversarial) that can adversely impact the normal operating conditions such as higher alert generation rates from a few IDSs, new alert patterns that decrease the throughput of the alert analysis process, analyst absenteeism, and internal system failures. The impact of all the above factors is that the alerts wait for a long duration before being analyzed, which impacts the readiness of the CSOC. It is imperative that (1) the readiness of the CSOC be quantified, which in this talk is defined as the level of operational effectiveness (LOE) of a CSOC, and (2) optimal actions are taken to bring the LOE to normal operating conditions. In order to return the CSOC to normal operating conditions, several actions can be taken such as increasing alert analysis time spent in a shift by delaying/canceling non-alert analysis related activities of the analysts, utilizing manager’s time to assist with the alert investigation, and summoning the on-call analyst workforce to boost the service rate of alerts. However, additional resources are limited in quantity over a 14-day work cycle, and the decision to determine when and how much action to take in the face of uncertainty is non-trivial. This talk presents a Reinforcement Learning (RL) model for optimizing the LOE throughout the entire 14-day work cycle. Results indicate that the RL model provides better decisions than current practices in determining when and how much resource to allocate when the LOE of a CSOC deviates from the normal operating conditions.

■ FB09

Student Commons 1500

Learning-Based Robotics and Control

General Session

Chair: Anil Aswani, UC Berkeley, San Francisco, CA, 94103, United States, aaswani@berkeley.edu

1 - Safe Trajectory Synthesis for Autonomous Driving in Unforeseen Environments

Ram Vasudevan, University of Michigan, 564 Kellogg Street, Ann Arbor, MI, 48105, United States, ramv@umich.edu

Path planning for autonomous vehicles in arbitrary environments requires a guarantee of safety, but this can be impractical to ensure in real-time when the vehicle is described with a high-fidelity model. To address this problem, this talk develops a method to perform trajectory design by considering a low-fidelity model that accounts for model mismatch. The presented method begins by computing a conservative Forward Reachable Set (FRS) of a high-fidelity model’s trajectories produced when tracking trajectories of a low-fidelity model over a finite time horizon. At runtime, the vehicle intersects this FRS with obstacles in the environment to eliminate trajectories that can lead to a collision, then selects an optimal plan from the remaining safe set. By bounding the time for this set intersection and subsequent path selection, this talk proves a lower bound for the FRS time horizon and sensing horizon to guarantee safety. This method is demonstrated on several real-world systems.

2 - Optimal Switching for Discrete-time Switched Systems in Finite Time

Qie He, University of Minnesota, 111 Church Street SE, Minneapolis, MN, 55455, United States, qhe@umn.edu

Given a discrete-time switched systems consisting of two n-th order linear time-invariant subsystems, we are interested in finding an optimal switching sequence to minimize a linear function of the system state after K periods. We show that this problem is NP-hard even when the system matrices are a pair of Boolean or stochastic matrices, if both n and K are input parameters. We develop an exact algorithm to solve this problem by sequentially constructing the convex hull of all possible states at each period. We show that the proposed algorithm runs in polynomial time if the system matrices satisfy certain conditions. In particular, the algorithm runs in polynomial time for a pair of 2-by-2 Boolean matrices.

3 - Combining Inverse Optimization with Reinforcement Learning

Anil Aswani, UC Berkeley, San Francisco, CA, 94103, United States, aaswani@berkeley.edu, Mo Zhou, Yonatan Mintz, Elena Flowers, Ken Goldberg, Philip Kaminsky, Yoshimi Fukuoka

Though reinforcement learning has found success in many applications, these approaches are less commonly used to engineer human-automation systems because utility maximization models of human decision-making are not compatible with existing reinforcement learning frameworks. This talk describes a new framework for reinforcement learning that integrates inverse optimization (for the purpose of estimating utility maximization models of human decision-making) into the reinforcement learning. We discuss how to formulate and solve the inverse optimization problem to ensure statistical consistency, and then discuss how to calculate a policy for a closed-loop system where the human is making decisions based on this utility maximization model. The key idea is to appropriately formulate an optimization model (whose solution describes a policy) that can be solved using mixed-integer numerical optimization software, and we use stochastic variational analysis to show that this formulation is such that as more data is collected the policy converges to the optimal one. We discuss how the effectiveness of our approach was demonstrated by a clinical trial we conducted on personalized goal-setting through smartphone apps to increase physical activity.

■ FB10

Student Commons 160

Stochastic Optimization for Reinforcement Learning

General Session

Chair: Alec Koppel, U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, MD, 20783, United States, aekoppel314@gmail.com

1 - Stochastic Primal-Dual Methods and Sample Complexity of Markov Decision Processes

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

We study the online estimation of the optimal policy of a Markov decision process (MDP). We propose a class of Stochastic Primal-Dual (SPD) methods which exploit the inherent minimax duality of Bellman equations. The SPD methods update a few coordinates of the value and policy estimates as a new state transition is observed. For MDPs that are ``sufficiently” ergodic, we show that SPD methods achieve a near-optimal sample/running-time complexity. We also provides the first computational complexity lower bound for solving MDPs.

2 - Stochastic Policy Gradient Ascent in Reproducing Kernel Hilbert Spaces

Santiago Paternain, 315 S. 45th Street, Philadelphia, PA, 19104, United States, spater@seas.upenn.edu

In this work, we consider the problem of policy optimization in the context of reinforcement learning. In order to avoid discretization, we select the optimal policy to be a continuous function belonging to a Reproducing Kernel Hilbert Space (RKHS) which maximizes an expected discounted reward (EDR). We design a policy gradient algorithm (PGA) in this context, deriving the gradients of the functional EDR and learning the unknown state transition probabilities on the way. In particular, we avoid the standard Q-learning step in PGA, since it requires running the system infinite times per iteration, rendering the overall algorithm unrealizable. Alternatively, we propose an unbiased stochastic approximation for the gradient that can be obtained in a finite number of steps. This unbiased estimator is the key enabler for a novel stochastic PGA, which provably converges to a critical point of the EDR. However, the RKHS approach increases the model order per iteration by adding extra kernels, which may render the numerical complexity prohibitive. To overcome this limitation, we prune the kernel dictionary using an orthogonal matching pursuit procedure and prove that the modified method keeps the model order bounded for all iterations while ensuring convergence to a neighborhood of the critical point.

3 - Nonparametric Stochastic Quasi-gradient Method for Policy Evaluation in Continuous Markov Decision Problems

Alec Koppel, U.S. Army Research Laboratory, Adelphi, MD, 20783, United States, akoppel@seas.upenn.edu, Ethan Stump, Garrett Warnell, Alejandro Ribeiro

We consider policy evaluation in infinite-horizon discounted Markov decision problems (MDPs) with continuous compact state and action spaces. We reformulate this task as a compositional stochastic program with a function-valued decision variable that belongs to a reproducing kernel Hilbert space (RKHS). We approach this problem via a new functional generalization of stochastic quasi-gradient methods operating in tandem with stochastic sparse subspace projections. The result is an extension of gradient temporal difference learning that yields nonlinearly parameterized value function estimates of the solution to the Bellman evaluation equation. We call this method Parsimonious Kernel Gradient Temporal Difference (PKGTD) Learning. Our main contribution is a memory-efficient non-parametric stochastic method guaranteed to converge exactly to the Bellman fixed point with probability 1 with attenuating step-sizes under the hypothesis that it belongs to the RKHS. Further, with constant step-sizes and compression budget, we establish mean convergence to a neighborhood and that the value function estimates have finite complexity. In the Mountain Car domain, we observe faster convergence to lower Bellman error solutions than existing approaches with a fraction of the required memory.

■ FB11

Student Commons 2500

New Techniques in Discrete and Mixed-discrete Optimization

General Session

Chair: Yuri Faenza, Columbia University, New York, NY, 10027, United States, yf2414@columbia.edu

1 - Mixed-integer Convex Representability

Juan Pablo Vielma, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E62-561, Cambridge, MA, 02139, United States, jvielma@mit.edu, Miles Lubin, Ilias Zadik

We consider the question of which nonconvex sets can be represented exactly as the feasible sets of mixed-integer convex optimization problems (MICP). We first show a complete characterization for the case when the number of possible integer assignments is finite. We then further study the characterization for the more general case of unbounded integer variables and introduce a simple necessary condition for representability. This condition can be used to show that the set of prime numbers is not MICP representable, even though it can be represented using polynomial equations and integrality constraints. While the result for the prime numbers suggests certain regularity of MICP representable sets, we show that even for subsets of the natural numbers, MICP representable sets can be significantly more irregular than rational mixed integer linear programming representable sets. Inspired by these irregular MICP representable sets we introduce a notion of rational MICP representability and show how this notion imposes regularity to MICP representable subsets of the natural numbers, for compact convex sets and the graphs and epigraphs of certain functions. Finally, we study other notions of regularity associated to infinite unions of convex sets with the same volume.

2 - Treewidth-based Extension Complexity Lower Bounds

Gonzalo Munoz, Polytechnique Montréal, 2920 Chemin de la Tour, Montréal, QC, Canada, gonzalo.munoz@polymtl.ca, Yuri Faenza

In this work, we study the extension complexity of 0-1 sets parametrized by treewidth: a graph-theoretical parameter that measures structured sparsity. If a 0-1 set can be formulated as the set of binary vectors that satisfy some polynomial inequalities, and those polynomial inequalities present a sparsity pattern whose treewidth is k, then it is known that the extension complexity of the convex hull of the set is O(n2^k). The goal of this work is to prove the existence of 0-1 sets that (nearly) meet this bound, for any arbitrary treewidth level k. To the best of our knowledge, this is the first work to provide parametric lower bounds on extension complexity.

3 - Learning Projections for Learning Faster

Swati Gupta, Simons Institute for the Theory of Computing, Berkeley, CA, 94708, United States, swatig@mit.edu

First-order projection based methods need to compute the minimizer of a certain Bregman divergence, potentially in each step of the method, whenever the decision set is constrained. We will talk about speeding up this computation by using various online learning and combinatorial optimization techniques.

■ FB12

Student Commons 2504

Integer Programming and Polyhedral Theory

Contributed Session

Chair: Charles Viss, University of Colorado Denver, 2300 E Wesley Ave, Denver, CO, 80210, United States, charles.viss@ucdenver.edu

1 - Optimal Solution Characteristics and Cutting Planes for the Asset Replacement Problem under Technological Advances

Esra Buyuktahtakin, Associate Professor, New Jersey Institute of Technology, Mechanical and Industrial Engineering Dept, Mechanical Engineering Center, 200 Central Av, Newark, NJ, 07102, United States, esratoy@njit.edu, Joseph C Hartman

The parallel replacement problem under economies of scale (PRES) determines minimum-cost keep-and-replace decisions for a group of assets. We study the MIP formulation of PRES under technological change (PRES-T), such that newer, technologically advanced assets have higher capacity than assets purchased earlier. We provide optimal solution characteristics and insights into the economics of the problem and derive cutting planes for solving it more efficiently. Computational experiments on the USPS fleet management case illustrate that the inequalities are quite effective in solving PRES-T instances.

2 - A New Approach for Extending Cover Inequalities for the Robust Knapsack Polytope

Timo Gersing, RWTH Aachen University, Pontdriesch 10-12, Aachen, 52062, Germany, gersing@math2.rwth-aachen.de, Christina Büsing, Arie Koster

The binary knapsack problem is one of the most frequently occurring substructures in integer linear programs. Thus, we are particularly interested in studying its polytope and finding valid inequalities in order to solve complex problems containing knapsack constraints more efficiently within a branch-and-cut framework. Such frameworks are also widely used in robust optimization. Therefore, polytopal insights into robust counterparts of the knapsack problem become more and more relevant in practice. In this talk, we study the robust knapsack with uncertain weights for which we consider the budget uncertainty set of Bertsimas and Sim. In literature, robust cover inequalities have been proposed alongside with two different extensions. We propose a novel class of extended cover inequalities, exploiting the structure of the uncertainty set by combining information of different scenarios. The new extensions are particularly suited for instances where the lengths of the uncertainty intervals around the nominal weights are not proportional to the nominal weights themselves. A preliminary computational study confirms the added value of the new approach, compared to the earlier versions of extended cover inequalities.

3 - Circuit Walks in Integral Polyhedra

Charles Viss, University of Colorado-Denver, Denver, CO, United States, charles.viss@ucdenver.edu, Steffen Borgwardt

As a generalization of the edge directions of a polyhedron, circuits play a fundamental role in the theory of optimization due to their connection to combinatorial algorithms and the efficiency of the Simplex method. To gain a better understanding of circuits, we introduce a hierarchy for integral polyhedra based on behaviors exhibited by their circuit walks. Many problems in combinatorial optimization fall into the most interesting categories of this hierarchy—-steps of circuit walks in the associated polyhedra only stop at integer points, at vertices, or follow actual edges. We classify several classical families of integral polyhedra within the hierarchy, including 0/1-polytopes, polyhedra defined by totally unimodular matrices, and more specifically matroid polytyopes, transportation polytopes, and partition polytopes. Finally, we prove several equivalent characterizations of the non-degenerate polytopes in the bottom level of the hierarchy where all circuit walks are edge walks.

■ FB13

Student Commons 2600

Stochastic Algorithms for Large Scale Optimization

General Session

Chair: Peter Richtarik, University of Edinburgh, Edinburgh, EH93JZ, United Kingdom, peter.richtarik@ed.ac.uk

Co-Chair: Filip Hanzely, King Abdullah University of Science and Technology, Jeddah, 23955, Saudi Arabia, filip.hanzely@kaust.edu.sa

1 - Relative-continuity for Non-lipschitz Non-smooth Convex Optimization using Stochastic (or Deterministic) Mirror Descent

Haihao Lu, MIT, 60 Wadsworth St, Apt 14E, Cambridge, MA, 02142, United States, haihao@mit.edu

The usual approach to developing and analyzing first-order methods for non-smooth (stochastic or deterministic) convex optimization assumes that the objective function is uniformly Lipschitz continuous with parameter M_f. However, in many settings the non-differentiable convex function $f(\cdot)$ is not uniformly Lipschitz continuous — for example (i) the classical support vector machine (SVM) problem, (ii) the problem of minimizing the maximum of convex quadratic functions, and even (iii) the univariate setting with $f(x) := \max\{0, x\} + x^2$. Herein we develop a notion of ``relative continuity" that is determined relative to a user-specified ``reference function" $h(\cdot)$ (that should be computationally tractable for algorithms), and we show that many non-differentiable convex functions are relatively continuous with respect to a correspondingly fairly-simple reference function $h(\cdot)$. We also similarly develop a notion of ``relative stochastic continuity" for the stochastic setting. We analysis two standard algorithms — the (deterministic) mirror descent algorithm and the stochastic mirror descent algorithm — for solving optimization problems in these two new settings, and we develop for the first time computational guarantees for instances where the objective function is not uniformly Lipschitz continuous. This paper is a companion paper for non-differentiable convex optimization to the recent paper by Lu, Freund, and Nesterov, which developed similar sorts of results for differentiable convex optimization.

2 - A Structural Approach to Hardness of Continuous Optimization Problems

Ohad Shamir, Weizmann Institute of Science, Rehovot, Israel, ohad.shamir@weizmann.ac.il, Yossi Arjevani

The field of continuous optimization stands out as a remarkable example where a systematic study of computational boundaries has repeatedly furnished an important source for many cornerstone algorithms. Arguably, the most important computational model used in generic unstructured continuous optimization is that of oracle complexity, where an optimization process is modeled as a sequence of oracle calls issued so as to reveal a given problem instance to a sufficient extent. Although this model has been used to establish significant complexity bounds for important problem classes, in many cases, not taking into account the computational resources required for issuing and processing oracle queries produces a limited and somewhat deceptive picture of the actual computational efficiency. In this work, we devise a novel framework which models the dynamics of optimization methods much more intimately by controlling the functional form allowed for producing new iterates. The added structure assumed by our framework is used to establish novel complexity bounds under new regimes and problem settings. In particular, we show a dimension-independent bounds for a wide class of first-order and coordinate-descent algorithms for various settings, including finite-sum minimization. Lastly, this formulation suggests a more refined taxonomy of optimization methods which imply several useful design principles.

3 - Randomized and Accelerated for Minimizing Relatively Smooth Functions

Filip Hanzely, PhD Student, King Abdullah University of Science and Technology, Thuwal, Jeddah, 23955, Saudi Arabia, filip.hanzely@kaust.edu.sa

Relative smoothness - a notion introduced by Birnbaum et al. (2011) and rediscovered by Bauschke et al. (2016) and Lu et al. (2016) - generalizes the standard notion of smoothness, i.e., Lipschitz continuity of the gradients, typically used in the analysis of gradient type methods. In this work we are taking ideas from the field of convex optimization and applying them to the relatively smooth setting. In particular, we develop stochastic and accelerated algorithms for minimizing relatively smooth functions which outperform the standard methods developed in the papers mentioned above.

Friday, 2:30PM - 4:00PM

■ FC01

North Classroom 1806

Stochastic and Distributed Algorithms for Nonlinear Optimization I

General Session

Chair: Necdet Serhat Aybat, Penn State University, University Park, PA, 16802, United States, nsa10@psu.edu

Co-Chair: Shiqian Ma, UC Davis, Davis, CA, United States, sqma@ucdavis.edu

1 - Optimal Algorithms for Distributed Optimization

Angelia Nedich, ASU, 650 E. Tayler Mall, Goldwater Center, Rm 311, Tempe, AZ, 85281, United States, Angelia.Nedich@asu.edu, C.a. Uribe, S. Lee, A. Gashikov

We will discuss the optimal convergence rates for distributed convex optimization problems in networks, where the objective is to minimize the sum of network agents' objective functions. We model the communication restrictions imposed by the network as a set of affine constraints and provide optimal complexity bounds for several different cases based on the properties of the agents' objective functions. Our approach is making use of the dual of an appropriately formulated primal problem which includes the underlying static graph that models the communication restrictions. Our results show that Nesterov's accelerated gradient method for the dual problem can be executed in a distributed manner and that it enjoys the same optimal rates as in the centralized version of the problem (up to a constant or logarithmic factors), with an additional cost related to the spectral gap of the interaction matrix.

2 - When Cyclic Coordinate Descent Outperforms Randomized Coordinate Descent

Nuri Denizcan Vanli, Massachusetts Institute of Tehnology, 22 Palermo Street, Cambridge, MA, 02141, United States, denizcan@mit.edu, Asuman Ozdaglar

The coordinate descent (CD) method is a classical optimization algorithm that has seen a revival of interest because of its competitive performance in machine learning applications. A number of recent papers provided convergence rate estimates for their deterministic (cyclic) and randomized variants that differ in the selection of update coordinates. These estimates suggest randomized coordinate descent (RCD) performs better than cyclic coordinate descent (CCD), although numerical experiments do not provide clear justification for this comparison. In this paper, we provide examples and more generally problem classes for which CCD (or CD with any deterministic order) is faster than RCD in terms of asymptotic worst-case convergence. Furthermore, we provide lower and upper bounds on the amount of improvement on the rate of CCD relative to RCD, which depends on the deterministic order used. We also provide a characterization of the best deterministic order (that leads to the maximum improvement in convergence rate) in terms of the combinatorial properties of the Hessian matrix of the objective function.

3 - Robust Distributed Optimization for Networks with Unreliable Nodes

Jinming Xu, Arizona State University, Tempe, AZ, United States, jinmingx@asu.edu, Anna Scaglione, Angelia Nedich

In the past ten years, we have seen significant growth in the development of distributed algorithms to optimize a network utility (usually the sum of local functions) over a peer-to-peer network. These algorithms make use of local network resources, such as computing power and message-passing, to allow agents to determine an optimal common solution. Examples include distributed sensing, resource allocation and machine learning. The flat structure of distributed algorithms makes it amenable to asynchronous computation and communication but also render it highly vulnerable when there are some nodes which are unreliable due to faults or attacks (even if only one node is compromised). In this talk, we will share with you some recent work on robust design of distributed algorithms in the presence of insider attacks which keep injecting false data to the network, e.g., attacks from stubborn nodes. In particular, we show that it is beneficial to not have all agents reaching consensus but intentionally make some of them to be arbitrary. This can be done by leveraging proper regularization techniques, such as total variation that automatically assigns more weight to those with small residual but less to those with large disagreement. In so doing, we are able to loosen the requirement of consensus and allow regular nodes to knowingly disagree with irregular or irrational nodes, which may in turn lead to identification and isolation of unreliable nodes.

4 - Multi-agent Constrained Optimization of a Strongly Convex Function Over Time-varying Directed Networks

Necdet Serhat Aybat, Penn State University, University Park, PA, 16802, United States, nsa10@psu.edu

We consider cooperative multi-agent consensus optimization problems over possibly directed, time-varying communication networks, where only local communications are allowed. The objective is to minimize the sum of agent-specific possibly non-smooth composite convex functions over agent-specific private conic constraint sets; hence, the optimal consensus decision should lie in the intersection of these private sets. Assuming the sum function is strongly convex, we provide convergence rates in sub-optimality, infeasibility and consensus violation; examine the effect of underlying network topology on the convergence rates of the proposed decentralized algorithm.

■ FC02

North Classroom 1602

Joint Session: Network/OPT Uncertainty, Optimizing Stochastic Emerging Transportation Problems

General Session

Chair: Siqian Shen, University of Michigan, Ann Arbor, MI, 48109, United States, siqian@umich.edu

1 - Wasserstein Distance and the Distributionally Robust TSP

John Gunnar Carlsson, University of Southern California, 3750 McClintock Avenue, Los Angeles, 90089, United States, jcarlsso@usc.edu

Motivated by a districting problem in multi-vehicle routing, we consider a distributionally robust version of the Euclidean travelling salesman problem in which we compute the worst-case spatial distribution of demand against all distributions whose Wasserstein distance to an observed demand distribution is bounded from above. This constraint allows us to circumvent common overestimation that arises when other procedures are used, such as fixing the center of mass and the covariance matrix of the distribution. Numerical experiments confirm that our new approach is useful when used in a decision support tool for dividing a territory into service districts for a fleet of vehicles when limited data is available.

2 - Approximate Dynamic Programming for Planning Driverless Fleets of Electric Vehicles

Lina Al-Kanj, Princeton University, Princeton, NJ, United States, lina.kanj@gmail.com, Warren B. Powell

By year 2021, almost every major auto company, along with fleet operators such as Uber and Lyft, have announced plans to put driverless vehicles on the road. At the same time, electric vehicles are quickly emerging as a next-generation technology that is cost effective, in addition to offering the benefits of reducing the carbon footprint. The combination of a centrally managed fleet of driverless vehicles, along with the operating characteristics of electric vehicles, is creating a transformative new technology that offers significant cost savings with high service levels. This problem involves a control problem for assigning requesters to cars, a planning problem for deciding on the fleet size and a pricing problem all of which are high dimensional stochastic dynamic programs. In this work, we propose to use approximate dynamic programming to develop high-quality operational control strategies to determine which car (given the battery level) is best for a particular trip (considering its length and destination), when a car should be recharged, and when it should be re-positioned to a different zone which offers a higher density of trips. We then propose to use outputs (in the form of value functions) from the operational planning model to optimize the distribution of battery capacities in the fleet. We wish to determine the number of cars required to provide a high level of service, and from this to understand the economics of a driverless fleet of electric vehicles.

3 - On the Values of Vehicle-to-grid Selling in Electric Vehicle Sharing

Yiling Zhang, University of Michigan, Ann Arbor, MI, United States, zyiling@umich.edu, Mengshi Lu, Siqian Shen

In recent years, the demand and support of fuel-efficient and sustainable transportation has undergone rapid growth. In this paper, we study an infrastructure planning and fleet management problem in electrical vehicle sharing systems with the integration of vehicle-to-grid (V2G) selling of electricity. Through extensive computational tests by solving a two-stage stochastic integer program, we analyze the benefits of such an integration from three different dimensions, including i) user benefit through improved service satisfaction, ii) service provider benefit through increased revenues, and iii) socio-environmental benefit through the reduction of private vehicle ownership and greenhouse gas emissions.

■ FC03

North Classroom 1202

Stochastic Programming: Sampling and Learning

General Session

Chair: Suvrajeet Sen, University of Southern California, Los Angeles, CA, 90089-0193, United States, s.sen@usc.edu

1 - Monte Carlo Tree Search with Sampled Information Relaxation Dual Bounds

Daniel Jiang, University of Pittsburgh, 1002 Benedum Hall, 3700 O'Hara Street, Pittsburgh, PA, 15261, United States, drjiang@pitt.edu, Lina Al-Kanj, Warren B Powell

MCTS is a well-known strategy for solving sequential decision problems, particularly in the area of game-play AI. We propose a new technique called Primal-Dual MCTS that utilizes sampled information relaxation (Brown et. al., 2010) bounds on potential actions in order to make tree expansion decisions. The approach shows promise when used to optimize the behavior of a driver navigating a graph while operating on a ride-sharing platform.

2 - Optimizing the Design of a Latin Hypercube Sampling Estimator

Alexander Zolan, University of Texas-Austin, Austin, TX, 78712, United States, alex.zolan@utexas.edu, John Hasenbein, David Morton

Latin hypercube sampling (LHS) reduces variance, relative to naive Monte Carlo sampling, by partitioning the support of a random vector into strata. When creating these estimators, we must determine: (i) the number of strata; and, (ii) the partition that defines the strata. In this paper, we address the second point by formulating a nonlinear optimization model that designs the strata to yield a minimum-variance LHS estimator for a univariate function. Under a discrete set of candidate boundary points, the optimization model can be solved via dynamic programming. We extend this method by using an approximation of estimator variance to obtain strata for the domain of a multivariate function. We solve this problem exactly via a dynamic program for each random component. We present sufficient conditions for this technique to guarantee variance reduction compared to using LHS with equal-probability strata. Finally, we show empirical evidence of variance and bias reduction in solutions to stochastic programs in the literature using our method, compared to commonly used variance reduction techniques.

3 - Compromise Decisions and Statistical Optimality

Suvrajeet Sen, University of Southern California, Daniel J. Epstein Dept. of, Industrial and Systems Engineering, Los Angeles, CA, 90089-0193, United States, s.sen@usc.edu, Yunxiao Deng

A compromise decision in stochastic programming represents a decision which is an optimum with respect to a proximal mapping of a grand mean function which is the sample mean of value functions obtained by replications. Such decisions reduce variance in stochastic programming solutions, and can be used to estimate the probability that the solution is with a given tolerance from an optimal solution. In this talk we discuss these concepts within the context of the stochastic decomposition algorithm, and report computations which demonstrate the practical use of the concept of statistical optimality. This type of optimality is ideally suited for infinite dimensional SP models where the random variables are allowed to be continuous. Applications in Learning Enabled Optimization will be presented.

■ FC04

North Classroom 1204

Solving ACOPF Problem Under Uncertainty

General Session

Chair: Haoxiang Yang, Northwestern University, Evanston, IL, 60208, United States, haoxiangyang2019@u.northwestern.edu

1 - Tight Piecewise Convex Relaxations for Global Optimization of Optimal Power Flow

Harsha Nagarajan, Scientist, Los Alamos National Laboratory, 3000 Trinity Drive, Los Alamos, NM, 87544, United States, harsha@lanl.gov, Mowen Lu, Russell Bent, Sandra D. Eksioglu, Scott J. Mason

Since the alternating current optimal power flow (ACOPF) problem was introduced in 1962, developing efficient solution algorithms for the problem has been an active field of research. In recent years, there has been increasing interest in convex relaxations-based solution approaches that are often tight in practice. We further improve the quality of these convex relaxations by employing extreme-point-based convex-hull representations for multilinear functions. Based on these approaches, we develop tight piecewise convex relaxations for convex-hull characterizations and provide a few useful results on the polyhedral properties of these relaxations. Further, using these tight

relaxations, we develop an adaptive, multivariate partitioning algorithm with bound tightening that progressively improves these relaxations and, given sufficient time, converges to the globally optimal solution. We illustrate the strengths of our algorithm using benchmark ACOPF test cases from the literature. Computational results show that our novel algorithm reduces the best-known optimality gap for some hard ACOPF cases.

2 - Global Solution of AC Optimal Power Flow Problems

Michael Bynum, Ph.D. Candidate, Purdue University, West Lafayette, IN, United States, bynumm@purdue.edu, Carl Laird, Anya Castillo, Jianfeng Liu, Jean-Paul Watson

Successful operation of the electricity grid relies on solution of optimal power flow problems. Typical approaches use a linear approximation of the nonlinear AC power flow, however, solutions from these approximations are not guaranteed to be optimal or even feasible for the original nonlinear problem. Given the need for efficient global solution of the nonlinear AC power flow problem, in this presentation, we will describe the suite of available relaxations and show results for a multi-tree approach based on bounds-tightening and piecewise linearization of the nonlinear constraints.

3 - Robust Optimization and Control for Infrastructure Network Applications

Haoxiang Yang, Northwestern University, 2145 Sheridan Road, Evanston, IL, 60208, United States, haoxiangyang2019@u.northwestern.edu, David Morton, Krishnamurthy Dvijotham, Chaithanya Bandi

In this talk we will present the theoretical results of solving a robust optimization problem for the power system under uncertainty. Solving the deterministic alternating current optimal power flow (ACOPF) problem has been considered a hard problem since 1960s because the optimization problem is nonlinear and highly nonconvex. Linear approximation of the AC power flow system (DC approximation) has been deployed in the industry but does not guarantee a physically feasible system configuration. In recently years, different convex relaxation schemes of the ACOPF problem have been researched, and under some assumptions, a physically feasible solution can be recovered. Based on these convex relaxation schemes, we construct a robust convex optimization problem to solve for optimal controllable injections (fossil fuel, nuclear etc.) in electricity power systems under uncertainty (renewable energy generation, demand fluctuation, etc.). Experiment results indicate that the robust convex relaxation of the ACOPF problem will provide a tight lower bound, and for the test cases where the nominal relaxation is tight, a solution to the non-convex robust ACOPF problem can be found.

■ FC05

North Classroom 2001

Applications in Network Optimization

General Session

Chair: Robert Mark Curry, Clemson University, Clemson, SC, 29634, United States, rmcurry@clemson.edu

1 - Branch-cut-and-price for Vehicle Routing Problems with Time Windows and Convex Node Costs

Yongjia Song, Virginia Commonwealth University, 1015 Floyd Avenue, P.O. Box 843083, Richmond, VA, 23284, United States, ysong3@vcu.edu, Qie He, Stefan Irnich

We study an extension of the vehicle routing problem with time windows (VRPTW) in which the objective is to minimize the sum of routing and customer inconvenience costs. The latter inconvenience costs are defined by general convex functions, one for each customer, that express the customer's preference for a specific service start time. We call this problem the VRP with time windows and convex node costs (VRPTW-CNC). The VRPTW-CNC is a three-level optimization problem with interdependent levels for clustering, routing, and schedule optimization. We propose the first effective branch-cut-and-price algorithm for the VRPTW-CNC that is able to solve to optimality instances with 100 customers. The novelty of our algorithm lies in the column-generation mechanism. We show that dynamic-programming labeling can be used to simultaneously solve the two lower levels of routing and schedule optimization, while the clustering in the first level is standard.

2 - The Weighted Target Set Selection Problem on Cycles

Rui Zhang, Unversity of Colorado Boulder, Boulder, CO, 80309, United States, rui.zhang@colorado.edu

The study of viral marketing strategies on social networks has become an area of significant research interest. In this setting we consider a combinatorial optimization problem referred to as the weighted target set selection (WTSS) problem. In the WTSS problem, we are given a connected undirected graph G = (V, E), where for each node i in V, there are a threshold gi which is between 1 and deg(i) (the degree of node i) and a positive weight, denoted by bi, which models the situation that different nodes might require differing levels of effort to become initial adopters. All nodes are inactive initially. We select a subset of nodes, the target set, and they become active. After that, in each step, we update the state of nodes by the following rule: an inactive node i becomes active if at least gi of its neighbors are active in the previous step. The goal is to find the minimum cardinality target set while ensuring that all nodes are active by the end of this activation process. Earlier research focused on approximation algorithms for the unweighted case of the problem which is known to be NP-hard. Motivated by the desire to develop a better understanding of fundamental problems in social network analytics, we focus on a special case where the underlying graphs are cycles. Our paper makes several contributions. First, we propose a linear time algorithm for the WTSS problem on cycles. More importantly, we present a tight formulation in the node space. The polytope of the WTSS problem on cycles is defined as the convex hull of 0/1-incidence vectors of all feasible solution of the WTSS problem on cycles. The proposed formulation provides a complete description of this polytope. Our work can be a building block for developing exact methods for tackling this important problem in social network analytics.

3 - Designing River Basin Storage using Optimization

Andy Burrow, Colorado School of Mines, Golden, CO, United States, aburrow@mines.edu, Alexandra M. Newman

The ways in which a growing population increases hydrologic demand are often evaluated using simulation models. This research uses data produced by the State of Colorado's Stream Simulation Model as input to an optimization model to determine the flow of unappropriated water so as to minimize the cost of water shortage mitigation while adhering to constraints that force the physical and topographical structures of the river. Solutions incorporate new storage infrastructure and provide prescriptive, rather than descriptive, flow information.

4 - An Augmenting Flow Algorithm for the Node-capacitated Maximum Flow Problem

Robert Mark Curry, Clemson University, 278 Freeman Hall, 211 Fernow St., Clemson, SC, 29634, United States, rmcurry@g.clemson.edu, Cole Smith

We consider maximum flow problems having node- and arc-capacity constraints, in which each unit of flow on arc (i,j) consumes a non-negative amount of capacity at node i. Rather than solving a linear program for these problems, we present an augmenting flow algorithm. In this algorithm, we augment flows along a path having positive residual capacities among all arcs and nodes along the path. When no such paths exist, we find and augment flow along flow-cycles that increase the residual capacity at an incapacitated node. We prove the optimality and computational complexity of our algorithm and apply it to solve various problems in wireless sensor network settings.

■ FC06

North Classroom 2002

Software and Implementation

Contributed Session

Chair: Christopher Lourenco, Texas A&M University, 999 West Villa Maria Road, Bryan, TX, 77801, United States, clouren@tamu.edu

1 - SAS® Viya™ Optimization Modeling Interface for Python

Sertalp B. Cay, SAS Institute, 500 SAS Campus Drive, Cary, NC, 27513, United States, sertalpbilal@gmail.com, Jared Erickson

Python has become a popular programming language for both data analytics and mathematical optimization. With SAS® Viya™ and its Python interface, Python programmers can use the state-of-the-art optimization solvers that SAS® provides. This talk features sasoptpy a new Python optimization modeling interface package for SAS® Viya™ optimization solvers and demonstrates approaches for Python programmers to naturally model their optimization problems, solve them by using SAS® Optimization solver actions, and view and interact with the results.

2 - Gravity: A Modeling Language for Mathematical Optimization and Machine Learning

Hassan Lionel Hijazi, Los Alamos National Laboratory, P.O. Box 1663, MS B284, Los Alamos, NM, 87545, United States, hlh@lanl.gov

Gravity is an open source, scalable, memory efficient modeling language for solving mathematical models in Optimization and Machine Learning. It exploits structure to reduce function evaluation time including Jacobian and Hessian computation. Gravity is implemented in c++ with a flexible interface allowing the user to specify the numerical accuracy of variables and parameters. It is also designed to handle iterative model solving, convexity detection, distributed algorithms, and constraint generation approaches. When compared to state-of-the-art modeling languages such as JuMP, Gravity is 5 times faster in terms of function evaluation and up to 60 times more memory efficient. It also dominates commercial languages such as Ampl on structured models including quadratically-constrained and polynomial programs. This short paper serves as a quick introduction to the language, presenting its main features along with some preliminary results, an extended version of the work is in progress.

3 - Practical Utilization of a Progressive Hedging Algorithm for Solving Optimization Problems in Energy Systems

Devon Sigler, Postdoc, National Renewable Energy Lab, Denver, CO, United States, devon.sigler@nrel.gov, Wesley Jones, Monte Lunacek

Solving optimization problems with sources of uncertainty is an important field in the discipline of optimization. Many optimization problems subject to uncertainty can be formulated as scenario based multi-stage stochastic problems which exhibit problem structure that can be exploited via parallel computing. Optimization frameworks in production software are now available which exploit the structure of these formulations to solve them in parallel. We consider the optimization for planning the expansion of infrastructure and investigate an application of one such optimization framework. Traditionally, this is limited by implementations that principally only take advantage of a single large server and are thus constrained with respect to problem complexity and size (i.e. limited by both memory and computing power). We investigate an implementation of this framework which enables the use of multiple servers and examine the computational performance.

4 - Roundoff-error-free Framework for the Exact Solution of Sparse Linear Systems

Christopher Lourenco, Texas A&M University, 3131 TAMU, College Station, TX, 77843, United States, clouren@tamu.edu, Adolfo Raphael Escobedo, Erick Moreno-Centeno, Timothy Davis

LU factorizations are the key tool used to solve the sparse systems of linear equations that arise in linear and integer programming. In many documented cases, however, nontrivial roundoff errors accrued during the construction and implementation of these factorizations can lead solvers to claim suboptimal bases as optimal, infeasible bases as optimal, or optimal bases as infeasible. To address this issue, we develop an exact left-looking LU factorization framework to solve sparse linear systems in which all operations are performed entirely in integer arithmetic. We also present computational results in which we show that the novel LU factorization framework significantly outperforms a modern state-of-the-art exact sparse solver.

■ FC07

North Classroom 3004

Pricing and Inventory

Contributed Session

Chair: Omid Nohadani, Northwestern University, 2145 Sheridan Road, Technological Institute M233, Evanston, IL, 60208-3119, United States, nohadani@northwestern.edu

1 - Dynamic Pricing of Time Slots for Attended Home Delivery

Nalan Gulpinar, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, United Kingdom, nalan.gulpinar@wbs.ac.uk, Arne Strauss, Yijun Zheng

In this paper, we study dynamic pricing of regular and flexible time slots offered for the attended home delivery management. A flexible slot is designed as a combination of several regular time slots in which customer's orders need to be delivered. Since the underlying dynamic pricing problem suffers due to curse of dimensionality, we propose an approximation approach to obtain the dynamic pricing policy. In order to illustrate performance of this approach, we design numerical experiments based on realistically-sized scenarios. The computational results show that expected profit increases when adding flexible time slots rather than using only regular time slots.

2 - Optimization & Alignment of Production Level, Quality and Warranty Contract for Return Policy

Shirsendu Nandi, Assistant Professor, Indian Institute of Management Rohtak, Rohtak, Haryana, Rohtak, 124001, India, shir.05math@gmail.com

The present research explores the coordination mechanism through warranty period optimization when buyback contract and quantity flexibility contract are implemented under stochastic demand. It is assumed that the product faces a stochastic demand and the demand is also dependent on the length of warranty period offered. The manufacturer offers a free replacement warranty to the customer if the product fails within a specified time interval after sales. The present study determines the risk expressions for the retailer, manufacturer and the entire supply chain in case of buyback contract, quantity flexibility contract in conjunction with warranty. Risk is measured by calculating the variances in absolute profits of the respective parties in the supply chain. It also investigates the effect of change in order quantity on the risks borne by the manufacturer, the retailer and the total supply chain while the warranty length is kept optimal. Similar investigation is performed to examine the impact of change in warranty length upon the risks borne by the manufacturer, the retailer and the total supply chain keeping the order quantity optimal. The study also examines how the solution of the centralised system is affected when the supply chain is risk-averse in nature. The study examines the necessity of aligning the quality decision, warranty policy and production level of a supply chain by investigating the impact of quality improvement of the product on the optimal order quantity, optimal warranty length, optimal supply chain profit and profits of both the parties of the supply chain. The study also provides a guideline to recoordinate the supply chain by redesigning the contract parameters after the quality development of the product in case of buyback contract and quantity flexibility contract.

3 - Finite-horizon Approximate Linear Programs for an Infinite-Horizon Revenue Management Problem

Thomas Vossen, University of Colorado-Boulder, Leeds School of Business, UCB419, Boulder, CO, 80309, United States, thomas.vossen@colorado.edu, Fan You, Dan Zhang

We consider a rolling-horizon revenue management problem that can be formulated as an infinite horizon discounted cost Markov Decision Process. We consider affine and finite-horizon approximations, and show these admit compact representations that can be solved efficiently. The resulting approximations can be used to construct control policies that probabilistically allocate demand, and we use a numerical study to evaluate their performance.

4 - Sustainable Inventory with Robust Periodic-affine Policies and Application to Medical Supply Chains

Omid Nohadani, Northwestern University, 2145 Sheridan Road, Technological Institute M233, Evanston, IL, 60208-3119, United States, nohadani@northwestern.edu, Chaithanya Bandi, Eojin Han

We introduce a new class of adaptive policies called periodic-affine policies, that allows a decision maker to optimally manage and control large-scale newsvendor networks in the presence of uncertain demand without distributional assumptions. These policies are data-driven and model many features of the demand such as correlation, and remain robust to parameter mis-specification. We present a model that can be generalized to multi-product settings and extended to multi-period problems. This is accomplished by modeling the uncertain demand via sets. In this way, it offers a natural framework to study competing policies such as base-stock, affine, and approximative approaches with respect to their profit, sensitivity to parameters and assumptions, and computational scalability. We show that the periodic-affine policies are sustainable, i.e. time consistent, because they warrant optimality both within subperiods and over the entire planning horizon. This approach is tractable and free of distributional assumptions, and hence, suited for real-world applications. We provide efficient algorithms to obtain the optimal periodic-affine policies and demonstrate their advantages on the sales data from one of India's largest pharmacy retailers.

■ FC08

North Classroom 3210

Complexity Analysis and Approximations

Contributed Session

Chair: Jamie Haddock, University of California-Davis, One Shields Avenue, Mathematics, Davis, CA, 95616, United States, jhaddock@math.ucdavis.edu

1 - Approximating the Worst Optimal Value in Interval Linear Programming

Mohsen Mohammadi Dehcheshmeh, University of Louisville, Louisville, KY, United States, m0moha15@louisville.edu, Monica Gentili

One of the basic tools to describe uncertainty in a linear programming model is interval linear programming (ILP), where we assume that there are a priori known intervals in which parameters of the linear model can vary. Interval linear programming has applications in several areas, including portfolio selection problems environmental management, and interval matrix games. One of the main topics addressed in this context is determining the optimal value range, that is best and the worst of all the optimal values of the objective function among all the realizations of the uncertain parameters. When the feasible set of the ILP is described by a set of equality constraints, computing the best optimal value is an easy task, while the computation of the worst case optimal value is known to be NP-hard. In this study, we propose new methods to determine bounds for the worst case optimal value. The proposed methods are evaluated on randomly generated instances.

2 - The Minimum Euclidean-norm Point on a Convex Polytope: Wolfe's Combinatorial Algorithm is Exponential

Jamie Haddock, University of California-Davis, One Shields Avenue, Mathematics, Davis, CA, 95616, United States, jhaddock@math.ucdavis.edu

The complexity of Philip Wolfe's method for the minimum Euclidean-norm point problem over a convex polytope has remained unknown since he proposed the method in 1974. The method is important because it is used as a subroutine for one of the most practical algorithms for submodular function minimization. We present the first example that Wolfe's method takes exponential time. Additionally, we improve previous results to show that linear programming reduces in strongly-polynomial time to the minimum norm point problem over a simplex.

■ FC09

Student Commons 1500

Recent Advances in First and Second Order Methods

General Session

Chair: Mingyi Hong, University of Minnesota, Minneapolis, MN, 55455, United States, mhong@umn.edu

1 - Condition Number Analysis of Logistic Regression and its Implications for First Order Solution Methods

Robert Freund, Massachusetts Institute of Technology, 77 Massachusetts Avenue E62-567, Cambridge, MA, United States, rfreund@mit.edu, Paul Grigas, Rahul Mazumder

We introduce a pair of condition numbers associated with the sample data for logistic regression that measures the degree of separability or non-separability of the data sample. When the sample data is not separable (as is routinely the case in logistic regression), the degree of non-separability naturally enters the analysis and the computational properties of standard first-order methods such as steepest descent, greedy coordinate descent, stochastic gradient descent, etc. When the sample data is separable — in which case the logistic regression problem has no solution — the degree of separability can be used to show rather surprisingly that standard first-order methods also deliver approximate-maximum-margin solutions with associated computational guarantees as well. The guarantees we develop hold for any dataset.

2 - When Does Stochastic Gradient Algorithm Work Well?

Lam M. Nguyen, PhD Student, Lehigh University, 751 E 6th Street, Apt. A2, Bethlehem, PA, 18015, United States, LamNguyen.MLTD@gmail.com, Nam H. Nguyen, Dzung Phan, Jayant Kalagnanam, Katya Scheinberg

We consider a general stochastic optimization problem which is often at the core of supervised learning, such as deep learning and linear classification. We consider a standard stochastic gradient descent method with a fixed step size and propose a set of assumptions on objective function. We then empirically demonstrate that these assumptions hold for logistic regression and standard deep neural networks on classical data sets. Thus our analysis helps explain when efficient behavior can be expected from the SGD method in training classification models and deep neural networks.

3 - A Smoothing First-order Method for Piecewise Linear Non-convex Optimization

Qihang Lin, University of Iowa, Iowa City, IA, 52242, United States, qihang-lin@uiowa.edu

Piecewise linear objective functions, convex or non-convex, have been introduced in optimization models from many applications including statistical machine learning, image processing and signal processing. We propose a smoothing first-order method for solving the unconstrained optimization with a piecewise linear objective function with a focus on the case where the objective function is non-convex. We characterize the convergence property of the iterative solutions generated by the smoothing first-order method under a homotopy updating scheme on the smoothing parameter.

4 - Learning Deep Models: Critical Points and Local Openness

Meisam Razaviyayn, University of Southern California, 3715 McClintock Ave, Los Angeles, CA, 90089, United States, razaviya@usc.edu, Maher Nouiehed

With the increasing interest in deeper understanding of the loss surface of many non-convex deep models , this paper presents a unifying framework to study the local/global equivalence of the optimization problem arising from training of such non-convex models. Using the local openness property of the underlying training models, we provide sufficient conditions under which any local optimum of the resulting optimization problem is global. Our result unifies and extends many of the existing results in the literature. For example, our theory shows that when the input data matrix X is full row rank, all non-degenerate local optima of the optimization problem for training linear deep model with squared loss error are global minima. Moreover, for two layer linear models, we show that all degenerate critical points are either global or second order saddles and the non-degenerate local optima are global. Unlike many existing results in the literature, our result assumes no assumption on the target data matrix Y . For non-linear deep models having certain pyramidal structure with invertible activation functions, we can show global/local equivalence with no assumption on the differentiability of the activation function. Our results are the direct consequence of our main theorem that provides necessary and sufficient conditions for the matrix multiplication mapping to be locally open in its range.

■ FC10

Student Commons 1600

Bayesian Optimization

General Session

Chair: Matthias Poloczek, University of Arizona, Tucson, AZ, 85721, United States, poloczek@email.arizona.edu

1 - Toward Fully Automated Bayesian Optimization

Roman Garnett, Washington University in St. Louis, St. Louis, MO, United States, garnett@wustl.edu, Gustavo Malkomes, Chip Schlaff

Despite the success of kernel-based nonparametric methods, kernel selection still requires considerable expertise, and is often described as a “black art.” We present a sophisticated method for automatically searching for an appropriate kernel from an infinite space of potential choices. Previous efforts in this direction have focused on traversing a kernel grammar, only examining the data via computation of marginal likelihood. Our proposed search method is based on Bayesian optimization in model space, where we reason about model evidence as a function to be maximized. We explicitly reason about the data distribution and how it induces similarity between potential model choices in terms of the explanations they can offer for observed data. In this light, we construct a novel kernel between models to explain a given dataset. Our method is capable of finding a model that explains a given dataset well without any human assistance, often with fewer computations of model evidence than previous approaches, a claim we demonstrate empirically.

2 - Bayesian Optimization of Integrated Response Surfaces

Saul Toscano, Cornell University, 113 Lake Street, Ithaca, NY, 14850, United States, st684@cornell.edu, Peter Frazier

We propose a one-step optimal Bayesian global optimization algorithm that takes advantage of the structure of expensive objective functions that can be written as a sums or integrals to optimize problems with as few evaluations as possible. That structure is commonly found in multi-task Bayesian optimization, sequential design of experiments with random environmental conditions, and optimization via simulation. We demonstrate that our method leads to state-of-the-art performance when the objective is noisy on realistic examples based on a queuing and inventory problem simulation, and produces excellent results in the noise-free setting when training convolutional neural networks and recommender systems.

3 - Bayesian Optimization of High-Dimensional Problems

Mickael Binois, The University of Chicago, Chicago, IL, United States, mbinois@uchicago.edu, David Ginsbourger, Olivier Roustant

The challenge of taking into account many variables in optimization problems may be overcome under the hypothesis of low effective dimensionality. Then, the search of solutions can be reduced to the random embedding of a low dimensional space into the original one, resulting in a more manageable optimization problem. Specifically, in the case of time consuming black-box functions and when the budget of evaluations is severely limited, Bayesian optimization with random embeddings appears as a sound alternative to random search. Yet, in the case of box constraints on variables, defining suitable bounds on a low dimensional domain appears to be complex. Indeed, a small search domain does not guarantee to find a solution even under restrictive hypothesis about the function, while a larger one may slow down convergence dramatically. Here we tackle this issue of low-dimensional domain selection based on a detailed study of the properties of the random embedding, giving insight on the aforementioned difficulties. In particular, we describe a minimal low-dimensional set in correspondence with the embedded search space. We additionally show that an alternative equivalent embedding procedure yields simultaneously a simpler definition of the low-dimensional minimal set and better properties in practice. Finally, the performance and robustness gains of the proposed enhancements for Bayesian optimization are illustrated on three examples.

4 - Multi-information Source Optimization

Matthias Poloczek, University of Arizona, 1127 East James E. Rogers Way, Room 268, Tucson, AZ, 85721, United States, poloczek@email.arizona.edu, Jialei Wang, Peter Frazier

Bayesian optimization methods have been applied with great success to global optimization of expensive-to-evaluate functions in machine learning, engineering, healthcare, and other areas. While traditional approaches only query the expensive-to-evaluate objective, we also often have access to other information sources: when optimizing an aerodynamic design, we may assess its performance by wind tunnel studies, or by CFD simulations with varying mesh sizes; when optimizing an inventory management system, we may evaluate it in real life at the client's warehouse, or by discrete-event simulations that vary in length and number of replications. These approximations are typically subject to an unknown bias in addition to common noise. This so-called model discrepancy results from an inherent inability to model the reality accurately, e.g., due to limited physical models. In this talk I will present a rigorous mathematical treatment of the uncertainties arising from model discrepancy and noisy observations that allows us to leverage these information sources effectively. I propose novel knowledge gradient algorithms for choosing which information source to query at each point in time. These algorithms rely on a value of information analysis and maximize the predicted information gain per unit cost. Experimental results demonstrate that utilizing additional information sources improves the performance significantly beyond what could be accomplished through traditional methods. In addition, I will discuss recent work on solving sequences of related problems by warm-starting Bayesian optimization.

■ FC11

Student Commons 2500

New Paradigms for Cut Generation

General Session

Chair: Thiago Serra, Carnegie Mellon University, Pittsburgh, PA, 15213, United States, tserra@cmu.edu

1 - Generalized Chvatal-Gomory Closures for Integer Programs with Bound Constraints

Dabeen Lee, Carnegie Mellon University, Pittsburgh, PA, 15213, United States, dabeenl@andrew.cmu.edu, Sanjeeb Dash, Oktay Gunluk

Integer programming problems that arise in practice often involve nonnegative or bounded decision variables. Using information about bounds on variables, one can generate, possibly stronger cuts valid for all integer feasible solutions. In this paper, we consider a natural extension of Chvatal-Gomory inequalities, which are obtained by strengthening Chvatal-Gomory inequalities based on bound constraints on some variables. These strengthened Chvatal-Gomory inequalities can also be viewed as cutting-planes from "wide split disjunctions", introduced by Bonami, Lodi, Tramontani, and Wiese recently. We prove that the closure of a rational polyhedron obtained after applying the generalized Chvatal-Gomory inequalities is also a rational polyhedron.

2 - Cutting Plane Techniques via Decision Diagrams

Danial Davarnia, Carnegie Mellon University, 5562 Hobart Street Pittsburgh, PA, 15217, United States, ddavarni@andrew.cmu.edu, Willem-Jan Van Hoeve

As an alternative to traditional integer programming (IP), decision diagrams (DD) provide a new solution technology based on the combinatorial structure of discrete problems using basics of dynamic programming. While the literature mainly focuses on the competitive aspects of DD as a stand-alone solver, we study IP techniques that can be derived from DD and used in conjunction with IP to enhance its performance. We develop linear programming and subgradient-type methods to generate valid inequalities for the convex hull of the feasible region described by DD. For convex IPs, these cutting planes dominate the so-called linearized cuts used in the outer approximation schemes. These cutting planes can also be derived for nonconvex IPs, which allows for an adaptation of the outer approximation framework. Computational experiments show significant gap improvement upon the traditional cutting plane methods employed in the state-of-the-art solvers.

3 - On Checking the Regularity of Lift-and-project Cuts from Non-split Disjunctions

Thiago Serra, Carnegie Mellon University, Pittsburgh, PA, 15213, United States, tserra@cmu.edu, Egon Balas

Many techniques to generate cutting planes for a Mixed-Integer Linear Program (MILP) are equivalent to one another under certain conditions. Since some are more general and thereby more expensive computationally, it is important to determine if and when they generate cuts that others cannot. Balas and Perregaard (2003) have shown that there is a correspondence between lift-and-project cuts obtained from basic solutions of the Cut Generating Linear Program (CGLP) and intersection cuts from basic solutions of the LP relaxation, feasible or not, and thus also to Gomory fractional cuts. More recently, Balas and Kis (2016) have shown that such correspondence may also hold for some lift-and-project cuts from arbitrary disjunctions. This work has four contributions. First, we state a result that simplifies the verification of regularity for basic CGLP solutions from Balas and Kis (2016) and show that it can also be used with CGLP solutions that are not basic. Second, we introduce and prove the validity of an MILP that checks whether there is a regular CGLP solution for a given cut. Third, we describe a numerical procedure based on such MILP that verifies if a lift-and-project cut is regular or not. Finally, we present and analyze computational results on the regularity of cuts from not-split disjunctions for several instances from the MIPLIB 2, 3, and 2003 benchmarks.

■ FC12

Student Commons 2504

Graphs and Trees

Contributed Session

Chair: David Phillips, U.S. Naval Academy, Math Department, Annapolis, MD, 21401, United States, dphillip@usna.edu

1 - Column Generation Approach to the Convex Recoloring Problem on a Tree

Sangho Shim, Assistant Professor, Robert Morris University, Department of Engineering, 6001 University Blvd., Moon Township, PA, 15108, United States, shim@rmu.edu, Sunil Chopra, Ergin Erdem, Eunseok Kim

The convex recoloring (CR) problem is to recolor the nodes of a colored graph at minimum number of color changes such that each color induces a connected subgraph. We adjust to the convex recoloring problem the column generation framework developed by Johnson et al. (Math Program 62:133-151, 1993). For the convex recoloring problem on a tree, the subproblem to generate columns can be solved in polynomial time by a dynamic programming algorithm. The column generation framework solves the convex recoloring problem on a tree with a large number of colors extremely fast.

2 - A Decomposition Approach to Solve the Selective Graph Coloring Problem in Some Graph Classes

Oylum Seker, Bogaziçi University, Bogaziçi Ünv. Muhendislik Fak. Endustri Muh., Istanbul, 34342, Turkey, oylum.seker@boun.edu.tr, Tınaz Ekim, Z. Caner Taskın

Graph coloring is the problem of assigning minimum number of colors to vertices of a graph such that no two adjacent vertices receive the same color. Selective Graph Coloring Problem is a generalization of the standard graph coloring problem; given a graph with a partition of its vertex set into clusters, the objective is to choose exactly one vertex per cluster so that, among all possible selections, the number of colors necessary to color the vertices in the selection is minimum. The selective graph coloring problem is known to be NP-hard, and remains so in many special classes of graphs. In this study, we focus on a decomposition based exact solution framework for selective coloring in perfect graphs. We test our method on graphs with various size and edge densities, present computational results for perfect graphs in general and some perfect graph families; in particular, permutation, generalized split and chordal graphs. Our computational experiments indicate that our decomposition approach significantly improves solution performance compared to a pure integer programming formulation. We also generalize our method to solve the selective coloring problem on graphs with no particular structure.

3 - Algorithms and Complexity Results for Designing Graphs with Extremal Randic Index

David Phillips, Associate Professor, U.S. Naval Academy, Chauvenet Hall, 572C Holloway Road, Annapolis, MD, 21401, United States, dphillip@usna.edu, M. Drew Lamar, Rex Kincaid, Sarah Kunkler

We show that finding a subgraph realization with the minimum generalized Randic index for a given base graph and degree sequence is solvable in polynomial time by formulating the problem as the minimum weight perfect b-matching problem of Edmonds. However, the realization found via this reduction is not guaranteed to be connected. Approximating the minimum weight perfect b-matching problem subject to a connectivity constraint is shown to be NP-hard. For instances in which the optimal solution to the minimum Randic index problem is not connected, we describe a heuristic to connect the graph using pairwise edge exchanges that preserves the degree sequence. Although we focus on finding graph realizations with minimum Randic index, our results extend to finding graph realizations with maximum Randic index as well. Applications of the Randic index are provided to normalizing cortical thickness networks in diagnosing individuals with dementia.

■ FC13

Student Commons 2600

Data Driven Optimization and Learning

General Session

Chair: Paul Grigas, UC Berkeley, Berkeley, CA, 94720-1777, United States, pgrigas@berkeley.edu

1 - A Dual Approach for Demand-side Platform Optimization

Alfonso Lobos Ruiz, University of California-Berkeley, Berkeley, CA, United States, alobos@berkeley.edu, Paul Grigas, Zheng Wen, Kuang-Chih Lee

We develop an optimization model and corresponding algorithm for the management of a demand-side platform (DSP), whereby the DSP acquires valuable impressions for its advertiser clients. The DSP aims to maximize its profit while having a proper budget spending for its advertisers, the latter represented by the usage of a utility over the budget spending. The DSP interacts with ad exchanges in a real-time bidding environment in a cost-per-click/cost-per-action pricing model. Our proposed formulation leads to a nonconvex optimization problem due to the joint optimization over both impression allocation and bid price decisions. We take the dual of our problem using Fenchel Conjugates and prove that under fairly general assumptions our dual formulation obtain the same optimal value as the original non-convex formulation. We also show in which cases we can recover a close to optimal solution for the non-convex problem using a primal-dual scheme. We perform simulations over datasets inspired in real data mimicking how our algorithm would be used in practice by a DSP. Our results show that our algorithm outperforms a greedy heuristic and how different utility functions alter the tradeoff between maximizing the DSP profit and the budget's usage of its advertisers.

2 - Smart "Predict, then Optimize"

Adam Elmachtoub, Assistant Professor, Columbia University, 500 W 120th St., New York, NY, 10027, United States, adam@ieor.columbia.edu, Paul Grigas

We consider a class of optimization problems where the objective function is not explicitly provided, but contextual information can be used to predict the objective based on historical data. A traditional approach would be to simply predict the objective based on minimizing prediction error, and then solve the corresponding optimization problem. Instead, we propose a prediction framework that leverages the structure of the optimization problem that will be solved given the prediction. We provide theoretical, algorithmic, and computational results to show the validity and practicality of our framework.

3 - Spectral Algorithms for Computing Fair Support Vector Machines

Matt Olfat, University of California-Berkeley, Berkeley, CA, United States, molfat@berkeley.edu, Anil Aswani

Classifiers and rating scores are prone to implicitly codifying biases, which may be present in the training data, against protected classes (i.e., age, gender, or race). So it is important to understand how to design classifiers and scores that prevent discrimination in predictions. This paper develops computationally tractable algorithms for designing accurate but fair support vector machines (SVM's). Our approach imposes a constraint on the covariance matrices conditioned on each protected class, which leads to a nonconvex quadratic constraint in the SVM formulation. We develop iterative algorithms to compute fair linear and kernel SVM's, which solve a sequence of relaxations constructed using a spectral decomposition of the nonconvex constraint. Its e ectiveness in achieving high prediction accuracy while ensuring fairness is shown through numerical experiments on several data sets.

4 - Lazy Conditional Gradients through Simpler Oracles

Sebastian Pokutta, Georgia Institute of Technology, H. Milton Stewart School of Isye, 765 Ferst Drive, Atlanta, GA, 30332-0205, United States, sebastian.pokutta@isye.gatech.edu

Conditional Gradient Descent methods are popular first-order methods for (smooth) constraint convex minimization. Relying on a linear programming oracle, these methods often outperform projected gradient descent methods whenever projections into the feasible region are expensive. Unfortunately, even those methods might suffer from prohibitive running times if the linear programming oracle itself is expensive, e.g., when the feasible region corresponds to a hard combinatorial optimization problem.
In this talk, we will explore a general method to significantly speed-up conditional gradient descent methods by replacing the linear programming oracle with a significantly easier and cheaper oracle leading to real-world speedups by several orders of magnitude while maintaining identical theoretical converge rates modulo (small!) constant factors. Moreover, we will further outline a conditionally accelerated lazy stochastic gradient descent method (CAL-SGD) that achieves optimal bounds in terms of required gradient evaluations and calls to the new oracle matching those for the more complex linear programming oracle. (based on joint works with Gábor Braun, George Lan, Yi Zhou, Daniel Zink)

Friday, 4:30PM - 5:30PM

■ Plenary II

Student Commons 2600

What Decision Diagrams Can do for You

Plenary Session

1 - What Decision Diagrams Can Do for You

John Hooker, Carnegie Mellon University, Tepper School of Business, Pittsburgh, PA, 15213, United States, jh38@andrew.cmu.edu

Decision diagrams have been used for decades as a compact representation of Boolean functions. More recently, they have emerged as a powerful tool for optimization. They provide a discrete relaxation of the problem that does not require linearity or convexity. The relaxation yields useful bounds and novel search strategies. This talk surveys recent applications of decision diagrams to discrete and nonlinear optimization, constraint programming, logic-based Benders decomposition, and comprehensive postoptimality analysis. Because a decision-diagram-based solver naturally accepts recursive dynamic programming (DP) models, it provides a new approach to solving DP problems by branch and bound rather than state space enumeration. In addition, use of a reduced decision diagram can sometimes lead to radical simplification of the state space.

Saturday, 8:30AM - 10:00AM

■ SA01

North Classroom 1806

First-Order Methods for Large-Scale Constrained Nonlinear Programming I

General Session

Chair: Yangyang Xu, Ressenlaer Polytechnic Institute, Troy, NY, 12180, United States, xuy21@rpi.edu

Co-Chair: Qihang Lin, The University of Iowa, Iowa City, IA, 52245, United States, qihang-lin@uiowa.edu

1 - Proximal Alternating Penalty Algorithms for Constrained Convex Optimization

Quoc Tran Dinh, Department of Statistics and Operations Research, UNC, USA, 333 Hanes Hall, CB# 3260 - UNC Chapel Hill, Chapel Hill, NC, 27599-3260, United States, quoctd@email.unc.edu

We develop two new proximal alternating penalty algorithms to solve a wide range class of constrained convex optimization problems. Our approach mainly relies on a novel combination of the classical quadratic penalty, alternating, Nesterov's acceleration, and homotopy techniques. The first algorithm is designed to solve generic and possibly nonsmooth constrained convex problems without requiring any Lipschitz gradient continuity or strong convexity, while achieves the best-known O(1/k)-convergence rate in the non-ergodic sense, where k is the iteration counter. The second algorithm is also designed to solve non-strongly convex problems, but with one strongly convex objective term. This algorithm achieves the O(1/k^2) - convergence rate on the primal constrained problem. Such a rate is obtained in two cases: (i) averaging only on the iterate sequence of the strongly convex term, or (ii) using two proximal operators of this term without averaging. In both algorithms, we allow one to linearize the second subproblem to use the proximal operator of the corresponding objective term. Then, we customize our methods to solve different convex problems, and lead to new variants. As a byproduct, these algorithms preserve the same convergence guarantees as in our main algorithms. Finally, we verify our theoretical development via different numerical examples and compare our methods with some existing state-of-the-art algorithms.

2 - New Lagrangian Dual Methods for Convex Programs with General Functional Constraints

Hao Yu, PhD, University of Southern California, yuhao@usc.edu, Michael Neely

We consider convex programs with a general (possibly non-differentiable) convex objective function and Lipschitz continuous convex inequality constraint functions. When the problem scale is large, these convex programs are difficult to solve by interior point methods or other Newton-type methods due to prohibitive computation and storage complexity for Hessians and matrix inversions. Instead, they are often solved by Lagrangian dual type methods which can often decompose the original problems. The dual subgradient method can deal nonlinear constraints but has slow O(1/ 2) convergence. The ADMM algorithm has faster O(1/) convergence but can only deal with linear constraints. In this talk, we present a new Lagrangian dual method that can deal with nonlinear constraint functions and has fast O(1/) convergence. Similar to the classical dual subgradient algorithm and the ADMM algorithm, this new Lagrangian dual method is parallel when objective and constraint functions are separable.

For large scale convex programs where the objective or constraint functions are not separable, the primal update of dual subgradient method can be difficult since it requires to solve unconstrained convex programs. Recall that the dual subgradient method has a close relative, primal-dual subgradient method, whose primal updates follow a gradient dynamic and is always parallel as long as primal variables are from a Cartesian set. However, the primal-dual subgradient method suffers the same slow O(1/ 2) converge as the dual subgradient method and requires to know an upper bound of optimal Lagrange multipliers. For convex programs with smooth or composite objective and constraint functions, we further propose another primal-dual version of our new Lagrangian dual method. This new primal-dual version is parallel even when the objective or constraint functions are not separable and is proven to enjoy the same O(1/) convergence as its original version. In addition, its implementation does not require any knowledge of the optimal Lagrange multiplier.

3 - Recent Advances in Coordinate Optimization

Mingyi Hong, University of Minnesota, 200 Union Street SE 4-174 Keller Hall, IMSE Department, Minneapolis, MN, 55455, United States, mhong@umn.edu

In this talk, we discuss a number of recent results on coordinate descent based optimization. First we present a double stochastic coordinate descent method, and show that for a number of convex problems including optimizing quadratic problems with polyhedral constraints, the algorithm is able to achieve linear convergence, while its deterministic counterparts can diverge. Second we present a block coordinate based primal-dual algorithm, and show that it is able to converges to second-order stationary solutions. An important implication of our latter result is that it gives rise to the first global convergence result (to the second-order stationary solutions) for two classes of unconstrained distributed non-convex optimization problems over multi-agent networks.

4 - A Stochastic Level Set Method for Convex Optimization with Expectation Constraints

Qihang Lin, The University of Iowa, 21 East Market Street, S380, Pappajohn Business Building, Iowa City, IA, 52245, United States, qihang-lin@uiowa.edu

We consider the problem of minimizing the expectation of a stochastic convex function subject to multiple inequality constraints. Each inequality in the constraints is also defined using the expectation of a stochastic convex function. The previous level-set framework extends the applicability of first-order methods to tackle problems with only deterministic objective and constraint functions. We develop a stochastic level-set method that finds an ϵ-optimal solution for the problem with stochastic objective and constraint functions. We establish the iteration complexity of our approach based on whether the strong convexity holds in the objective function and constraints.

■ SA02

North Classroom 1602

Decomposition and Parallelization for Stochastic Programs

General Session

Chair: Siqian Shen, University of Michigan, Ann Arbor, MI, 48109, United States, siqian@umich.edu

1 - Optimal Black Start Allocation for Power System Restoration

Georgios Patsakis, University of California Berkeley, 1460 Cedar St, Berkeley, CA, 94702, United States, gpatsakis@berkeley.edu, Ignacio Aravena, Deepak Rajan, Shmuel S Oren

Equipment failures, operator errors, natural disasters or cyber-attacks have caused extended blackouts of the electric grid in the past. Since most of the generating units cannot restart without connecting to an energized power grid, the system operator relies on a few units with the ability to start autonomously, called Black Start (BS) units, to restore the power system. Allocating and maintaining these units across the grid is costly, but can severely impact the restoration time for critical loads and infrastructure, as well as the restoration security. System operators currently employ expert judgment to make decisions regarding the placement of these units (BS Allocation) and heuristically predefined sequences for restoring the grid back to its initial state (Power System Restoration). In this work, we formulate an optimization problem for BS Allocation, while simultaneously optimizing over the restoration sequence. We extend existing models for restoration by including grid considerations such as active power nodal balance, transmission switching, reactive power support and voltage limits. The resulting problem is a large scale Mixed Integer Program (MIP) that utilizes binary variables for the allocation of BS resources and for the energization state of generators, lines and nodes in a finite time horizon. In order to aid the branch and bound tree we develop a randomized heuristic that is executed multiple times in parallel on a high-performance computing environment to find feasible solutions to the problem. The approach is tested on the IEEE-39 and on the IEEE-118 systems and the results are reported. In order to handle larger systems and stochastic variants, we recognize the need to decompose the problem by time as well. We first examine a special case, the Transmission Switching problem, that only has binary variables for line energization and assumes energized generators and nodes. A combination of Bender's cuts, No-Good-Cuts and Integer L-shaped cuts are tested to solve the problem. Our ultimate goal is to generalize this approach for the Power System Restoration problem.

2 - Level Regularization for Multistage Stochastic Programs

Yongjia Song, Virginia Commonwealth University, 1015 Floyd Avenue, P.O. Box 843083, Richmond, VA, 23284, United States, ysong3@vcu.edu, Wim van Ackooij, Welington de Oliveira

We consider well-known decomposition techniques for multistage stochastic programming and a new stabilization scheme based on normal solutions. The proposed algorithms combine ideas from finite perturbation of convex programs and level bundle methods to regularize the so-called forward step of these decomposition methods. Numerical experiments on a hydrothermal scheduling problem indicate that our algorithms are competitive with the state-of-the-art, e.g., the multistage regularized decomposition and the stochastic dual dynamic programming methods in the literature.

3 - Parallel Decomposition of Nonlinear Time-discretized Systems

J. Santiago Rodriguez, PhD Candidate, Purdue University, West Lafayette, IN, United States, rodri324@purdue.edu, Carl Laird, Bethany Nicholson

As the need for rapid solution of increasingly complex nonlinear dynamic optimization problems grows, the size of the resulting optimization problem can easily outstrip the computational capability of a single workstation. In this presentation, we describe multiple decomposition strategies for nonlinear time-discretized systems. Focusing on NLPs, we can parallelize the structured linear algebra (using Schur-complement-based decompositions), providing efficient scalability while retaining the convergence properties of the host algorithm. Problem-level decomposition approaches (e.g., progressive hedging, Benders' decomposition) can also be used. We will also demonstrate the scalability of PH on time-discretized systems with both continuous and discrete variables.

4 - Scenario Grouping and Decomposition Algorithms for Chance-constrained Programs

Siqian Shen, University of Michigan, Industrial & Operations Engineering, 1205 Beal Ave., Ann Arbor, MI, 48109, United States, siqian@umich.edu, Yan Deng, Shabbir Ahmed, Jon Lee

A lower bound for a finite-scenario chance-constrained problem is given by the quantile value corresponding to the sorted optimal objective values of scenario subproblems. This quantile bound can be improved by grouping subsets of scenarios at the expense of larger subproblems. The quality of the bound depends on how the scenarios are grouped. We formulate a mixed-integer bilevel program that optimally groups scenarios to tighten the quantile bounds. For general chance-constrained programs we propose a branch-and-cut algorithm to optimize the bilevel program, and for chance-constrained linear programs, we derive a mixed-integer linear programming reformulation. We also propose several heuristics for grouping similar or dissimilar scenarios. Our computational results show that optimal grouping bounds are much tighter than heuristic bounds, resulting in smaller root node gaps and better performance of the scenario decomposition algorithm for chance-constrained 0-1 programs. Moreover, the bounds from feasible grouping solutions obtained after solving the optimal grouping model for 20%-50% of the total time are sufficiently tight, having gaps under 10% of the corresponding optimal grouping bounds. They outperform heuristic grouping bounds both in tightness and solving time, and can be significantly strengthened using larger group size.

■ SA03

North Classroom 1202

Stochastic Programming Applications in Renewable Energy

General Session

Chair: Gokce Kahvecioglu,Northwestern University, Evanston, IL, 60208, United States, gokcekahvecioglu2014@u.northwestern.edu

1 - A Multi-agent Stochastic Programming Framework

Harsha Gangammanavar, Southern Methodist University, Department of EMIS, P.O. Box 750123, Dallas, TX, 75275, United States, harsha@smu.edu

In this talk we present a multi-agent stochastic programming framework to study a power system with a central grid connected to multiple microgrids. The microgrids are equipped to control their local generation and demand in the presence of uncertain renewable generation and heterogeneous energy management systems. The model is an extension of the classical two-stage stochastic program where the central grid occupies the role of the master, and the microgrids' energy management system optimization problems form the subproblems. We propose a sequential sampling-based optimization algorithm to tackle this framework. Our computational experiments, conducted on the US western interconnect (WECC-240) data set, illustrates that our approach can be used as a systematic optimization-simulation tool to gauge (a) the impact of energy management settings in efficiently utilizing renewable generation and (b) the role of flexible demands in reducing system costs.

2 - The Feasibility of Ambitious Renewable-integration Targets

Semih Atakan, University of Southern California, 3715 McClintock Ave, GER 240, Los Angeles, CA, 90089-0193, United States, atakan@usc.edu, Harsha Gangammanavar, Suvrajeet Sen

There is a growing debate among academics regarding the feasibility of ambitious renewable-integration targets. While the scope of the debate is broad, a particular concern, namely the uncertainty in renewable supplies, deserves better attention. The majority optimization models, which produces day-to-day decisions for the power industry, have 'deterministic' nature, therefore are ill-suited for uncertain environments. In this talk, we will demonstrate how 'stochastic' optimization frameworks can lead to better decisions, and assist the progress towards greater renewable integration.

3 - Stochastic Dispatch Optimization for Concentrated Solar Power Systems

Gokce Kahvecioglu, Northwestern University, 2145 Sheridan Road, Room C210, Evanston, IL, 60208, United States, gokcekahvecioglu2014@u.northwestern.edu, David Morton

Concentrated solar power (CSP) technology with thermal energy storage is dispatchable, i.e., it has the flexibility to adapt power output to the demand, even on a cloudy day or at night. While there are several key advantages of CSP, so-called direct normal irradiance (DNI), the source of power for CSP, is more challenging to accurately forecast compared to global irradiance that more pervasive PV systems use. We focus on producing scenario-based DNI estimates rather than just a point estimate of DNI via a new sampling method that allows temporal dependence and non-identical errors. We then input the scenarios into the stochastic dispatch optimization model that states how the system should operate over the next few hours, and positions the system to be well-hedged against the set of DNI scenarios.

■ SA04

North Classroom 1204

Stochastic Optimization

Contributed Session

Chair: Caleb Bugg, University of California-Berkeley, 4141 Etcheverry Hall, Industrial Engineering & Operations Research, Berkeley, CA, 94720, United States, caleb_bugg@berkeley.edu

1 - Distributionally Robust CVaR Formulation of SVMs using Wasserstein Metric

Ammon Washburn, Graduate Student, University of Arizona, 2525 N Los Altos Avenue, Apt 155, Tucson, AZ, 85705, United States, wammonj@email.arizona.edu, Neng Fan, Helen Hao Zhang

There has been many variants of Support Vector Machines that use different norms or penalty functions in order to accomplish feature selection, good generalization and robustness to outliers. We propose to unify all of them under a framework that comes from CVaR in risk management. Then we will show properties of the resulting quadratic program as a function of parameters and provide an algorithm that will compute the whole solution path for two out of three of the parameters in about the time it takes to solve one using interior points methods.

2 - Logarithmic Sample Bounds for Sample Average Approximation

Caleb Bugg, Graduate Student, University of California, Berkeley, Industrial Engineering & Operations Research, Berkeley, CA, United States, caleb_bugg@berkeley.edu, Anil Aswani, Deepak Rajan

The Sample Average Approximation (SAA) method is a commonly used approach for solving stochastic optimization problems approximately, and often works better in practice than existing theoretical bounds suggest for the number of samples needed to ensure the SAA minimum value is close to the true minimum value. In this paper, we derive new theoretical bounds for SAA that, for certain types of constraint sets, are logarithmic in problem dimension, whereas existing bounds are polynomial in dimension. Our approach characterizes the stability of random instances of the optimization problem using stochastic process theory, and then uses this characterization to construct confidence intervals using concentration of measure techniques. Notably, for single stage stochastic optimization problems, we find that the presence of an L1 constraint yields logarithmic bounds on the number of samples needed. This provides theoretical explanation for the success of SAA for capacity- or budget-constrained problems.

3 - Combined Worst and Average Case Considerations in an Integrated Emergency Preparedness Network Design Problem

Halit Uster, Southern Methodist University, Lyle School of Eng., Dept. EMIS, Dallas, TX, 75275-0123, United States, uster@smu.edu, Jyotirmoy Dalal

We study an emergency response network design problem that integrates relief (supply) and evacuation (demand) sides under disaster location and intensity uncertainties which, in turn, dictate uncertainty in terms of the location and the amount of demand. Representing these uncertainties by discrete scenarios, we present a stochastic programming framework in which two second stage objectives, the average and worst case costs, are combined. In our model, we minimize, over all the scenarios, the fixed costs of opening supply centers and shelters, and the weighted sum of average and worst case flow costs. Thus, the model gives the decision maker the flexibility to put relative emphasis on the worst case and average flow cost minimization and explore the outcomes in terms of total costs and network configurations. To solve large scale instances with varying relative weights, we devise alternative Benders Decomposition (BD) approaches. We implement these by using an advanced callback feature of the solver while simultaneously incorporating several performance-enhancing steps that help to improve runtimes significantly. We conduct a detailed computational study to highlight the efficiency of our proposed solution methodology. Furthermore, we also apply our approach in a case study based on data on coastal Texas and present interesting insights about the problem and the resulting network structures for varying weights assigned to objectives.

■ SA05

North Classroom 2001

Cliques, Clubs, and Dominating Sets

General Session

Chair: Austin Buchanan, Oklahoma State University, Stillwater, OK, 74078, United States, buchanan@okstate.edu

1 - Parsimonious Formulations for Low-diameter Clusters

Hosseinali Salemi, Oklahoma State University, 15 North University Place, Stillwater, OK, 74075, United States, hosseinali.salemi@okstate.edu, Austin Buchanan

In the analysis of social and biological networks, one often searches for tightly knit clusters. One property of a "good" cluster is that it have a small diameter, which leads to the concept of a k-club. We propose a new IP formulation for detecting k-clubs and show that a relatively simple implementation of it often outperforms previous approaches.

2 - Why is Maximum Clique Easy in Practice?

Jose Luis Walteros, University at Buffalo, SUNY, 413 Bell Hall, Buffalo, NY, 14260, United States, josewalt@buffalo.edu, Austin Buchanan

Recently, researchers have developed algorithms for solving the maximum clique problem that run rather quickly for million-node real-life graphs despite the computational intractability of the problem in the worst case. A natural explanation for the success of these approaches on real-life graphs is their sparsity, allowing one to safely delete many low-degree vertices in a preprocessing step. Related ideas lead to algorithms that solve maximum clique in time polynomial in the size of the graph, but exponential in the graph's degeneracy d, which is a measure of sparsity. This is encouraging given that d is orders of magnitude smaller than the number n of nodes on most real-life graphs. However, given that the approach is exponential in d and the fact that d is often in the hundreds on these "easy" instances, this explanation for their easiness is somewhat unsatisfying. This paper provides an alternative explanation based on the empirically observed proximity of the clique number to the graph's degeneracy d. We develop an algorithm for the maximum clique problem that runs in time polynomial in the graph's size, but exponential in the difference between d and the clique number. When this difference is a constant, the running time is O(dm)=O(m1.5). Key subroutines in our approach include existing kernelization and fixed-parameter tractable algorithms for vertex cover. Since for 70% of common test instances the difference between d and the clique number is less than 3, our implementation performs rather well and is competitive with previous approaches. Key subroutines in our approach include existing kernelization and fixed-parameter tractable algorithms for vertex cover. We feel that the tools and insights of parameterized complexity deserve a larger role in the operations research and mathematical programming communities and hope that this paper serves as an illustrative case study.

3 - The Power Domination Problem in Reconfigurable Graphs

Logan Smith, Rice University, 6100 Main St., Houston, TX, United States, logan.smith@rice.edu

The study of power domination in graphs arises from a monitoring problem in electrical networks. Various placements of sensors, known as Phasor Measurement Units (PMUs), allow for the observation of electrical networks. Due to the high monetary cost of PMUs, it is desirable to find arrangements of PMUs that can fully observe networks, while deploying a minimum number of sensors. When represented in a graph, the sets of sensor locations allowing full network observation are known as the graph's power dominating sets, and the cardinality of a minimum power dominating set is the power domination number of the graph. In addition to the incorporation of PMUs, several innovations for smart grids allow electrical networks to undergo spontaneous reconfigurations. This motivates a novel varient of the power domination problem. In this presentation, the power domination problem in reconfigurable graphs is formulated, and techniques for identifying minimum power dominating sets in reconfigurable graphs are discussed.

4 - The Optimal Design of Low Latency Virtual Backbones

Hamidreza Validi, Oklahoma State University, 72 South University Place, Stillwater, OK, 74075, United States, hamidreza.validi@okstate.edu, Austin Buchanan

In wireless networks, two nodes may not be close enough to communicate directly, necessitating the use of intermediate nodes for relaying information. Often, one seeks a (smallest possible) subset ofnodes to serve as designated relay nodes. In this talk, we discuss a hop-constrained variant of this optimization problem and propose an integer programming approach to solve it.

■ SA06

North Classroom 2002

Bilevel and Two-Stage Optimization

Contributed Session

Chair: Liang Xu, University of Pittsburgh, 3700 O'Hara Street, 1048 Benedum Hall, Pittsburgh, PA, 15261, United States, lix21@pitt.edu

1 - A Computational Method for Pessimistic Semivectorial Bilevel Optimization

Bo Zeng, University of Pittsburgh, 3700 O'Hara Street, Benedum Hall, Pittsburgh, PA, 15260, United States, bzeng@pitt.edu, Liang Xu

In this talk, we study the challenging semivectorial bilevel optimization based on the pessimistic assumption. Through analysis and reformulations, we convert it into a computationally friendly form and develop an effective solution procedure. Numerical results will be presented to demonstrate the strength of our computational method.

2 - Techniques for Scenario Creation in Two-stage Stochastic Programming Applied to Economic Dispatch under Uncertainty

Matthew Reynolds, National Renewable Energy Laboratory, Golden, CO, United States, matthew.reynolds@nrel.gov, Ryan King, Devon Sigler, Wesley Jones

We present a technique for creating scenarios to solve two-stage stochastic programming problems utilizing an importance distribution derived from the solution of a simpler, deterministic problem. Such an approach is useful for applications involving large problems, where computational complexity can limit the number of scenarios that can be used. Solving two-stage stochastic optimization problems with a small number of scenarios requires careful scenario selection to take into account not only the probability distribution of the random variables but also the second-stage loss function. Our importance sampling approach addresses this requirement by using both the nominal probability distribution and a surrogate loss function to compute a new probability distribution from which to draw scenarios. The target application for this technique is stochastic economic dispatch for power grids with uncertain generation (e.g. from renewable energy sources such as wind). If the grid under consideration has a large number of recourse variables, e.g. reserve generation and uncertain generation dispatch decisions, and we take into account reliability requirements, the size of the problem will grow rapidly with the inclusion of a large number of scenarios. Thus we use our importance sampling technique to carefully select a small set of scenarios that provide adequate tests of grid reliability. Applying our technique to a stochastic form of the classical economic dispatch problem results in improved convergence to the expectation of the second-stage cost and cheaper, yet more reliable economic dispatch decisions.

3 - Semi-pessimistic Bilevel Programs

Liang Xu, University of Pittsburgh, 3700 O'Hara Street, 1048 Benedum Hall, Pittsburgh, PA, 15261, United States, lix21@pitt.edu, Bo Zeng

For bilevel programs, the concepts of optimism and pessimism have been studied. However, it is not clear if the response variables appear in the upper level constraints. In this talk, we consider a few variants of classical bilevel programs. Both theoretical and numerical results will be discussed.

■ SA07

North Classroom 3004

Underground Mining Applications II

General Session

Chair: Tulay Flamand, Colorado School of Mines, Division of Economics and Business, Engineering Hall 816 15th Street, Room 313, Golden, CO, 80401, United States, tulayvarol@gmail.com

1 - Optimal Selection of Support Pillars in an Underground Mine

Levente Sipeki, Colorado School of Mines, Golden, CO, 80401, United States, levente.sipeki@gmail.com

One method to extract high-value ore that lies far below the earth's surface is called top-down open stope retreat mining. Using this mining method, a large volume of earth below the surface is notionally divided into three-dimensional rectangular blocks on each of several horizontal layers. Some of the blocks are left behind as pillars, satisfying geotechnical structural stability constraints, and the remainder are extracted and processed to obtain ore. We present an optimization-based iterative heuristic to determine which blocks are extracted and which are left in-situ with the objective of maximizing total profit. The four main types of constraints to ensure geotechnical stability are: (i) limits on the size of contiguous voids between adjacent pillars, (ii) restrictions on the length-to-width ratios of the pillars; (iii) the extraction ratio, i.e., the number of blocks designated as stopes in a given area; and (iv) the minimum distance between two semi-parallel ore bodies to be treated independently. Our integer programming optimization model contains 40,000 variables and 11,000 constraints for the instance that motivated our study. The heuristic converges to a geotechnically stable mine design in under one hour when applied to a mine in Africa.

2 - Analysis of an Evacuation Plan After an Earthquake in the Sector "El Progreso" in Carabayllo

Miguel Eduardo Ramírez Huamán, Student, Pontificia Universidad Católica del Perú, Av. Universitaria 1801 - San Miguel, Lima, 51, Peru, mramirezh@pucp.pe, Miguel Rodríguez Anticona, Daniel Yupanqui Santiago

Lima is a city of 10 million inhabitants, where approximately 60% live in slums, which lack any urban plan and even less a plan for natural disasters despite the fact that Lima is located in the belt of fire of the Pacific where seismic activity is high. After decades of neglect by Latin American governments towards this population, it is imperative to demonstrate quantitatively the level of risk in which they are exposed and to propose an evacuation plan to mitigate, as far as possible, the effects of a post-earthquake. The "El Progreso" sector located in Carabayllo in the district of Lima in Peru was chosen as a case study because it is one of the slums with the greatest potential risk because it is a basin surrounded by hills that due to the effects of informal constructions (such as ceilings). Filled with rocks and walls their slopes have suffered much more deterioration than in other hillsides. In addition to being in an area prone to debris avalanches, rocks, debris flows and other types of geological hazards presented in this area. First we will identify all the risks such as the height of the buildings near the escape routes, illegally parked vehicles (mainly taxis and Peruvian motorcycle taxis), current traffic on the streets, distance to the evacuation points, flow capacity of the escape routes, flow of people evacuating in the opposite direction, slope, type of floor of the escape route, danger of avalanche, falling rocks. Subsequently the risks will be classified using the hierarchical analysis process to obtain a cost function of the roads, then we will formulate a mathematical model to find the most optimal evacuation plan, minimizing the risk function and the time, having as one of the most important restrictions. The capacity of the evacuation areas that are in this slum include sports fields and green areas. This mathematical model could be replicated in real time to be able to guide evacuation with greater efficiency.

3 - Heuristic Approaches for the Optimal Deployment Problem

Tulay Flamand, Colorado School of Mines, Division of Economics and Business, Engineering Hall 816 15th Street, Golden, CO, 80401, United States, tulayvarol@gmail.com

We address a deployment problem, in the case of a natural disaster, for the optimal assignment of commodities by several modes from various origins to the disaster area while minimizing travel time. Effective column generation based heuristics are proposed to solve challenging instances for this problem.

■ SA08

North Classroom 3210

Stochastic and Distributed Algorithms for Nonlinear Optimization II

General Session

Chair: Shiqian Ma, University of California-Davis, Davis, CA, United States, sqma@math.ucdavis.edu

Co-Chair: Necdet Serhat Aybat, Penn State University, University Park, PA, 16802, United States, nsa10@psu.edu

1 - Directly and Efficiently Optimizing Prediction Error and AUC of Linear Classifiers

Hiva Ghanbari, Lehigh University, 11 Duh Drive, Bethlehem, PA, 18015, United States, hig213@lehigh.edu, Katya Scheinberg

The predictive quality of most machine learning models is measured by the expected pre- diction error or so-called Area Under the Curve (AUC). However, these functions are not used in the empirical loss minimization, because their empirical approximations are non- convex and nonsmooth, and more importantly have zero derivative almost everywhere. Instead, other loss functions are used, such as the logistic loss. In this work, we show that in the case of linear predictors, and under the assumption that the data has normal distribution, the expected error and the expected AUC are not only smooth, but have well defined derivatives, which can be computed given on the first and second moments of the normal distribution. We show that these derivatives can be also approximated and used in empirical risk minimization, thus proposing a gradient-based optimization methods for direct optimization of prediction error and AUC. Moreover the proposed algorithm has no dependence on the size of the dataset, unlike logistic regression and all other well known empirical risk minimization techniques.

2 - A Stochastic Trust Region Method

Rui Shi, Lehigh University, Bethlehem, PA, United States, rus415@lehigh.edu

In this talk, we present a new stochastic trust region method, deemed TRish, for solving stochastic and finite-sum minimization problems. We motivate our approach by illustrating how it can be derived from a trust region methodology. However, we also illustrate how a direct adaptation of a trust region methodology might fail to lead to general convergence guarantees. Hence, our approach involves a modified update scheme, which we prove possesses convergence guarantees that are similar to those for a traditional stochastic gradient (SG) method. We also present numerical results showing that TRish can outperform SG when solving convex and nonconvex machine learning test problems.

3 - A Varying-coefficient Regularized Dual Averaging Algorithm for Regularized Stochastic Optimization

Shiqian Ma, UC Davis, Davis, CA, 95616, United States, sqma@ucdavis.edu

In this talk, we propose a varying-coefficient regularized dual averaging (RDA) algorithm that combines the advantages of SGD and RDA with convergence guarantees. A novel adaptive scaling scheme is incorporated to further accelerate this algorithm. Numerical results indicate that our new method outperforms existing algorithms such as SGD, RDA and FTRL-Proximal.

■ SA09

Student Commons 1500

Continuous and Discrete Optimization in Statistics

General Session

Chair: Hongbo Dong, Washington State University, Pullman, WA, 99163-5142, United States, hongbo.dong@wsu.edu

1 - Nonconvex Statistical Learning of Hierarchical Interactions

Lingzhou Xue, Penn State University, State College, PA, United States, lingzhou@psu.edu

In this talk, I will provide the nonconvex statistical learning of higher-order interactions in high-dimensional graphical models. I will introduce the folded concave penalized composite conditional likelihood approach for learning sparse Ising models with higher-order interactions, where folded concave functions are used to approximate the affine sparsity constraints. Simulation studies and real applications will be presented to demonstrate the power of our proposed methods.

2 - New Algorithms for Statistical Learning Problems with Structured Sparsity

Sam Davanloo Tajbakhsh, The Ohio State University, 210 Baker Systems Building, 1971 Neil Ave, Columbus, OH, 43210, United States, sam.tajbakhsh@gmail.com, Dewei Zhang

Sparsity-inducing penalties are commonly used in statistical learning for variable selection. While these penalty functions help to obtain a sparse solution, they cannot guarantee specific sparsity structures. In some problems, however, there exist logical sparsity relations between variables or groups of variables that are known a priori. Being able to utilize such structures helps to obtain more interpretable models. Our study focuses on hierarchical sparsity structures that can be represented as Directed Acyclic Graphs (DAG). Designed penalty functions exploit group overlaps to induce solutions with desired hierarchical structures. These overlaps, however, complicate the problem from the optimization perspective. In this talk, we will present new algorithms to solve the underlying optimization problems with established complexity bounds. Some numerical results supporting the proposed algorithms will be provided.

3 - A New Estimator for the High-Dimensional Linear Regression Model

Abhishek Kaul, Assistant Professor, Washington State University, Neil Hall Rm 217, Pullman, WA, 99164, United States, akaul@math.wsu.edu

We propose a new estimator for the high-dimensional linear regression model with observation error in the design where the number of coefficients is potentially larger than the sample size. The main novelty of our procedure is that the choice of penalty parameters is pivotal. The estimator is based on applying a self-normalization to the constraints that characterize the estimator. Importantly, we show how to cast the computation of the estimator as the solution of a convex program with second order cone constraints. This allows the use of algorithms with theoretical guarantees and reliable implementation. Under sparsity assumptions, we derive Lq-rates of convergence and show that consistency can be achieved even if the number of regressors exceeds the sample size. We further provide a simple to implement rule to threshold the estimator that yields a provably sparse estimator with similar L2 and L1-rates of convergence.

■ SA10

Student Commons 1600

Algorithms for Large-scale Problems

General Session

Chair: Angelia Nedich, UIUC, Urbana, IL, 61801, United States, angelia@illinois.edu

Co-Chair: Farzad Yousefian, Oklahoma State Univeristy, Stillwater, OK, 74074, United States, farzad.yousefian@okstate.edu

1 - Curvature-aided Incremental Gradient Method

Hoi To Wai, Arizona State University, Tempe, AZ, United States, hwai2@asu.edu, Shi (Wilbur) Wai, Angelia Nedich

We propose the curvature-aided incremental aggregated gradient (CIAG) method for finite sum, d-dimensional optimization problems, where the objective function is a sum of m convex functions and $m \gg d \gg 1$. The CIAG method accelerates incremental aggregated gradient (IAG) methods using aids from the Hessian information, while avoiding the evaluation of matrix inverses required by the incremental Newton (IN) method. Our idea is to exploit the incrementally aggregated Hessian matrix to trace the full gradient vector at every incremental step, therefore achieving an improved linear convergence rate over the state-of-the-art IAG methods. For strongly convex problems, we show that running one iteration of the CIAG method yields the same improvement to optimality gap as running one iteration of the full gradient method, while the complexity per iteration is O(d^2) for CIAG and O(md) for the full gradient. Overall, the CIAG method strikes a balance between the high computation complexity incremental Newton-type methods and the slow IAG method. Our numerical results support the theoretical findings and show that the CIAG method often converges with much fewer iterations than IAG, and requires much shorter running time than IN when the problem dimension is high. We also discuss a Nesterov type acceleration to the CIAG method.

2 - Random Gradient Extrapolation for Distributed and Stochastic Optimization

Guanghui Lan, Georgia Institute of Technology, 755 Ferst Drive NW, Atlanta, GA, 30332, United States, george.lan@isye.gatech.edu, Yi Zhou

In this talk, we study the distributed finite-sum optimization problems defined over a star topology. By proposing an optimal randomized incremental gradient method, we show that without any exact gradient evaluation, the optimal linear sampling complexity bound can be achieved for solving deterministic finite-sum problems. Moreover, for stochastic finite-sum problems, the proposed algorithm maintains the optimal sublinear sampling complexity (up to a certain logarithmic factor), and attains a linear communication complexity.

3 - Randomized Block Coordinate Stochastic Mirror Prox Methods for Stochastic Cartesian Variational Inequalities

Farzad Yousefian, Oklahoma State University, 317D Engineering North, School of Industrial Engineering & Management, Stillwater, OK, 74074, United States, farzad.yousefian@okstate.edu, Angelia Nedich, Uday Shanbhag

Motivated by multi-user optimization problems and non-cooperative Nash games in uncertain regimes, we consider stochastic Cartesian variational inequalities (SCVI) where the set is given as the Cartesian product of a collection of component sets. When the number of the component sets is large, the classical stochastic approximation methods and their prox generalizations are computationally inefficient as each iteration becomes very costly. To address this challenge, we develop a randomized block stochastic mirror-prox (B-SMP) algorithm, where at each iteration only a randomly selected block coordinate of the solution is updated through implementing two consecutive projection steps. Under standard assumptions on the problem and settings of the algorithm, we show that when the mapping is strictly pseudo-monotone, the algorithm generates a sequence of iterates that converges to the solution of the problem almost surely. To derive rate statements, we assume that the maps are strongly pseudo-monotone and obtain a non-asymptotic mean squared error O(d/k), where k is the iteration number and d is the number of component sets. Second, we consider large-scale stochastic optimization problems with convex objectives. For this class of problems, we develop a new averaging scheme for the B-SMP algorithm. Unlike the classical averaging stochastic mirror-prox (SMP) method where a decreasing set of weights for the averaging sequence is used, here we consider a different set of weights that are characterized in terms of the stepsizes and a parameter. We show that using such weights, the objective values of the averaged sequence converges to the optimal value in the mean sense with the rate O(d/k^0.5). Both of the rate results appear to be new in the context of SMP algorithms.

■ SA11

Student Commons 2500

Constraint Programming and Integrated Methods

General Session

Chair: Willem-Jan Van Hoeve, Carnegie Mellon University, Pittsburgh, PA, 15213, United States, vanhoeve@andrew.cmu.edu

1 - A Tolerant Algebraic Side-channel Attack on AES using CP

Laurent Michel, University of Connecticut, Computer Science & Engineering, 371 Fairfield Road, Storrs, CT, 06269-2155, United States, ldm@engr.uconn.edu, Fanghui Liu, Waldemar Cruz, Chujiao Ma, Greg Johnson

AES is a mainstream block cipher used in many protocols and whose resilience against attack is essential for cybersecurity. Oren et al. recently discussed a Tolerant Algebraic Side-Channel Analysis (TASCA) and showed how to use optimization technology to exploit side-channel information and mount a computational attack against AES. This talk revisits the results and posits that Constraint Programming is a strong contender and a potent optimization solution. It extends bit-vector solving, develops a CP and an IP model and compares them with the original Pseudo-Boolean formulation. The empirical results establish that CP can deliver solutions with orders of magnitude improvement in both run time and memory usage, traits that are essential to potential adoption by cryptographers.

2 - A Local Search Framework for Compiling Relaxed Decision Diagrams

Michael Römer, Martin Luther University, Saale, Germany, michael.roemer@cirrelt.ca, Andre Augusto Cire, Louis-Martin Rousseau

This work presents a local search framework for constructing an improving relaxed decision diagrams (DDs). The framework consists of a set of elementary DD manipulation operations including a redirect operation introduced in this paper and a general algorithmic scheme. We show that the framework can be used to reproduce several standard DD compilation schemes and to create new compilation and improvement strategies. In computational experiments for the 0—1 knapsack problem, the multidimensional knapsack problem and the set covering problem we compare different compilation methods. It turns out that a new strategy based on the local search framework consistently yields better bounds, in many cases far better bounds, for limited-width DDs than previously published heuristic strategies proposed in the literature.

3 - Hybrid Optimization for Time-dependent Sequencing Problems

Willem-Jan Van Hoeve, Carnegie Mellon University, Tepper School of Business, 5000 Forbes Avenue, Pittsburgh, PA, 15213, United States, vanhoeve@andrew.cmu.edu, Joris Kinable, Andre Augusto Cire

We present a hybrid optimization method for sequencing problems in which the setup time between a pair of tasks depends on the relative position of the tasks in the ordering. Our hybrid approach extends a constraint programming model with two relaxations: One discrete relaxation based on multi-valued decision diagrams, and one continuous relaxation based on linear programming. We show how these relaxations can be combined to generate stronger bounds and enhance constraint propagation. Experiments conducted on three variants of the time-dependent traveling salesman problem indicate that our techniques substantially outperform mixed-integer linear programming and constraint programming models.

■ SA12

Student Commons 2504

MINLP

General Session

Chair: Akshay Gupte, Clemson University, Clemson, SC, 29634-0975, United States, agupte@clemson.edu

1 - Volume Computation for Sparse Boolean-Quadric Relaxations

Daphne Skipper, University of Michigan, 1205 Beal Avenue, Ann Arbor, MI, 48109-2117, United States, skipper@usna.edu, Jon Lee

Motivated by understanding the quality of tractable convex relaxations of intractable polytopes, Ko et al. gave a closed-form expression for the volume of a standard relaxation Q(G) of the BQP P(G) of the complete graph. We extend to structured sparse graphs, giving: (i) an efficient algorithm for calculating vol(Q) when G has bounded tree width, (ii) closed-form expressions (and asymptotic behaviors) for vol(Q) for all stars, paths, and cycles, and (iii) a closed-form expression for vol(P) for all cycles. We show that when G is a cycle, the light Q is a very close model for the heavy P.

2 - Cutting Planes for Linear Programs with Complementarity Constraints

Haoran Zhu, University of Wisconsin-Madison, 1513 University Avenue, 3226 Mechanical Engineering Building, Madison, WI, 53706-1572, United States, hzhu94@wisc.edu, Alberto Del Pia, Jeffrey Linderoth

Our focus is on finding the globally optimal solution to a linear program with additional complementarity restrictions on certain pairs of its variables. We study an extended formulation of the problem arising from a special form of the Reformulation-Linearization-Technique. We prove structural properties about this extended formulation and then augment the formulation with cutting planes derived from the Boolean Quadric Polytope in order to improve computational performance of branch-and-bound based methods.

3 - Some Characterizations of Bilinear Functions with Small Extended Formulations

Akshay Gupte, Clemson University, Dept. of Mathematical Sciences, O-321 Martin Hall, Clemson, SC, 29634-0975, United States, agupte@clemson.edu

We consider the problem of characterizing the convex hull of the graph of a bilinear function f(x) over [0,1]^n. The Boolean Quadric Polytope (BQP) is an extended formulation of this convex hull. However, the BQP has a rich combinatorial structure that is not known completely. We propose a systematic study of the properties of f(x) that guarantee that a small subset of some well-known facet-defining inequalities for the BQP is sufficient for an extended formulation. In particular, we identify three classes of bilinear functions, represented in terms of their incidence graphs, for which polynomial-sized relaxations of the BQP project onto the convex hull. Our proof technique uses a measure-theoretic characterization that we simplify from literature for 0/1 polytopes and establish for graphs of nonlinear functions. We also provide a computational study wherein we apply the BQP facets to random graphs and QCQP instances to determine which of the inequalities are strongest in practice.

■ SA13

Student Commons 2600

AMPL

Tutorial Session

Chair: Robert Fourer, AMPL Optimization Inc., 2521 Asbury Ave, Evanston, IL, 60201, United States, 4er@ampl.com

1 - A Guide to Identifying Good Near-Optimal Formulations for Hard Mixed-Integer Programs

Robert Fourer, AMPL Optimization Inc., 2521 Asbury Ave, Evanston, IL, 60201, United States, 4er@ampl.com

When an exact mixed-integer programming formulation resists attempts at solution, sometimes much better results can be achieved by “cheating” a bit on the formulation. Typically, a judicious choice of reformulation, restriction, or decomposition serves to make the problem easier, in a way not guaranteed to preserve the solution’s optimality but highly unlikely to make much of a difference given the model and data of interest. This tutorial illustrates such an approach through a series of case studies. All rely on trial and error, a flexible modeling language, and a good general-purpose solver, and each is seen to be founded on one or two simple ideas that have the potential to be more broadly applied.

Saturday, 10:30AM - 11:20AM

■ Plenary III

Student Commons 2600

Value of Multi-Stage Stochastic Optimization in Power Systems Operations

Plenary Session

1 - Value of Multi-Stage Stochastic Optimization in Power Systems Operations

Shabbir Ahmed, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA, 30332, United States, sahmed@isye.gatech.edu

Day-ahead scheduling of electricity generation or unit commitment is an important and challenging operational activity in power systems. Increasing penetration of renewable technologies in recent years has motivated addressing uncertainty in this already difficult optimization problem. Existing approaches adopt a two-stage decision structure, where the day-ahead commitment is decided before the uncertainty is realized and the power dispatch is adapted to the uncertainty. In the first part of this talk, we present theoretical results on the value of multi-stage or dynamic generation scheduling in risk-neutral and risk-averse settings. The second part of the talk is on algorithmic developments. In particular, we present a Stochastic Dual Dynamic Integer Programming (SDDiP) algorithm for multistage stochastic unit commitment problems.

Saturday, 11:30AM - 12:20PM

■ Plenary IV

Student Commons 2600

Hybrid Optimization Algorithms to Solve Real-world Problems

Plenary Session

1 - Hybrid Optimization Algorithms to Solve Real-world Problems

Karla L. Hoffman, George Mason University, System Eng and Operations Research Dept., 4400 University Drive Mailstop 4a6, Fairfax, VA, 22030, United States, khoffman@gmu.edu

Optimization algorithms are playing an increasingly important role in improving the operations of many standard corporate activities such as supply chain management, real-time scheduling and routing, and the pricing of goods and services through auctions. In this talk, we examine several large-scale real-world problems recently solved using hybrid optimization techniques. First, we describe a high-profile government auction where the Federal Communications Commission (FCC) bought back spectrum from TV stations, packed the remaining broadcasters into a smaller swath of spectrum and sold the acquired spectrum to the wireless industry. Optimization was used before, during and after this auction to assure that multiple governmental goals are met. Our second example considers how the military can use similar optimization strategies to allocate limited spectrum during combat operations when spectrum is scarce and communication vital. In both examples, the underlying structure is that of a graph-coloring problem with multiple side constraints. Our approach is to use a combination of combinatorial optimization, heuristics, decompositions, and constraint programming to create an overall algorithm capable of solving to global or near global optimality problems with millions of variables and hundreds of thousands of constraints. Our third example explores real-time routing and scheduling where one needs near-optimal solutions in less than a second. In this case, we compare and seek to determine the conditions under which the following algorithms provide the best overall performance: enumeration, constraint programming, heuristics, and global optimization techniques.
In each application, we use realistic data sets that we make available to the research community. We conclude with suggestions for other areas that seem ripe for exploitation by similar hybrid optimization approaches.

Saturday, 2:00PM - 3:30PM

■ SB01

North Classroom 1806

First-Order Methods for Large-Scale Constrained Nonlinear Programming II

General Session

Chair: Yangyang Xu, Ressenlaer Polytechnic Institute, Troy, NY, 12180, United States, xuy21@rpi.edu

Co-Chair: Qihang Lin, The University of Iowa, Iowa City, IA, 52245, United States, qihang-lin@uiowa.edu

1 - New Level Set Method for Convex Optimization

Negar Soheili, University of Illinois-Chicago, 601 S Morgan Street, University Hall 2416, Chicago, IL, 60607, United States, nazad@uic.edu

First order methods have emerged as an effective solution strategy for tackling large-scale convex optimization problems. The success of these methods has been greatest for such problems with simple feasible sets. We propose a novel level set method that extends the applicability of these methods to convex optimization problems with potentially complicated constraint sets.

2 - Convergence Rates for Stochastic Subgradient Methods without Lipschitz Continuity or Convexity

Benjamin Grimmer, Cornell University, 114 Parker St, Ithaca, NY, 14850, United States, bdg79@cornell.edu

We extend the classic convergence rate theory for subgradient methods to apply to non-Lipschitz functions. For the deterministic projected subgradient method, we present a global O(1/√ T) convergence rate for any convex function with at most exponential growth. Further, we show a O(1/√ T) convergence rate for the stochastic projected subgradient method on convex functions with at most quadratic growth, which improves to O(1/T) under strong convexity. Finally, we introduce a nonconvex variation of the stochastic subgradient method where convergence to a stationary point can be guaranteed at the same rate as the stochastic gradient method for smooth nonconvex problems.

3 - Iteration Complexity of First Order Augmented Lagrangian Methods for Convex Conic Programming

Zirui Zhou, Simon Frazier University, Surrey, BC, Canada, ziruiz@sfu.ca, Zhaosong Lu

In this paper we consider a class of convex conic programming problem. We propose an inexact augmented Lagrangian (I-AL) method for this problem, where the subproblems are solved approximately by a variant of Nesterov's optimal first-order method. We show that the number of overall first-order iterations of the proposed I-AL method for computing an ϵ-approximate Karush-Kuhn-Tucker (KKT) solution is at most $\mathcal{O}(\epsilon^{-\frac{7}{4}})$. We also propose a modified I-AL method and show that it has an improved iteration-complexity $\mathcal{O}(\epsilon^{-1}\log(\epsilon^{-1}))$. Our analysis of iteration-complexities is based on the results that the proposed two I-AL methods can be regarded as inexact proximal point algorithm (PPA) applied to two different maximally monotone operators respectively. As a byproduct, we establish some results on the iteration-complexity of inexact PPA, which could be of independent interest. Compared to existing works on I-AL methods, our algorithms can be used practically and apply to a broader class of problems, and our analysis provides the sharpest complexity bounds.

4 - Linearized Augmented Lagrangian Method for Convex Programs with Functional Constraints

Yangyang Xu, Assistant Professor, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 12180, United States, xuy21@rpi.edu

First-order methods have been popularly used for solving large-scale problems. However, many existing works only consider unconstrained problems or those with simple constraint. In this talk, I will introduce two first-order methods for constrained convex programs, for which the constraint set is represented by affine equations and smooth nonlinear inequalities. Both methods are based on the classic augmented Lagrangian function. They update the multipliers in the same way as the augmented Lagrangian method (ALM) but employ different primal variable updates. The first method, at each iteration, performs a single proximal gradient step to the primal variable, and the second method is a block update version of the first one. Global iterate convergence as well as global sublinear and local linear convergence will be shown. Numerical experiments are carried out on the basis pursuit denoising and a convex quadratically constrained quadratic program to show the empirical performance of the proposed methods. Their numerical behaviors closely match the established theoretical results.

■ SB02

North Classroom 1602

Data-Driven Stochastic Program: Theory and Applications

General Session

Chair: Weijun Xie, Virginia Tech, Blacksburg, VA, 24060, United States, wxie@vt.edu

1 - A Copositive Approach for Two-stage Adjustable Robust Optimization with Uncertain Right-hand Sides

Guanglin Xu, University oif Minnesota, 207 Church Street SE, Lind Hall 432, Minneapolis, MN, 55108, United States, guanglin-xu@uiowa.edu, Samuel Burer

We study two-stage adjustable robust linear programming in which the right-hand sides are uncertain and belong to a convex, compact uncertainty set. This problem is NP-hard, and the affine policy is a popular, tractable approximation. We prove that under standard and simple conditions, the two-stage problem can be reformu- lated as a copositive optimization problem, which in turn leads to a class of tractable, semidefinite-based approximations that are at least as strong as the affine policy. We investigate several examples from the literature demonstrating that our tractable ap- proximations significantly improve the affine policy. In particular, our approach solves exactly in polynomial time a class of instances of increasing size for which the affine policy admits an arbitrarily large gap.

2 - Distributionally Robust Expectation using Dominance Information

Ruiwei Jiang, University of Michigan, 1205 Beal Ave., Ann Arbor, MI, 48109, United States, ruiwei@umich.edu, Yuanyuan Guo

This talk discusses the expectation of a random function when the distributional information of the uncertain parameters consists of moment (e.g., mean, covariance, support) and probabilistic dominance information. We find that the expectation in this setting can be bounded using conic programming. Finally, we demonstrate the theoretical results via case studies on appointment scheduling.

3 - Effective Scenarios in Multistage Distributionally Robust Stochastic Programs with Total Variation Distance

Hamed Rahimian, The Ohio State University, 1971 Neil Avenue, 210 Baker Systems Building, Columbus, OH, 43210, United States, rahimian.1@osu.edu, Guzin Bayraksan, Tito Homem-de-Mello

Traditional multistage stochastic programs assume perfect information about the underlying probability distribution of uncertainty is available and optimize the expected cost. However, in practice, decision makers have imperfect information about the underlying distribution of uncertainty. An alternative modeling approach to handle uncertainty in these cases is to use distributionally robust stochastic programming (DRSP), which assumes the true probability distribution lies in an ambiguity set of distributions and hedges against the worst-case probability distribution in the ambiguity set. In this work, we study the multistage DRSP with a finite number of scenarios. We formulate this problem in a nested form and use total variation distance to construct the stagewise ambiguity sets of conditional probability distributions. We show the resulting problem is equivalent to a multistage risk-averse optimization with nested coherent risk measures—-formed by a convex combination of the worst-case and conditional value-at-risk of the cost function. We propose a decomposition-based algorithm to obtain the optimal policy and the worst-case probability distribution. In this work, we also investigate the question of which scenarios have “effect” on the optimal value of the multistage DRSP. To identify these effective scenarios, we conduct the perturbation analysis of the problem with respect to a collection of scenarios being excluded and propose easy-to-check sufficient conditions for them. We explore the effectiveness for a scenario path as well as for a scenario conditional on the history of the stochastic process. Computational experiments illustrate the results on finance and energy problems.

■ SB03

North Classroom 1202

Robust Optimization: Theory and Applications

General Session

Chair: Vineet Goyal, Columbia University, New York, NY, 10027, United States, vgoyal@ieor.columbia.edu

1 - Benders Decomposition for Adjustable Robust Optimization Problems

Dimitri Papadimitriou, Nokia Bell Labs, Rue du Charme 24, Brussels, 1190, Belgium, dimitri.papadimitriou@nokia-bell-labs.com

In this paper, we show that in natural situations, when the past is known but present and future are unknown, the Affinely Adjustable Robust Counterpart (AARC) approximation of a very general class of uncertain mixed-integer linear optimization problems is subject to a fundamental tradeoff between model complexity, efficiency of the computational method and solution quality. In the context of two-stage robust programs (with continuous second-stage variables), such problems translate situations where allocation (online/dynamic) decisions taken to satisfy uncertain demands depend on anterior production decisions derived themselves from past observations. We analyze the properties of the resulting Affinely Adjustable Robust Counterpart (AARC) formulation when the decision-making policies are limited to (piecewise) affine rules, i.e., the continuous adjustable variables are approximated by affine function of the uncertain data. Then, we propose a formulation of the uncertain mixed-integer linear program that exploits its decomposable structure into first and second stage decision problems. Next, we propose an exact algorithm for the solving of such problems that relies on the Benders decomposition method. This method relies on dynamic cutting-plane generation. More precisely, under the relatively complete recourse assumption, it performs by iteratively generating so-called optimality cuts/cutting planes (obtained from the dual subproblem), adding them to the master problem and solving the resulting master problem such that its lower and upper bounds converge and thus, an optimal solution of the original uncertain problem can be obtained. This paper also investigates possible strategies to reduce the convergence time of the Benders decomposition algorithm to the optimal solution by maintaining balance between the number of iterations and the number as well as the type of cuts produced at each iteration. Finally, we provide a first characterization of the fundamental tradeoff between decomposability and robustification for this class of problems.

2 - Robust Markov Decision Processes under Non-rectangular Uncertainty

Julien Grand Clement, Columbia University, 195 Claremont Avenue, New York, NY, 10027, United States, jg3728@columbia.edu, Vineet Goyal

Markov decision processes are a common approach for sequential decision problem with a stochastic dynamic. However, in most real world problems, there is uncertainty in the model parameters. We consider a robust approach where the uncertainty in probability transitions is modeled as an adversarial selection from an uncertainty set. Most prior work considers the case where uncertainty between different states is uncoupled, potentially leading to highly conservative solutions. On the other hand, the case of general uncertainty sets is known to be intractable. We consider the model of rank-r uncertainty sets where probability transition matrices depends on only a few underlying vectors / factor models. We show that our model leads to a tractable approach for the Policy Evaluation and the Policy Improvement problem using a rectangularity assumption and thus overcome the conservativeness of using uncoupled uncertainty sets. We also describe an FPTAS for the Policy Evaluation problem if one does relax the assumption of rectangularity.

3 - On Optimal Approximations via Affine Policies for Two-stage Adjustable Robust Optimization under Budget of Uncertainty Sets

Vineet Goyal, Columbia University, 304 S.W. Mudd Building, 500w 120th Street, New York, NY, 10027, United States, vgoyal@ieor.columbia.edu, Omar El Housni

The performance of affine policies for two-stage adjustable robust optimization problem under a budget of uncertainty set. This important class of uncertainty sets provide the flexibility to adjust the level of conservatism in terms of probabilistic bounds on constraint violations. The two-stage adjustable robust optimization problem is hard to approximate within a factor better than $\Omega(\log n)$ even for budget of uncertainty sets where n is the number of decision variables. We show affine policies provide the best possible approximation of $O(\log n)$ for budget of uncertainty sets. We discuss performance of affine when the uncertainty set is given by intersection of budget constraints.

■ SB04

North Classroom 1204

Robust and Discrete Optimization

Contributed Session

Chair: Adolfo Raphael Escobedo, Arizona State University, P.O. Box 878809, Tempe, AZ, 85287-8809, United States, adRes@asu.edu

1 - Robust Binary Linear Programming Under Implementation Uncertainty

Jose Ernesto Ramirez Calderon, Texas A&M University, 1501 Harvey Rd., College Station, TX, 77840, United States, ramirez.jose@tamu.edu

This paper studies robust binary linear programming in the presence of uncertainties that may prevent the implementation of solutions exactly as prescribed. A formal model of this type of uncertainty, termed implementation uncertainty, is presented and used to develop a robust binary linear programming formulation under implementation uncertainty. The solution approach involves a mixed-binary linear program reformulation used in conjunction with identified structural properties, allowing the generation of a set of optimal solutions. Three methods to control the level of conservatism are developed including a cardinality-constrained reformulation of the original problem. Experimental results compare the deterministic solutions and robust solutions in terms of the feasibility protection percentage and optimality loss.

2 - An Adiabatic Quantum Optimization Approach for the Mixed Integer Nonlinear Programming Problem

Maxwell Henderson, QxBranch LLC, Washington, DC, 20036, United States, max.henderson@qxbranch.com, John Kelly

We present a method of using adiabatic quantum optimization (AQO) to solve a mixed integer nonlinear programming (MINLP) problem instance. The MINLP problem is a general form of a set of NP-hard optimization problems that are critical to many business applications. It requires optimizing a set of discrete and continuous variables with nonlinear and potentially nonconvex constraints. Obtaining an exact, optimal solution for MINLP problem instances of non-trivial size using classical computation methods is currently intractable. Current leading algorithms leverage heuristic and divide-and-conquer methods to determine approximate solutions. Creating more accurate and efficient algorithms is an active area of research. Quantum computing (QC) has several theoretical benefits compared to classical computing, through which QC algorithms could obtain MINLP solutions that are superior to current algorithms. AQO is a particular form of QC that could offer more near-term benefits compared to other forms of QC, as hardware development is in a more mature state and devices are currently commercially available from D-Wave Systems Inc. It is also designed for optimization problems: it uses an effect called quantum tunneling to explore all lowest points of an energy landscape where classical approaches could become stuck in local minima. Our work used a novel algorithm formulated for AQO to solve a special type of MINLP problem. The research focused on determining: 1) if the problem is possible to solve using AQO, 2) if it can be solved by current hardware, 3) what the currently achievable performance is, 4) what the performance will be on projected future hardware, and 5) when AQO is likely to provide a benefit over classical computing methods. Two different methods, integer range and 1-hot encoding, were investigated for transforming the MINLP problem instance constraints into a mathematical structure that can be embedded directly onto the current D-Wave architecture. For testing and validation a D-Wave 2X device was used, as well as QxBranch's QxLib software library.

3 - An Improved Branch and Bound Algorithm for Robust Ranking Aggregation

Adolfo Raphael Escobedo, Assistant Professor, Arizona State University, P.O. Box 878809, Tempe, AZ, 85287-8809, United States, adRes@asu.edu, Yeawon Yoo

We introduce a new fundamental social-choice related property for distributed decision-making, called the relativized extended Condorcet criterion, which can be regarded as a natural extension of the well known extended Condorcet criterion. We prove that this property is satisfied by the scaled Kendall tau-extended correlation coefficient and the normalized projected Kemeny Snell distance, but not by any other measures. Unlike its parent property, the relativized extended Condorcet criterion is adequate for both complete and incomplete rankings. This new property also provides computational advantages. Namely, it allows us to simplify the solution process for certain types of instances of the NP-hard consensus ranking problem via a combinatorial branch and bound algorithm. We also derive additional enhancements for this algorithm including branching strategies, initial-solution generation procedures, and an upper-bound refinement scheme. Together, these enhancements allow the algorithm to solve instances with large numbers of objects.

■ SB05

North Classroom 2001

Joint Session: OPT Uncertainty/Network, Shared Mobility System Design and Optimization

General Session

Chair: Siqian Shen, University of Michigan, Ann Arbor, MI, 48109, United States, siqian@umich.edu

1 - Models and Algorithms to Balance Bike-sharing Systems

Daniel Freund, Cornell University, Ithaca, NY, United States, danielrfreund1990@gmail.com, Shane Henderson, David B. Shmoys

The collaboration between Cornell's School of Operations Research and Information Engineering (ORIE) and the bike-sharing company Motivate gives us access to a wealth of interesting operational questions with the potential of real-world impact. For Motivate, operator of the systems in NYC, Chicago, DC, and Boston, among others, this collaboration provides access to a range of analytical tools. In this talk, I will present one of the models we use to support Motivate's operations. In particular, I will describe a simple bike-sharing inventory model defined by Raviv and Kolka (2013) and present various applications we developed for it in rebalancing, incentives, and system design.

2 - Designing and Optimizing an Integrated Car-and-ride Sharing System for Mobilizing Underserved Populations

Miao Yu, University of Michigan, 1891 IOE, 1205 Beal Avenue, Ann Arbor, MI, 48109, United States, miaoyu@umich.edu, Siqian Shen

Carsharing and ridesharing are experiencing a fast growth and have shown their benefits in long-term social and economic changes. However, both types of service have their restrictions in coverage areas. In this paper, we consider an integrated car-and-ride sharing system and optimize its operations to improve the mobility of underserved communities with transportation scarcity. We consider two types of demands: Type 1 users who rent shared cars and Type 2 users who need rides served by Type 1 users. We propose a two-phase model to maximize served demand: In Phase I, we model car movements and match supply and demand on a spatial-temporal network; in Phase II, we find optimal matching and scheduling for matched Type 1 and Type 2 users under stochastic driving time via solving a two-stage stochastic mixed-integer programming model. We maximize the total profit with penalties on users' waiting time and system overtime, and develop a decomposition algorithm based on finite samples of driving time. We conduct computational studies on diverse instances and data related to underserved communities in Washtenaw County, Michigan. Both in-sample and out-of-sample test results demonstrate high demand fulfillment rates and effective matching and scheduling with low risk of waiting and overtime, as well as better results after allowing vehicle relocation.

3 - Online Vehicle Routing, the Edge of Optimization in Large-scale Applications

Sebastien Martin, Massachusetts Institute of Technology, 238 Prospect St, Cambridge, MA, 02139, United States, semartin@mit.edu, Patrick Jaillet, Dimitris Bertsimas

With the emergence of ride-sharing companies that offer transportation on demand at a large scale and the increasing availability of corresponding demand datasets, new challenges arise to develop routing optimization algorithms that can solve massive problems in real time. In this talk, we discuss an optimization framework, coupled with a novel backbone algorithm, that allows us to dispatch in thousands of taxis serving more than 25,000 customers per hour using the New York City demand data and routing network. We provide evidence from historical simulations to show that our algorithms improve upon the performance of existing heuristics in such real-world settings.

■ SB06

North Classroom 2002

Derivative-free and Simulation-based Optimization

General Session

Chair: Faruque Hasan, Texas A&M University, College Station, TX, 77843, United States, hasan@tamu.edu

1 - Optimization using Surrogate Models - Which Surrogate Model?

Selen Cremaschi, B. Redd Associate Professor, Auburn University, 212 Ross Hall, Auburn, AL, 36849, United States, selen-cremaschi@auburn.edu, Sarah Davis, Mario Eden

Surrogate models - simple functional approximations of complex models - can facilitate optimization of complicated systems by greatly reducing computational expense. The quality of surrogate models depends on two main characteristics: (1) the surrogate model form, and (2) the size and location of the samples that are used to determine the parameters of the selected model form. There are a number of machine learning techniques that can be used to construct surrogate models, e.g., Extreme Learning Machines (ELM), Artificial Neural Networks (ANN), and Automated Learning of Algebraic Models using Optimization (ALAMO), but little work has been done to systematically compare their abilities to learn the responses and the extremums of complicated models with different characteristics for optimization purposes. This talk aims to fill in this gap by comparing different surrogate-model construction approaches using computational experiments. Eight surrogate-model construction approaches are evaluated, and they are Artificial Neural Networks (ANN), Automated Learning of Algebraic Models using Optimization (ALAMO), Radial Basis Networks (RBN), Extreme Learning Machines (ELM), Gaussian Progress Regression (GPR), Random Forests (RF), Support Vector Regression (SVR), and Multivariate Adaptive Regression Splines (MARS). Each approach is used to construct surrogate models for solving thirty-four optimization test problems with various function shapes and number of decision variables. These test problems can be found in Virtual Library of Simulation Experiments (https://www.sfu.ca/~ssurjano/optimization.html). The data set used for training the surrogate models were generated using Sobol sequences. The performance is measured using two metrics: (1) the distance between the global minimum of the optimization test problem and the one obtained using the surrogate model; and (2) the mean absolute percentage error of the surrogate model predictions for a set of 100000 data points generated using Sobol sequences.

2 - Integrating a Concentrating Solar Power Plant Availability Simulation with an Optimization Dispatch Model

Jesse Wales, Colorado School of Mines, Golden, CO, United States, jwales@mines.edu, Alexander Zolan, William Hamilton, Michael J. Wagner, Alexandra Newman, Robert Braun

Concentrating Solar Power (CSP) is an emerging large-scale renewable technology consisting of a field of mirrors that directs sunlight to a small area where it heats molten salt. The molten salt is held in a storage device where it can be converted to power via a steam-Rankine cycle using heat-exchangers and turbines. We describe a maintenance, failure, and repair simulation that interfaces with an optimization dispatch model for a CSP production plant. The optimization model determines an hourly energy dispatch strategy on a rolling horizon to maximize profit. The simulation generates planned and unplanned down-times based on dispatch decisions from the optimization model; these down-times are incorporated into constraints in the optimization model. This allows wear-and-tear impacts from dispatch decisions, such as the number of start-ups of the CSP plant, to be realized at the component level. The results of running the availability simulation integrated with the CSP plant optimization dispatch model are presented.

3 - Branch-and-bound Optimization using Best Surrogate Approximations

Fani Boukouvala, Georgia Institute of Technology, Atlanta, GA, 30332, United States, fani.boukouvala@chbe.gatech.edu, Jianyuan Zhai, Sun Hye Kim, Alexandra Brown

Decision-making for an increasing number of systems today depends on data, generated either by high-fidelity simulations or by advanced sensors embedded within complex engineering systems. Despite the high potential of applicability of data-driven optimization, currently there is a lack of efficient and scalable methods, which can provide high quality solutions for a general class of data-dependent problems. In equation-based optimization, the existence of algebraic relationships allow for global convergence and guarantee of optimality, through the calculation of derivatives and exploration of the model structure. However, when equations are not available, optimality has only been guaranteed at an asymptotic rate of sampling, which is unrealistic for real applications with many variables, constraints and high sampling cost. A very common approach in data-driven optimization is fitting of surrogate-functions using collected data, which have been found to speed up the search for better solutions. When optimized, these surrogates can lead to new and promising sampling locations, while new data points are used to update the fitted surrogate functions iteratively (adaptive sampling). In this work, we present results of a new data-driven optimization approach, which uses statistical information from fitted surrogate models to derive over and under estimators that are used within a custom-based branch-and-bound framework. We compare this new idea with existing data-driven optimization methods and provide results on the computational efficiency, sampling requirements and convergence of the new methodology for a general class of benchmark problems.

4 - New Projection-based Derivative-free Optimization

Faruque Hasan, Texas A&M University, 3122 TAMU, College Station, TX, 77843, United States, hasan@tamu.edu

We propose a new derivative-free method to solve multi-dimensional black-box problems based on a projection of the original function onto a univariate space. This space is defined using a linear combination of the decision variables. Projecting the objective function values onto this space leads to point-to-set map. The lower envelope of this map contains the global minima of the original function. Sensitivity theorem can be employed to predict the values on the lower envelope of this map and a trust-region based algorithm can be used to correct the predicted point. Once the lower envelope of this map is identified, a single dimensional search can be used to find its minima. The resulting solution is also the minima of the original multi-dimensional problem. Based on these key ideas, we propose a predictor-corrector algorithm for bound-constrained derivative-free optimization. The proposed algorithm is applied to a large suite of test problems comprising of sets of convex smooth, convex nonsmooth and nonconvex smooth black-box problems, and performance is compared to those of other competitive model-based solvers. The algorithm solves more number of nonconvex smooth and convex nonsmooth problems compared to established derivative-free solvers.

■ SB07

North Classroom 3004

Production Planning

Contributed Session

Chair: Joshua Betz, Colorado School of Mines, Golden, CO, United States, jkbetz@mymail.mines.edu

1 - Guaranteed Delivery Times for Scheduling Jobs under Uncertainty

Arie Koster, RWTH Aachen University, Pontdriesch 10-12, Aachen, 52056, Germany, koster@math2.rwth-aachen.de, Cole Smith

From a mathematical optimization point of view, many (production) planning problems can be modelled as scheduling jobs on a single machine. Here, the processing times are often uncertain and a schedule have to be prepared without complete knowledge about them. Recently, there have been an increasing interest in applying robust optimization to these and related problems. For this, an uncertainty set have to be defined containing all considered processing time vectors. Goal is to determine a processing sequence of the jobs that yields good performance in the worst case. The completion times of the jobs can only determined after the actual processing times are revealed.
In this talk, we follow a customer oriented approach, where we have to determine guarenteed delivery times, i.e., the jobs have to be completed by their delivery time regardless the actual processing times of the jobs. For the objective to minimize the weighted sum of delivery times, we present first formulations for this problem setting as well as a combinatorial algorithm, extending the well-known Smith's rule for the sum of weighted completion times.

2 - Coordinating Decentralized Negotiations Over Factory Capacity in the Semiconductor Industry

Ankit Bansal, North Carolina State University, 2713 Brigadoon Drive, Apt 21, Raleigh, NC, 27606, United States, abansal3@ncsu.edu, Reha Uzsoy, Karl Kempf

We model the negotiations between a product development organization and a production organization for access to manufacturing capacity for product development activities in the semiconductor industry. We develop a negotiation framework based on Lagrangian decomposition that maximizes overall firm contribution subject to the resource constraints of both organizations. The approach aims to achieve coordinated decisions between the two organizations, and provides a benchmark for alternative models of negotiations.

3 - Scheduling Optimization for a Continuous Steel Caster

Joshua Betz, Colorado School of Mines, Golden, CO, 80401, United States, jkbetz@mymail.mines.edu

In continuous steel casting operations, molten steel is batched into heats inside a ladle that is cast into slabs, which are then rolled into coils. We present a mixed integer program to produce a daily casting schedule that is solved using state-of-the-art software. This model minimizes penalties received by violating plant best practices while strictly adhering to safety and logical constraints. A heuristic produces an initial feasible solution to expedite the generation of near-optimal schedules.

■ SB08

North Classroom 3210

Augmented Lagrangian, Operator Splitting, and Alternating Direction Methods

General Session

Chair: Jonathan Eckstein, Rutgers University, 100 Rockafeller Road, Room 5145, Piscataway, NJ, 08854, United States, jeckstei@rci.rutgers.edu

1 - On the Convergence and Complexity of Nonconvex ADMM

Shiqian Ma, University of California-Davis, Department of Mathematics, Davis, CA, United States, sqma@math.ucdavis.edu

The alternating direction method of multipliers (ADMM) has been successfully used in solving problems arising from different fields such as machine learning, image processing, statistics and so on. However, most existing works on analyzing the convergence and complexity of ADMM are for convex problems. In this talk, we discuss several recent results on convergence behavior of ADMM for solving nonconvex problems. We consider two nonconvex models. The first model allows the objective function to be nonconvex and nonsmooth, but the constraints are convex. The second model allows the constraints to be Riemannian manifolds. For both models, we propose ADMM variants for solving them and analyze their iteration complexities for obtaining an ϵ-stationary solution. Numerical results on tensor robust PCA, maximum bisection problem and community detection problem are reported to demonstrate the efficiency of the proposed methods.

2 - Projective Splitting with Forward Steps: A Flexible, Distributed, and Asynchronous Operator-splitting Scheme

Patrick R. Johnstone, PhD, Rutgers University, Piscataway, NJ, 08854, United States, patrick.r.johnstone@gmail.com, Jonathan Eckstein

Within the block-iterative projective operator splitting framework recently proposed by Combettes and Eckstein, we show how to replace the original backward step calculation with one based on two forward steps for operators that are Lipschitz continuous. The resulting algorithms have the same kind of coordination procedure and can be implemented in the same asynchronous manner, but may perform backward steps on some operators and forward steps on others. The method has a stepsize constraint which resembles those for other splitting methods with forward steps. When Lipschitz constants are unknown, a simple backtracking linesearch procedure can be used. We discuss overall convergence, convergence rates, various applications, and some computational tests on data fitting problems.

3 - OPAL: Building Parallel Augmented Lagrangian Solvers without Decomposition

Jonathan Eckstein, Rutgers University, 100 Rockafeller Road, Room 5145, Piscataway, NJ, 08854, United States, jeckstei@rci.rutgers.edu, Gyorgy Matyasfalvi

This talk presents OPAL ("Object Parallel Augmented Lagrangian"), a software framework for building application-specific optimization solvers that can take advantage of parallelism but use a form of the classical augmented Lagrangian method, as opposed to a decomposition variant like the ADMM. OPAL implements an algorithm closely resembling the Birgin-Martinez "ALGENCAN" augmented Lagrangian solver, which uses a form truncated Newton method to solve the augmented Lagrangian subproblems. But rather than being a monolithic solver, OPAL is a template that supports flexible parallelism in whatever form the application warrants, requiring only efficient parallel implementation of the underlying matrix-vector multiplication operations. We describe applying OPAL to large-scale multistage continuous stochastic programming problems using a data layout resembling that of progressive hedging methods. The necessary interprocessor communication can be reduced to operations that take time logarithmic in the number of processors, including an innovative use of parallel prefix "scan" operations.

■ SB09

Student Commons 1500

Nonconvex Optimization

General Session

Chair: Alejandro Ribeiro, University of Pennsylvania, 101 Crosshill Rd., Wynnewood, PA, 19096, United States, aribeiro@seas.upenn.edu

Co-Chair: Santiago Paternain, University of Pennsylvania, 315 S 45th Street, Philadelphia, PA, 19104, United States, spater@seas.upenn.edu

Co-Chair: Aryan Mokhtari, University of Pennsylvania, Philadelphia, PA, 19104, United States, aryanm@seas.upenn.edu

1 - Geometry of Non-convex Landscapes: Deep Learning, Matrix Completion, and Saddle-points

Jason Lee, University of Southern California, San Marino, CA, United States, jasonlee@marshall.usc.edu

We show that saddlepoints are easy to avoid for even Gradient Descent — arguably the simplest optimization procedure. We prove that with probability 1, randomly initialized Gradient Descent converges to a local minimizer. The same result holds for a large class of optimization algorithms including proximal point, mirror descent, and coordinate descent.
Next, we study the problems of learning a two-layer ReLU network and the matrix completion problem. Despite the non-convexity of both problems, we prove that every local minimizer is a global minimizer. By combining with the previous algorithmic result on gradient descent, this shows that simple gradient-based methods can find the global optimum of these non-convex problems.Step Complete Status: Complete

2 - Learning to Optimization: Training Deep Neural Networks for Signal Processing

Mingyi Hong, University of Minnesota, 200 Union Street SE 4-174 Keller Hall, Minneapolis, MN, 55455, United States, mhong@umn.edu

In this work, we propose a new learning-based approach for wireless resource management. The key idea is to treat the input and output of a resource allocation algorithm as an unknown non-linear mapping and to use a deep neural network (DNN) to approximate it. If the non-linear mapping can be learned accurately and effectively by a DNN of moderate size, then such DNN can be used for resource allocation in almost real time, since passing the input through a DNN to get the output only requires a small number of simple operations. In this work, we first discuss a few theoretical issue related to this approach. We characterize a class of 'learnable algorithms' and then design DNNs to approximate some algorithms of interest in wireless communications. Further, we rigorously characterize how the approximation error scale as a function of the size of DNN. Finally, we use extensive numerical simulations to demonstrate the superior ability of DNNs for approximating a state-of-the-art algorithm that is designed for power allocation in wireless transmit signal design, while giving orders of magnitude speedup in computational time.

3 - Complexity Analysis of Second-order Line-search Algorithms for Smooth Nonconvex Optimization

Clement W. Royer, University of Wisconsin-Madison, 330 N Orchard Street, Madison, WI, 53715, United States, croyer2@wisc.edu, Stephen J Wright

There has recently been a growth of interest in algorithms dedicated to solving unconstrained smooth nonconvex optimization problems, partly due to the outbreak of such instances in neural networks and robust statistics. For these schemes, the derivation of complexity guarantees has become an increasingly popular area of research: motivated by promising result for cubic regularization methods, one line of work has focused on deriving iteration complexity bounds for Newton-based algorithms. A number of recently proposed algorithms, based on first-order methodology but with second-order guarantees, have also been endowed with precise guarantees regarding their computational cost. Among those classes of methods, it can be difficult to build a comparison based on the sole complexity bounds, or to determine what algorithmic features are instrumental in guaranteeing an optimal complexity order. In this talk, we describe a generic line-search framework that explicitly deals with nonconvexity through the use of second-order information. We first describe the algorithm in its basic version, and show that its complexity bounds are of optimal order among a large class of Newton-type methods: the proof simply relies upon the standard principle of backtracking line search, together with a suitable choice of a search direction. We then study inexact variants of the method based on the conjugate gradient and Lanczos algorithms, two popular and efficient iterative linear algebra tools. For the resulting schemes, only Hessian-vector products are required, while complexity guarantees are preserved. Our generic analysis allows us to position our framework within the landscape of optimization methods with second-order complexity guarantees, in order to compare them on both a theoretical and a practical level. Our algorithm is easily implementable, and amenable to several variations that do not jeopardize the theoretical results: we will provide numerical illustrations on how those modifications impact the practical behavior of our method when applied to nonconvex problems.

4 - A Second-order Method for Non-convex Optimization

Santiago Paternain, University of Pennsylvania, 315 S 45th Street, Apt 1F, Philadelphia, PA, 19104, United States, spater@seas.upenn.edu, Aryan Mokhtari, Alejandro Ribeiro

Machine learning problems such as neural network training, tensor decomposition, and matrix factorization, require local minimization of a nonconvex function. This local minimization is challenged by the presence of saddle points, of which there can be many and from which descent methods may take an inordinately large number of iterations to escape. In this talk, we present a second-order method that modifies the update of Newton's method by replacing the negative eigenvalues of the Hessian by their absolute values and uses a truncated version of the resulting matrix to account for the objective function's curvature. The method is shown to escape saddles exponentially with base 1.5 regardless of the condition number of the problem. Adding classical properties of Newton's method, the paper proves convergence to a local minimum with high probability after a number of iterations that is logarithmic in the target accuracy.

■ SB10

Student Commons 1600

Nonlinear Optimization

General Session

Chair: Farzad Yousefian, Oklahoma State Univeristy, Stillwater, OK, 74074, United States, farzad.yousefian@okstate.edu

1 - Conditional Proximal Gradient for Nonsmooth Convex Optimization

Niao He, University of Illinois Urbana-Champaign, 117 Transportation Building, Room 117, Urbana, IL, 61801, United States, niaohe@illinois.edu

We propose a new class of first-order methods to solve high-dimensional non-smooth convex composite problems with mixed regularization penalties, which can be characterized as semi-structured variational inequalities. Typical examples of such problems have an objective that decomposes into a non-smooth convex loss function, a proximal-friendly regularization penalty, and a linear-minimization-friendly penalty. The proposed algorithm, called \emph{Conditional Proximal Gradient}, leverages prox mapping computation and linear minimization, and attains optimality in both worlds: when the loss function admits a smooth saddle point representation, the algorithm achieves the $O(1/\epsilon)$ complexity bound for the number of calls to proximal operators, and $O(1/\epsilon^2)$ for the number of calls to linear minimization oracles. We also present a stochastic approximation extension and a randomized block-decomposition extension of the algorithm along with their theoretical complexity bounds for stochastic optimization problems and problems with finite sum structure, respectively. We present promising experimental results demonstrating the interest of the approach in comparison to competing methods in particular based on smoothing techniques.

2 - Optimal First Order Method for Structured Convex Optimization

Yuyuan Ouyang, Clemson University, O-207 Martin Hall, Clemson, SC, 29634, United States, yuyuano@clemson.edu

We consider a class of convex optimization with certain structure. In particular, the structure of the problem may involve the smoothness of certain loss functions, saddle-point formulation, and linear equality constraints. In order to design efficient first-order methods, it is important to study the impact of the specific structure on the complexity of the problem. We will study some first-order methods that explore the structures of large optimization problems and their respective iteration complexity analysis.

3 - An Iterative Regularized Mirror Descent Method for Stochastic Ill-posed Optimization Problems

Mostafa Amini, Oklahoma State University, Stillwater, OK, United States, moamini@okstate.edu, Farzad Yousefian

A wide range of applications arising in machine learning and signal processing are formulated as an ill-posed optimization problem. To address these problems, we consider a bilevel model, where the goal is to find among optimal solutions of the original problem, a solution that achieves the minimum value of a regularizer. We assume that the objective function of both lower and upper level problems are nonsmooth convex functions, and each one is given as an expected value of a stochastic function. To solve this bilevel problem, we develop an iterative regularized stochastic mirror descent method. At each iteration, the stepsize and regularization parameter are updated. Under suitable assumptions, we establish the convergence of the iterate generated by the algorithm to the optimal solution of the bilevel problem in both an almost sure and a mean sense. Moreover, we derive the iteration complexity of the method in solving the lower level problem and show that it is nearly optimal.

4 - Stochastic Quasi Newton Methods for Nonstrongly Convex Optimization

Farzad Yousefian, Oklahoma State University, Stillwater, OK, United States, farzad.usefian@gmail.com, Angelia Nedich, Uday Shanbhag

Motivated by applications arising from large scale optimization and machine learning, we consider stochastic quasi-Newton (SQN) methods for solving unconstrained convex optimization problems. Much of the convergence analysis of SQN methods, in both full and limited-memory regimes, requires the objective function to be strongly convex. However, this assumption is fairly restrictive and does not hold in many applications. To the best of our knowledge, no rate statements currently exist for SQN methods in the absence of such an assumption. Also, there is no rate statement available for stochastic gradient methods when applied to non-strongly convex problems with unbounded gradients. Motivated by these gaps, we consider optimization problems with non-strongly convex objectives with Lipschitz but possibly unbounded gradients. To address large scale stochastic optimization problems, we develop a regularized stochastic limited-memory BFGS algorithm, where the stepsize, regularization parameter, and the Hessian inverse approximate matrix are updated iteratively. We establish convergence of the iterates to an optimal solution of the original problem both in an almost-sure sense and a mean sense. The convergence rate is derived in terms of the objective function. We present our numerical experiments performed on a large scale text classification problem.

■ SB11

Student Commons 2500

Theoretical and Computational Advances in IPCO

General Session

Chair: Manish Bansal, Virginia Tech., Blacksburg, VA, 24060, United States, bansal@vt.edu

1 - Adaptive Submodular Routing

Fatemeh Navidi, University of Michigan, navidi@umich.edu, Viswanath Nagarajan, Prabhanjan Kambadur

We study a general stochastic routing problem where we are given a metric (V,d) and an algorithm needs to adaptively select a path so as to “cover” a random scenario (drawn from a known distribution) at minimum expected length. The coverage of each scenario is captured by an individual submodular function, where the scenario is said to be covered when its function value goes above some threshold. We obtain a polylogarithmic factor approximation algorithm for this adaptive ranking problem, which almost matches the best-known result. This problem is a significant generalization of Adaptive Traveling Salesman and Travailing Repairman problem. It has many applications in robotics, data collection and disaster management.

2 - Two-dimensional Maximum P-coverage Problem with Partial Coverage

Manish Bansal, Assistant Professor, Virginia Tech, 227 Durham Hall, 1145 Perry Street, Blacksburg, VA, 24060, United States, bansal@vt.edu

We introduce a new generalization of the classical maximum p-coverage problem (MCP) in which p geometric objects of known dimensions are to be located such that their union covers maximum weight distributed on a two-dimensional plane using another set of geometric objects such as circles, rectangles, polygons, etc. We allow partial coverage in the foregoing generalization and denote this problem by MCP-PC. Using greedy approach, we present an approximation algorithm to solve the MCP-PC and showcase that this algorithm has approximation ratio of 1 – 1/e where e is the base of natural logarithm. We then provide theoretical properties for MCP-PC with rectangular objects and an exact algorithm to solve it along with our computational results.

3 - Decentralized Online Integer Programming

Ezgi Karabulut, Rensselaer Polytechnic Institute, Troy, NY, United States, ezgi.karabulut@gatech.edu, Shabbir Ahmed, George L. Nemhauser

We consider a set of agents that need to coordinate their actions to optimize the sum of their objectives while satisfying a common resource constraint. The objective functions of the players are unknown to them a priori and are revealed in an online manner. The resulting problem is an online optimization problem to optimally allocate the resource among the players prior to observing the item values. We show that for any deterministic online algorithm for this problem, there exists a lower bound of $\Omega(T)$ on regret. Furthermore, when the players' integer programs satisfy a special concavity condition, we propose a decentralized randomized online algorithm that guarantees an upper bound of $O(\sqrt T)$ on the expected regret.

4 - On Solving Single-module and Multi-module Capacitated Lot-sizing Problems without Backlogging

Manish Bansal, Assistant Professor, Virginia Tech, 227 Durham Hall, 1145 Perry Street, Blacksburg, VA, 24060, United States, bansal@vt.edu

We provide sufficient conditions under which the (k,l,S,I) inequalities of Pochet and Wolsey (1993), the mixed (k,l,S,I) inequalities, derived using mixing procedure of Gunluk and Pochet (2001), and the paired (k,l,S,I) inequalities, derived using sequential pairing procedure of Guan et al. (2007), are facet-defining for the single-module capacitated lot-sizing problem without backlogging. We also provide conditions under which the inequalities derived using the sequential pairing and the n-mixing procedure of Sanjeevi and Kianfar (2012) are facet-defining for the multi-module capacitated lot-sizing problem without backlogging. Then we present our recent algorithmic advances for solving the multi-module capacitated lot-sizing problem without backlogging.

■ SB12

Student Commons 2504

Linear and Nonlinear Optimization for Large Scale Data Analysis

General Session

Chair: Yu Du, University of Colorado Denver, Denver, CO, 80202, United States, duyu197@gmail.com

Co-Chair: Steffen Borgwardt, University of Colorado-Denver, Denver, CO, 80202, United States, steffen.borgwardt@ucdenver.edu

1 - The Value of Optimization Based Bound Tightening for Power Network Optimization

Carleton Coffrin, PhD, Los Alamos National Laboratory, Los Alamos, NM, 87545, United States, cjc@lanl.gov, Hassan Lionel Hijazi, Pascal Van Hentenryck

Convexification is a fundamental technique in (mixed-integer) nonlinear optimization and many convex relaxations are parametrized by variable bounds, i.e., the tighter the bounds, the stronger the relaxations. This work investigates how optimization based bound tightening can improve convex relaxations for power network optimization. In particular, this work shows that the Quadratic Convex relaxation of AC power flows, enhanced by bound tightening, results in very small variable domains in the seminal Optimal Power Flow problem. Leveraging these improved bounds, experimental results demonstrate root-node optimality gaps that are less than one percent in the majority of Optimal Power Flow benchmarks.

2 - Discrete Barycenters

Stephan Patterson, University of Colorado Denver, Denver, CO, United States, stephan.patterson@ucdenver.edu, Steffen Borgwardt

Discrete barycenters are the optimal solutions to mass transport problems for discrete measures, and can be computed by linear programming. However, the programs scale exponentially in the number of measures, and so it is of interest to design approximation algorithms. We study a heuristic for the computation of approximate barycenters that will maintain two notable, favorable properties of exact barycenters: a sparse support and a non-mass-splitting optimal transport. This divide-and-conquer approach splits the computation of a barycenter into several parts. Exact barycenters are computed for subsets of the measures, which are then merged using another barycenter computation. We exhibit a significant advantage in computation times and favorable practical results. The approach gives an exact optimum in dimension one, and we show that a single merging step provides a 3-approximation independent of the dimension.

3 - The Quadratic Unconstrained Binary Optimization Problem

Yu Du, University of Colorado Denver, 1475 Lawrence St, Room 5021, Denver, CO, 80202, United States, duyu197@gmail.com

The Quadratic Unconstrained Binary Optimization problem (QUBO), has become a unifying model for representing a wide range of combinatorial optimization problems, and for linking a variety of disciplines that face these problems such as engineering and social sciences. In this talk, we introduce new algorithms to improve the bounds Qx of xQx and an algorithm for partitioned QUBO problems. Also we conduct computational study on linear representations of QKP based test problems with multiple knapsack constraints and large scale variables.

4 - An Introduction to Discrete Barycenters

Steffen Borgwardt, University of Colorado Denver, 3500 Rockmont Drive, Apt 5-303, Denver, CO, 80202, United States, steffen.borgwardt@ucdenver.edu

We provide an introduction to Discrete Wasserstein barycenters. These are the optimal solutions of transportation problems for several marginals of discrete measures. They arise in applications from economics to statistics.
Discrete barycenters exhibit some notable properties: All barycenters have finite support, and there always is a barycenter for which this support is provably sparse. Further, each barycenter allows for a non-mass splitting optimal transport to each of the discrete marginals. Discrete barycenters can be computed by linear programming, but the sizes of these programs scale exponentially. This makes approximation algorithms interesting that trade a small error for a significant reduction in computational effort. We devise a strongly polynomial time 2-approximation that is based on restricting the possible support of the barycenter to the union of supports of the measures. The resulting measure has sparse support, but its transport to the measures will generally split mass. The non-mass split property can be recovered in strongly polynomial time through a local improvement step, but at the cost of a larger support. The combination of sparse support and non-mass split can be achieved by alternating these two steps. We conclude with some computational results.

■ SB13

Student Commons 2600

CPLEX

Tutorial Session

Chair: Ed Klotz, IBM, P.O. Box 4670, Incline Village, NV, 89450, United States, klotz@us.ibm.com

1 - Performance Tuning for CPLEX's Spatial Branch-and-Bound Solver for Global Nonconvex Mixed Integer Quadratic Programs

Ed Klotz, IBM, P.O. Box 4670, Incline Village, NV, 89450, United States, klotz@us.ibm.com

MILP solvers have been improving for more than 40 years, and performance tuning tactics regarding both adjusting solver strategies and model formulations have evolved as well. State-of-the-art global nonconvex MIQP solvers have improved dramatically in recent years, but they lack the benefit of 40 years of evolution. Also, they use a broader notion of branching that can create different performance challenges. This talk will assess the effectiveness of existing MILP tuning tactics for solving nonconvex MIQPs, as well as consider more specific strategies for spatial branching. It will also examine in detail some tightening strategies specific to nonconvex MIQPs involving bilinear terms of binary variables and their associated linearizations.

Saturday, 4:00PM - 5:00PM

■ Plenary V

Student Commons 2600

Discrete Optimization and Network Analysis

Plenary Session

1 - Discrete Optimization and Network Analysis

Illya V Hicks, Rice University, Houston, TX, 77005-1892, United States, ivhicks@rice.edu

Graphs or networks are everywhere and network analysis has garnered significant attention in diverse fields as an effective tool for studying complex, natural and engineered systems. Novel network models of data arising from internet analytics, systems biology, social networks, computational finance, and telecommunications have led to many interesting insights. In this talk, we explore discrete optimization techniques for finding cohesive data within these network-based models. The goal is to detect cohesiveness in spite of missing information (linkages). In this regard, we will explore different aspects of cohesiveness and how they are used for different applications. In particular, we will focus in on three particular structures: k-plex, k-core, and k-club. All three are generalizations of the clique structure and were first utilized for social network analysis. This talk is based on joint work with John Arellano, Baski Balasundaram, Sergiy Butenko, Ben McClosky, and Foad Pajouh.

Sunday, 8:30AM - 9:20AM

■ Plenary VI

Student Commons 2600

A Globally Convergent Cutting-Plane Method for Simulation-Based Optimization with Integer Constraints

Plenary Session

1 - A Globally Convergent Cutting-Plane Method for Simulation-Based Optimization with Integer Constraints

Sven Leyffer, Argonne National Laboratory, MCS Division 9700 South Cass Avenue, Argonne, IL, 60439, United States, leyffer@mcs.anl.gov

Many design and engineering applications result in optimization problems that involve so-called black-box functions as well as integer variables, resulting in mixed-integer derivative-free optimization problems (MIDFOs). MIDFOs are characterized by the fact that a single function evaluation is often computationally expensive (requiring a simulation run for example) and that derivatives of the problem functions cannot be computed or estimated efficiently. In addition, many problems involve integer variables that are non-relaxable, meaning that we cannot evaluate the problem functions at non-integer points.
In the first part of our talk, we survey applications of MIDFO from a range of Department of Energy applications. The design of nano-photonic devices involves integer decision variables due to manufacturing limitations, and each function evaluation requires a finite-element simulation that takes several hours to run on a cluster. Similarly, automatic performance tuning of code snippets for high-performance computing involves non-relaxable integer variables such as loop-unroll-factors and compiler options and require several runs to eliminate random measurement errors. Finally, the design and operation of concentrating solar plants, requires forward simulations that take hours on a desktop and involve unrelaxable decision such as the number of panels on the receiver.
In the second part of our talk, we present a new method for non-relaxable MIDFO that enables us to prove global convergence under idealistic convexity assumptions. To the best of our knowledge this is the first globally convergent method for non-relaxable MIDFO apart from complete enumeration. Our method constructs hyperplanes that interpolate the objective function at previously evaluated points. We show that in certain portions of the domain, these hyperplanes are valid underestimators of the objective, resulting in a set of conditional cuts. The union of these conditional cuts provide a nonconvex underestimator of the objective. We show that these nonconvex cuts can be modeled as a standard mixed-integer linear program (MILP). Unfortunately, this MILP model turns out to be prohibitively expensive to solve even with state-of-the-art MILP solvers. We develop an alternative approach that is computationally tractable, and provide some early numerical experience with our new method.
Co-Authors: Prashant Palkar, Jeffrey Larson, and Stefan Wild

Sunday, 9:30AM - 10:20AM

■ Plenary VII

Student Commons 2600

Large-scale Open Pit Mine Production-Scheduling

Plenary Session

1 - Large-scale Open Pit Mine Production-Scheduling

Marcos Goycoolea, Universidad Adolfo Ibañez, Diagonal Las Torres 2640, oficina 534 C., Santiago, 7941169, Chile, marcos.goycoolea@uai.cl

The central concern of strategic mine planning is the construction of a tentative production schedule. This is a life-of-mine plan (20-50 years) specifying which part of a mineral deposit should be extracted, when it should be extracted, and how it should be extracted so as to maximize the net present value of the mining project (easily in the hundreds of millions of dollars).
Strategic mine planning is a complex optimization problem made daunting both by the great number of activities that must be scheduled in time and the great uncertainty concerning key economic, geological and operational parameters. Despite being a fairly standard problem regularly faced my mining projects throughout the world, it has received very little attention from the operations research community. In this talk we describe our efforts modelling this problem as an RCPSP (Resource Constrained Project Scheduling Problem), the integer-programming techniques we have been using to tackle them, and our experience working with Newmont and Barrick gold, the world's biggest gold producers. This talk covers joint work with Andrea Brickey, Daniel Espinoza, Barry King, Eduardo Moreno, Gonzalo Muñoz, Alexandra Newman and Orlando Rivera.

Sunday, 11:00AM - 12:00PM

■ SUN01

North Classroom 1806

Second-Order Cone Programming

Contributed Session

Chair: Ali Mohammad Nezhad, Lehigh University, Bethlehem, PA, 18015, United States, alm413@lehigh.edu

1 - Parametric Semidefinite and Second Order Conic Optimization

Ali Mohammad Nezhad, Doctoral Student, Lehigh University, 200 W. Packer Ave, Mohler Lab, Bethlehem, PA, 18015, United States, alm413@lehigh.edu, Tamas Terlaky

We consider the parametric and stability analysis of semidefinite and second-order cone optimization problems when the objective function is perturbed along a fixed direction. We study the continuity and Lipschitz continuity of the optimal set mapping in terms of set convergence. We compute the length of the nonlinearity interval of the optimal value function when strict complementarity and nondegeneracy conditions hold. Finally, we show that under the primal and dual nondegeneracy conditions, the change in the optimal partition can be predicted using the standard implicit function theorem.

■ SUN03

North Classroom 1202

Stochastic and Robust Integer Programs: Theory and Applications

General Session

Chair: Bismark Singh, Sandia National Laboratories, Albuequerque, TX, United States, bsingh@sandia.gov

1 - Two-stage Stochastic P-order Conic Mixed Integer Programs: Tight Second-stage Formulations

Manish Bansal, Virginia Tech., 227 Durham Hall, 1145 Perry Street, Blacksburg, VA, 24060, United States, bansal@vt.edu, Yingqiu Zhang

We study two-stage stochastic p-order conic mixed integer programs (TSS-CMIPs) in which the second-stage problems have sum of lp-norms in the objective function along with integral constraints. We present sufficient conditions under which the addition of globally valid parametric non-linear inequalities provides convex programming equivalent for the second stage CMIPs. In addition, we introduce TSS-CMIPs with structured p-order CMIPs in the second stage and derive new classes of valid inequalities that satisfy the foregoing conditions. We perform computational experiments by solving randomly generated structured TSS-CMIPs and and observe that after adding our parametric cuts, there is a significant reduction in the total solution time taken to solve these problems instances.

2 - Chance-Constrained Optimization Models for Securing Critical Electricity Grid Infrastructure

Bismark Singh, Sandia National Laboratories, Albuqurque, NM, United States, bsingh@sandia.gov, Jean-Paul Watson

We develop two-stage stochastic mixed integer programs with chance-constraints for designing a secure electrical grid. We are motivated by classical attacker-defender interdiction models. We demonstrate preliminary results on national scale data.

■ SUN07

North Classroom 3004

Inverse Optimization

Contributed Session

Chair: Chaosheng Dong, University of Pittsburgh, 4200 Fifth Avenue, Pittsburgh, PA, 15260, United States, chaosheng@pitt.edu

1 - Inverse Optimization Based Learning of Unknown Metrics in Large Networks

Rupei Xu, University of Texas at Dallas, Richardson, TX, United States, rxx130430@utdallas.edu

In numerous systems, especially in communication networks, it is a key functionality to find optimal paths in massive graphs. In order to understand the system, we need to know the specific path metric under which the optimal paths are chosen. In many cases, however, no explicit knowledge of the metric is available, due to the multitude of factors that implicitly influence the selections. We can only observe the eventual path choices that are made by an unknown mechanism between various end nodes, but we are not familiar with the underlying metric. Our goal is to learn the unknown metric, as accurately as possible, purely from the observed path choices. We present an inverse optimization based mathematical model, along with a solution algorithm, to handle this problem. Our main result is that the unknown path metric can be optimally learned from the observed path choices by a polynomial time algorithm, if we assume that the metric is additive, but otherwise arbitrary. Thus, the powerful tool of inverse optimization offers an efficient learning method for the considered problem.

2 - Inverse Optimization through Online Learning

Chaosheng Dong, University of Pittsburgh, 4200 Fifth Avenue, Pittsburgh, PA, 15260, United States, chaosheng@pitt.edu, Bo Zeng

We study a learning challenge to infer decision makers' utility functions (and restrictions), based on observed decisions. We formulate such problem as an inverse optimization model, and develop an Online Learning algorithm to handle noisy data. Numerical results show that the algorithm can learn the parameters with great accuracy and is very robust to noises.

■ SUN11

Student Commons 2500

Combinatorial Optimization in Sports

Invited: Optimal Control

Invited Session

Chair: David Morton, Northwestern University, IEMS Department, 2145 Sheridan Road, Evanston, IL, 60208, United States, david.morton@northwestern.edu

1 - NBA Playoff Picture with Tie-breaking Criteria

Mark A Husted, Colorado School of Mines, Golden, CO, 80401, United States, markahusted@gmail.com, Eli Olinick, Nelson Winbush

The National Basketball Association (NBA) is divided into two conferences, each of which is comprised of fifteen teams. At the end of the regular season, the top eight teams from each conference, based on winning percentage, compete in the playoffs. An integer-programming model determines when a team has guaranteed its position in the playoffs, or, conversely, when it has been eliminated from the playoffs before the completion of the regular season. At the end of the regular season, there are instances in which teams' winning percentages are tied. Ties are broken using seven independent criteria based on the number of teams tied, and we implement these by determining: (i) when each a has been eliminated from being first place in the conference; (ii) how many games a team must win in order to clinch first place in the conference; (iii) when a team has been eliminated from the playoffs; and (iv) how many games a team must win in order to clinch a playoff position. The results are published on the RIOT website so fans can follow their favorite teams' playoff standings.

2 - Optimal Clustering on a Graph

David Morton, Northwestern University, IEMS Department, 2145 Sheridan Road, Evanston, IL, 60208, United States, david.morton@northwestern.edu, Gokce Kahvecioglu

We study a hierarchical clustering problem on an undirected graph with a weight function assigning nonnegative weights to the edges. We remove a subset of edges to break the graph into a number of smaller pieces, i.e., clusters. We consider a bicriteria graph clustering problem, in which we maximize the number of clusters while minimizing the weight of deleted edges. Solving this bicriteria problem parametrically identifies solutions that lie on the concave envelope of the efficient frontier, and the breakpoints on this envelope are nested, yielding a hierarchical family of clusters. We illustrate our ideas using NCAA football schedules, attempting to identify conferences, divisions, etc.

■ SUN12

Student Commons 2504

Decomposition for Integer Programming

Contributed Session

Chair: Ragheb Rahmaniani, Universite de Montreal, Pavillon Andre-Aisenst, CIRRELT P.O. Box 6128, Montreal, QC, H3T 1J4, Canada, Ragheb.rahmaniani@gmail.com

1 - Accelerating Mathematical Programming Techniques with the Corridor Method

Marco Caserta, IE University, Maria de Molina 31B, Madrid, 28006, Spain, marco.caserta@ie.edu, Stefan Voss

In this paper we investigate how three widely used decomposition techniques can be accelerated when intertwined with a matheuristic. More precisely, we consider the Benders decomposition, Lagrangean relaxation, and Dantzig-Wolfe reformulation techniques and hybridize them with the corridor method. We exemplify the proposed approaches on the capacitated lot sizing problem with setups. Our computational analysis on a set of benchmark instances shows that all the decomposition methods benefit from the hybridization with the corridor method, and that this is especially evident in the case of the Benders decomposition algorithm.

2 - A Decomposition Method for Mixed-integer Problems

Ragheb Rahmaniani, Universite de Montreal, Pavillon Andre-Aisenst, CIRRELT P.O. Box 6128, Montreal, QC, H3T 1J4, Canada, Ragheb.rahmaniani@gmail.com, Shabbir Ahmed, Teodor Gabriel Crainic, Michel Gendreau, Walter Rei

In this talk, we present the Benders dual decomposition (BDD) method for mixed-integer linear problems. The development of the BDD is based on a specific reformulation of the Benders subproblem, where local copies of the master variables are introduced and then priced out into the objective function. We show that the BDD is capable of: (i) generating stronger feasibility and optimality cuts compared to the classical Benders method, (ii) can converge to the optimal integer solution at the root node of the Benders master problem and (iii) is capable of generating high quality incumbent solutions at the early iterations of the algorithm. Moreover, we demonstrate that the BDD method is also capable of handling nonlinearities and integrality requirements in the subproblem's formulation when applied to a certain class of optimization problems. We present encouraging numerical results in comparison with the classical decomposition methods and CPLEX 12.7.

■ SUN13

Student Commons 2600

Castle Labs

Tutorial Session

Chair: Warren B. Powell, Princeton University, 230 Sherrerd Hall, Dept. of Operations Research and Financial Eng, Princeton, NJ, 08544, United States, powell@princeton.edu

1 - A Unified Modeling and Algorithmic Framework for Optimization under Uncertainty

Warren B. Powell, Princeton University, 230 Sherrerd Hall, Dept. of Operations Research and Financial Eng, Princeton, NJ, 08544, United States, powell@princeton.edu

Stochastic optimization is a fragmented field that spans topics such as stochastic programming, Markov decision processes, reinforcement learning, bandit problems and stochastic search, featuring different modeling styles and notational systems, and competing solution strategies. I will give a common mathematical framework for modeling all of these problems, illustrated using a range of applications. This framework consists of five fundamental elements, and requires optimizing over policies, which is the major point of departure. I will identify four (meta)classes of policies which will work, in some combination, over all problems. I will then highlight the role of offline vs. online learning, and state-independent vs. state-dependent problems.