The 20th LIDS Student Conference

is held on Jan 29-30, 2015 in Rooms 32-141 and 32-155.

The Conference Program

more information on `lidsconf.mit.edu`

This event is sponsored by

DRAPER LABORATORY ENGINEERING INNOVATIONS

with additional support from

Prof. Peter Falb

Welcome to the 20th LIDS Student Conference!

On behalf of everyone at the Laboratory for Information and Decision Systems (LIDS) at MIT, it is our great pleasure to welcome you to the 2015 LIDS Student Conference. The two days of the conference feature student presentations, four plenary talks, and a panel discussion, showcasing many of the research themes and fields of study pursued at LIDS.

The conference is an excellent opportunity for students to present their work to both their peers and the wider community, gaining valuable feedback and presentation experience. The conference also provides an overview of the work done at LIDS, serving to foster communication and collaboration between research groups within LIDS and across MIT.

We are delighted to have with us as distinguished plenary speakers:

Jennifer Tour Chayes
Managing Director,
Microsoft Research New England and Microsoft Research New York City.

Ramesh Johari
Associate Professor, Management Science and Engineering,
Stanford University.

Sridevi V. Sarma
Assistant Professor of Biomedical Engineering, Institute for Computational Medicine,
Johns Hopkins University.

Martin J. Wainwright
Professor of Electrical Engineering and Computer Science, Professor of Statistics,
UC Berkeley.

The entire conference will take place in the Stata Center (building 32) at MIT. All plenary and oral student presentations will be in rooms 32-155 and 32-144, poster presentations and the student social will be held in the 6th floor lounge in LIDS, and the banquet will be held in the Catalyst Restaurant.

Finally, we would like to acknowledge those whose help has been invaluable in planning this conference and making it happen: Jennifer Donovan, Lynne Dell, and Brian Jones at LIDS headquarters, and our student and plenary speakers and our panelists for their willingness to tell us about their work and ideas! We would also like to thank all student volunteers for their help in organizing this event. We would like to thank Draper Laboratory for their ongoing sponsorship of this event over many years, and Prof. Peter Falb for providing additional support. Finally we would like to thank the attendees for making this event a success.

The 20th LIDS Student Conference Co-chairs,

Christina Lee, Shreya Saxena and Omer Tanovic

Thursday, Jan 29, Room 32-141

08:30 - 09:30   **Breakfast and Registration**

09:30 - 09:45   **Opening remarks**

09:45 - 10:45   **Plenary Talk**

Prof. Sridevi Sarma. On the Therapeutic Mechanisms of Deep Brain Stimulation for Parkinson's disease: Why High Frequency?

10:45 - 11:00   **Coffee break**

11:00 - 12:15   **Student Session on Controls and Communication**

Jordan Romvary. A Robustness Approach to the Stability Analysis of Linear Systems in the Presence of Model Uncertainty

Shreya Saxena. Neuronal Decoding in the Motor System: Considerations in Feedback

Omer Tanovic. Discrete-Time Models Resulting From Dynamic Continuous-Time Perturbations In Phase-Amplitude Modulation-Demodulation Schemes

Andrew Young. Combinatorial Joint Source-Channel Coding for the BSC

Dawsen Hwang. Distributed Multi-Depot Routing without Communications

12:15 - 01:30   **Lunch and Poster Session**

01:30 - 02:30   **Plenary Talk**

Dr. Jennifer Tour Chayes. The Power of Locality for Network Algorithms

02:30 - 02:45   **Coffee break**

02:45 - 03:45   **Student Session on Networks**

Yongwhan Lim. A Simple Model of Cascades in Networks

Marzieh Parandehgheibi. Can Interdependency make the Networks Robust?

Elie Adam. Towards an Algebra for Cascade Effects

Quan Li. On the Max-Cut over sparse random graph

03:45 - 04:00   **Coffee break**

04:00 - 05:00   **Student Session on Inference and Information**

Hajir Roozbehani. When is efficient inference (im)possible?

Jennifer Tang. Analysis of Models for Physical Redundancy

Abigail Horn. A Framework For Attributing Probabilities to Possible Sources of Foodborne Disease Outbreaks

Beipeng Mu. Focused Information Gathering and an application in SLAM

05:15   **LIDS Social**

Friday, Jan 30, Room 32-155

`08:30 - 09:30`   **Breakfast**

`09:30 - 10:30`   **Plenary Talk**

Prof. Ramesh Johari. Can I Take a Peek? Continuous Monitoring of Online A/B Tests

`10:30 - 10:45`   **Coffee break**

`10:45 - 12:15`   **Student Session on Optimization and Approximation Algorithms**

Diego Cifuentes. Chordal structure and polynomial systems

Frank Permenter. The linear images of convex cones: face lattices, lower bounds, and extension complexity

Nikita Korolko. Multiperiod Optimization for Fleet Defense: Centralized and Distributed Approaches

Alexandre Jacquillat. An Integrated Scheduling and Operations Approach to Airport Congestion Mitigation

Christina Lee. Solving Systems of Linear Equations: Locally and Asynchronously

Qingqing Huang. A greedy algorithm for approximate non-negative matrix factorization

`12:15 - 01:30`   **Lunch (by invitation only)**

`01:30 - 02:30`   **Panel Discussion**

`02:30 - 02:45`   **Coffee break**

`02:45 - 04:00`   **Student Session on Decision Systems and Game Theory**

Vivek Sakhrani. Collaborative Design with Boundary Objects

Ellen Czaika. Model use in sustainability negotiations

Stephen Zoepf. Decision Models of Technology

Maite Pena-Alcaraz. Analysis of capacity pricing and allocation mechanisms in shared railway systems

Swati Gupta. Games people (could not) play

`04:00 - 04:15`   **Coffee break**

`04:15 - 05:15`   **Plenary Talk**

Prof. Martin Wainwright. Statistics meets Computation: Some vignettes from the interface

`06:00`   **Banquet (by invitation only)**

Plenary Talk I.

Thursday, Jan 29, Room 32-141

Sridevi V. Sarma,

Assistant Professor of Biomedical Engineering, Institute for Computational Medicine,
Johns Hopkins University

On the Therapeutic Mechanisms of Deep Brain Stimulation for Parkinson's disease: Why High Frequency?

Abstract.   Deep brain stimulation (DBS) is clinically recognized to treat movement disorders in Parkinsons disease (PD), but its therapeutic mechanisms remain elusive. One thing is clear though: high frequency periodic DBS (130-180Hz) is therapeutic, while low frequency DBS is not therapeutic and may even worsen symptoms. So, what is so special about high frequency? In this talk, we address this question by discussing our viewpoint supported by recent results from our key studies of the thalamo-cortical-basal ganglia motor network.

First, thalamic cells play a pivotal role in performing movements by selectively relaying motor-related information back to cortex under the control of modulatory signals from the basal ganglia (BG). Through computational models of thalamic cells, systems theory and analysis, and single unit recordings from primates, we show that (i) there is a set of BG signals ("Proper Relay Set", PRS), under which the thalamic cells can reliably relay the motor commands, and that (ii) the BG signals belong to the PRS in healthy conditions but are outside the PRS under PD conditions.

Then, we use a detailed computational model of the motor network under PD conditions to study the effects of DBS on the BG signals projecting to the thalamic cells. We show that high frequency periodic DBS steers the BG signals back to the PRS while lower frequency regular DBS and irregular DBS do not. Furthermore, we show that DBS pulses evoke inputs that propagate through the motor circuit both orthodromically (i.e., forward) and antidromically (i.e., backward) and fade away within a few milliseconds, thus having little effects on the BG signals. However, when the latency between consecutive DBS pulses is small (i.e., DBS is high frequency) and constant over time (i.e., DBS is periodic), then orthodromic and antidromic effects overlap in the loop and result in a strong, long-lasting perturbation that ultimately drives the BG signals back to the PRS.

Taken together, these results provide a holistic view of motor control in healthy and PD conditions, account for the neural mechanisms of therapeutic DBS, and suggest that the merit of DBS critically depends on loop delays in the closed-loop thalamo-cortical-basal ganglia system.

Biography. Sridevi V. Sarma (M04) received the B.S. degree in electrical engineering from Cornell University, Ithaca NY, in 1994; and an M.S. and Ph.D. degrees in Electrical Engineering and Computer Science from Massachusetts Institute of Technology in, Cambridge MA, in 1997 and 2006, respectively. She was a Postdoctoral Fellow in the Brain and Cognitive Sciences Department at the Massachusetts Institute of Technology, Cambridge, from 2006-2009. She is now an assistant professor in the Institute for Computational Medicine, Department of Biomedical Engineering, at Johns Hopkins University, Baltimore MD. Her research interests include modeling, estimation and control of neural systems using electrical stimulation. She is a recipient of the GE faculty for the future scholarship, a National Science Foundation graduate research fellow, a LOreal For Women in Science fellow, the Burroughs Wellcome Fund Careers at the Scientific Interface Award, the Krishna Kumar New Investigator Award from the North American Neuromodulation Society, and a recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE).

Plenary Talk II.

Thursday, Jan 29, Room 32-141

Jennifer Tour Chayes,

Managing Director,
Microsoft Research New England and Microsoft Research New York City

The Power of Locality for Network Algorithms

Abstract. Given the massive size of many networks, conventional algorithms which scale as polynomials in the network size are woefully inadequate. In the first part of this talk, we consider how to use locality to deliver much more efficient algorithms (quasi-linear or even sub-linear in the network size) for quantities and questions like pagerank, coverage, diffusion, and determining the most influential nodes. In the second part of this talk, we consider another aspect of locality, namely the question of local data access. Many large networks are encoded locally, e.g., with adjacency lists. How does the nature of the data access impact the efficiency of algorithms? Surprisingly, we show that small differences in data access can lead to very large differences in efficiency and approximability of network algorithms. As an example, we consider a covering problem which arises naturally for recruiters on social networks, and show how small differences between the information on neighbors in LinkedIn and Facebook lead to large differences in their utility to recruiters.

Biography. Jennifer Tour Chayes is Distinguished Scientist, Managing Director and Co-founder of Microsoft Research New England and Microsoft Research New York City. Before joining Microsoft in 1997, Chayes was for many years Professor of Mathematics at UCLA. Chayes is the author of over 125 academic papers and the inventor of over 30 patents. Her research areas include phase transitions in discrete mathematics and computer science, structural and dynamical properties of self-engineered networks, graph theory, graph algorithms, algorithmic game theory, and computational biology.

Chayes holds a BA in biology and physics from Wesleyan, where she graduated first in her class, and a PhD in mathematical physics from Princeton. She did postdoctoral work in the Mathematics and Physics Departments at Harvard and Cornell. She is the recipient of an NSF Postdoctoral Fellowship, a Sloan Fellowship, the UCLA Distinguished

Teaching Award, and the ABI Women of Leadership Vision Award. She has twice been a member of the IAS in Princeton. Chayes is a Fellow of the American Association for the Advancement of Science, the Fields Institute, the Association for Computing Machinery, and the American Mathematical Society, and an Elected Member of the American Academy of Arts and Sciences.

Plenary Talk III.

Friday, Jan 30, Room 32-155

Ramesh Johari,

Associate Professor, Management Science and Engineering,
Stanford University

Can I Take a Peek? Continuous Monitoring of Online A/B Tests

Abstract.   A/B testing is a hallmark of Internet services: from e-commerce sites to social networks to marketplaces, nearly all online services use randomized experiments as a mechanism to make better business decisions. Such tests are generally analyzed using classical frequentist statistical measures: p-values and confidence intervals. Despite their ubiquity, these reported values are computed under the assumption that the experimenter will not continuously monitor their test—in other words, there should be no repeated peeking at the results that affects the decision of whether to continue the test. On the other hand, one of the greatest benefits of advances in information technology, computational power, and visualization is precisely the fact that experimenters can watch experiments in progress, with greater granularity and insight over time than ever before.

We ask the question: if users will continuously monitor experiments, then what statistical methodology is appropriate for hypothesis testing, significance, and confidence intervals? We present recent work addressing this question. In particular, building from results in sequential hypothesis testing, we present analogues of classical frequentist statistical measures that are valid even though users are continuously monitoring the results.

Joint work with Leo Pekelis and David Walsh. (This work was carried out with Optimizely, a leading A/B testing platform.)

Biography.   Ramesh Johari is an Associate Professor at Stanford University and the Cisco Faculty Scholar in the School of Engineering, with a full-time appointment in the Department of Management Science and Engineering (MS and E), and courtesy appointments in the Departments of Computer Science (CS) and Electrical Engineering (EE). He is a member of the Operations Research group in MS and E, the Information Systems Laboratory in EE,

and the Institute for Computational and Mathematical Engineering. He received an A.B. in Mathematics from Harvard (1998), a Certificate of Advanced Study in Mathematics from Cambridge (1999), and a Ph.D. in Electrical Engineering and Computer Science from MIT (2004).

Plenary Talk IV.

Friday, Jan 30, Room 32-155

Martin Wainwright,

Professor of Electrical Engineering and Computer Science, Professor of Statistics,
UC Berkeley

Statistics meets Computation: Some vignettes from the interface

Abstract. In the modern era of massive data sets, computational considerations have become increasingly important in statistics. In this talk, we discuss some problems that lie at the interface between statistics and algorithms, including rigorous guarantees for non-convex optimization problems in statistics, and various forms of optimality in randomized sketching methods.

Biography. Martin Wainwright joined the faculty at University of California at Berkeley in Fall 2004, and is currently a Professor with a joint appointment between the Department of Statistics and the Department of Electrical Engineering and Computer Sciences. He received his Bachelor's degree in Mathematics from University of Waterloo, Canada, and his Ph.D. degree in Electrical Engineering and Computer Science (EECS) from Massachusetts Institute of Technology (MIT), for which he was awarded the George M. Sprowls Prize from the MIT EECS department in 2002. He is interested in high-dimensional statistics, information theory and statistics, and statistical machine learning. He has received an Alfred P. Sloan Foundation Research Fellowship (2005), IEEE Best Paper Awards from the Signal Processing Society (2008) and Communications Society (2010); the Joint Paper Award from IEEE Information Theory and Communication Societies (2012); a Medallion Lecturer (2013) of the Institute for Mathematical Statistics; a Section Lecturer at the International Congress of Mathematicians (2014); and the COPSS Presidents' Award in Statistics (2014). He is currently serving as an Associate Editor for the Annals of Statistics, Journal of Machine Learning Research, Journal of the American Statistical Association, and Journal of Information and Inference.

Panel Discussion

We will be holding a panel discussion on Friday Jan 30, 1:30 to 2:30 pm in Room 32-155.

LIDS Storytelling: History, Impact, and Vision of LIDS.

For the 20th anniversary of the LIDS student conference, speakers who have played important roles in the history of LIDS will share their insights about the past, present, and future of LIDS. Come hear about the key developments in LIDS history and the impact of LIDS in transforming academia and industry. There will be opportunities to ask their perspectives towards the direction of LIDS moving forward, and hear their wisdom and advice for young researchers looking for ideas and inspiration.

Panelists. The panel will be moderated by Prof. Munther Dahleh, and composed of:

> Prof. Dimitri Bertsekas
>
> Dr. John Dowdle
>
> Prof. Bob Gallager
>
> Prof. Sanjoy Mitter
>
> Prof. John Tsitsiklis

Session on Controls and Communication.

JORDAN ROMVARY
A Robustness Approach to the Stability Analysis of Linear Systems in the Presence of Model Uncertainty

Abstract.   The problem of model uncertainty in the study of dynamical systems has been proven to be essential to both the understanding of stability as well as in other important applications, most notably adaptive control settings like flight control. In this work, we will explore model uncertainty in the domain of linear (possibly time-varying) systems. We will utilize insights from geometry (numerical ranges) and the theory of differential equations (logarithmic norms) to develop a more computationally efficient robustness approach to the stability analysis of a dynamical system in the presence of model uncertainty. Our major application domain will be in the area of adaptive control.

SHREYA SAXENA
Neuronal Decoding in the Motor System: Considerations in Feedback

Abstract.   Driving a prosthetic device using signals recorded from the motor cortex has been demonstrated in several experimental settings with remarkable success. However, in the regime of fast movements, additional tools may need to be developed in order to decode the movement accurately. In accordance with previous work, we formulate movement generation as a feedback control problem. In these formulations, the output is usually the movement or the muscle activity, and the plant and the controller are relevant structures of the brain and body. Despite the knowledge that information passes through neurons in order to make movements, one usually sidesteps the binary nature of neuronal encoding and treats the structures of the brain as continuous input - continuous output systems in closed loop. While preliminary results are available for decoding a signal encoded by a realistic binary neuron model in open loop, with the caveat that the decoder needs access to all spikes, in practical applications, we need to understand the effect of binary neuronal encoding and real-time decoding on the closed loop feedback control formulation. In neural decoding of fast movements, we would also like to reconstruct the reference signal given a finite set of samples of the neural activity. The motor neurons are modeled here as integrate-and-fire (IAF) encoders, which can be cast as irregular sampling devices. We use concepts in sampling theory and robust control theory to, i) develop a real-time decoder for neuronal spikes with clear convergence guarantees, and ii) provide explicit conditions for the well-posedness and stability of the encoder-decoder pair embedded in a feedback loop. This may help us further to develop conditions for the robust tracking of a reference movement signal by the output of the muscle model. The stable implementation of a binary encoder and a causal decoder in a feedback system gives us some insight into a possible method of neuronal decoding in the human body, with the muscle model and the feedback systems being tunable

in order to reflect more realistic models of the corresponding biological systems. Moreover, the accurate reconstruction of the driving signal from neuronal spikes leads us to believe that the concepts developed here may be used to drive further research in prosthetic control via neuronal decoding.

OMER TANOVIC
Discrete-Time Models Resulting From Dynamic Continuous-Time Perturbations In Phase-Amplitude Modulation-Demodulation Schemes

Abstract. We consider discrete-time (DT) systems S in which a DT input is first transformed to a continuous-time (CT) format by phase-amplitude modulation, then modified by a non-linear CT dynamical transformation F, and finally converted back to DT output using an ideal de-modulation scheme. Assuming that F belongs to a special class of CT Volterra series models with fixed degree and memory depth, we provide a complete characterization of S as a series connection of a DT Volterra series model of fixed degree and memory depth, and an LTI system with special properties. The result suggests a new, non-obvious, analytically motivated structure of digital compensation of analog nonlinear distortions (for example, those caused by power amplifiers) in digital communication systems. Results from a MAT-LAB simulation are used to demonstrate effectiveness of the new compensation scheme, as compared to the standard Volterra series approach.

ANDREW YOUNG
Combinatorial Joint Source-Channel Coding for the BSC

Abstract. Traditional error correction and source coding has focused on the stochastic setting where separation based schemes are optimal, and current solutions for applications requiring both lossy compression and noise resilience reflect this approach. However, in the combinatorial setting, with worst case errors, separation based schemes are far from being even asymptotically optimal. This work investigates fundamental limits, achievability and converse bounds, duality and monotonicity for combinatorial joint source-channel codes (CJSCC). Particular attention is paid to the binary symmetric channel (BSC) where it is shown, among other things, that coding with greater than unit bandwidth expansion factor probably yields no gain when the channel parameter is greater than a forth.

DAWSEN HWANG
Distributed Multi-Depot Routing without Communications

Abstract. We consider and formulate a class of distributed multi-depot routing problems, where servers are to visit a set of requests, with the aim of minimizing the total distance travelled by all servers. These problems fall into two categories: distributed offline routing

problems where all the requests that need to be visited are known from the start; distributed online routing problems where the requests come to be known incrementally. A critical and novel feature of our formulations is that communications are not allowed among the servers, hence posing an interesting and challenging question: what performance can be achieved in comparison to the best possible solution obtained from an omniscience planner with perfect communication capabilities? The worst-case (over all possible request-set instances) performance metrics are given by the approximation ratio (offline case) and the competitive ratio (online case).

Our first result indicates that the online and offline problems are effectively equivalent: for the same request-set instance, the approximation ratio and the competitive ratio differ by at most an additive factor of 2, irrespective of the release dates in the online case. Therefore, we can restrict our attention to the offline problem. For the offline problem, we show that the approximation ratio given by the Voronoi partition is m (the number of servers). For two classes of depot configurations, when the depots form a line and when the ratios between the distances of pairs of depots are upper bounded by a sublinear function f(m) (i.e., f(m) = o(m)), we give partition schemes with sublinear approximation ratios O(log m) and Θ(f(m)) respectively. We also discuss several interesting open problems in our formulations: in particular, how our initial results (on the two deliberately chosen classes of depots) shape our conjecture on the open problems.

Poster Session.

MARZIEH PARANDEHGHEIBI
Investigating the Impact of Communication Loss on Power Grid Stability during a Large Disturbance

Abstract.   In the literature of interdependency between power grid and communication network (e.g. Buldyrev2010) it is assumed that failures from communication network cascade to the power grid and it is modeled as a point-wise failure; i.e. every power node that loses its connection to the communication network fails. However, this model is too simplistic to capture the complex interaction between the two networks. In this work, we study the impact of communication loss on the power grid. In particular, we propose a wide-area control scheme implemented by the communication network that is responsible for the control of the grid in the presence of a disturbance. We apply our control scheme on several power-communication failure scenarios and our results indicate that the communication loss could lead to lower yield (percentage of served load) compared to scenarios that we have full communication. Moreover, we show that the impact of communication loss on power grid is a function of several factors including the size and topology of uncontrollable areas in power grid and the size of power disturbance. Finally, we consider two pre-defined scenarios for the operation of uncontrollable nodes; (1) operating in the latest state, before communication loss (2) tripping. We show a counter-intuitive result that there exist scenarios that it is better to trip nodes in an uncontrollable area than operating them in their latest states.

HAMZA FAWZI
Equivariant semidefinite lifts and sum-of-squares hierarchies

Abstract.   A central question in optimization is to maximize (or minimize) a linear function over a given polytope P. To solve such a problem in practice one needs a concise description of the polytope P. In this paper we are interested in representations of P using the positive semidefinite cone: a positive semidefinite lift (psd lift) of a polytope P is a representation of P as the projection of an affine slice of the dxd positive semidefinite cone. Such a representation allows linear optimization problems over P to be written as semidefinite programs of size d. Such representations can be beneficial in practice when d is much smaller than the number of facets of the polytope P. In this paper we are concerned with so-called equivariant psd lifts (also known as symmetric psd lifts) which respect the symmetries of the polytope P. We present a representation-theoretic framework to study equivariant psd lifts of a certain class of symmetric polytopes known as regular orbitopes. Our main result is a structure theorem where we show that any equivariant psd lift of size d of a regular orbitope is of sum-of-squares type where the functions in the sum-of-squares decomposition come from an invariant subspace of dimension smaller than d^2. We use this framework to study two well-known families of polytopes, namely the parity polytope and the cut polytope, and we prove exponential lower bounds for any equivariant psd lifts of these polytopes.

YONGWHAN LIM
Competition over Product Diffusion in Social Networks

Abstract. We consider a setting in which two firms compete to diffuse their products in a social network. Firm seed their products simultaneously and products diffuse according to the linear threshold model. Consumers have (randomly drawn) heterogeneous thresholds for each product. Using the concept of cascade centrality introduced by [1], we provide a sharp characterization of networks that admit pure-strategy Nash equilibria (PSNE). We provide tight bounds for the efficiency of these equilibria and for the inequality in firms' equilibrium payoffs. We apply our results to competition over product diffusion in tree networks. In this case, a PSNE always exists and can be efficiently found. The model can be extended to arbitrary budgets and to products of different quality.

MICHAEL DAVIDSON
Incorporating Data-Rich Wind Power Forecasts in an Adaptive Robust Unit Commitment

Abstract. Intermittent generation increases the variability and uncertainty of commitment and dispatch schedules, raising costs to maintain load balance and adequate reserves. Furthermore, analytical methods to minimize operational costs based on assumed probability distributions for forecast error such as stochastic programming are generally intractable for real-world systems. Thus, when running the day-ahead unit commitment (UC) optimization power system operators typically use heuristics based on historical data to calculate required levels of costly reserves to mitigate intermittency. Robust optimization, which minimizes the cost of the worst-case realization of uncertain variables, is a tractable method of incorporating large amounts of data from stochastic phenomena. Uncertainty sets the feasible space of uncertain variables can be constructed to account for complicated relationships among the data and, beneficially, no probability distribution must be assumed. Typically, robust UC models calculate forecast errors as deviations of actual from a most probable forecast with a fixed time horizon. Uncertainty sets may be bounded by these or may include a budget of uncertainty to allow for correlation between consecutive forecast errors. However, these ignore the often rich data produced by forecast models and the different quality of forecasts for the beginning and end of the commitment day. A novel uncertainty set construction for wind power uncertainty is presented here that accounts for multiple forecasts per hour and the time-varying accuracy of forecasts within a single 24-hour period. Hourly forecast and actual wind generation data over 2013-2014 were obtained from ERCOT, which include both the most probable and 20th percentile forecasts. A two-stage robust UC minimizing the worst-case dispatch of a reduced ERCOT-like system was then constructed using a polyhedral uncertainty set from the intersection of convex hulls of forecast errors parameterized by (1) the time of the forecast (10 to 34 hours ahead) and (2) the difference between the most probable and 20th percentile forecasts. Affine adaptive robust solutions (affine in wind forecast error) were compared to nominal deterministic solutions for 10 sample days, achieving some gains in worse cases but too conservative on average.

Calculations show that at most 40a budget of uncertainty will be implemented based on consecutive historical forecast errors, and compared to the nominal case as well as formulations without rich forecast data. Scaling this formulation to incorporate more outputs from forecast models will also be demonstrated.

CHRISTINA LEE
Solving Systems of Linear Equations: Locally and Asynchronously

Abstract. We consider approximating a single component of the solution to a system of linear equations $Ax = b$, where A is an invertible real matrix and b is a n-dimensional real vector. If A is either diagonally dominant or positive definite, we can equivalently solve for x_i in $x = Gx + z$ for some matrix G and vector z such that the spectral radius of G is less than 1. Existing algorithms either focus on computing the full vector x or use Monte Carlo methods to estimate a component x_i under the condition that the maximum L1 norm of any row of G is less than 1. We consider the setting where n is large, yet G is sparse, i.e., each row has at most d nonzero entries. We present synchronous and asynchronous randomized variants of a local algorithm which relies on the Neumann series characterization of the component x_i, and allows us to limit the sparsity of the vectors involved in the computation, leading to improved convergence rates. Both variants of our algorithm produce an estimate with an error bounded by epsilon times the L2 norm of x. We provide convergence guarantees when the 2-norm of G is less than 1, thus encompassing a larger class of systems. The asynchronous local algorithm adaptively samples one coordinate to update among the nonzero coordinates of the current iterate in each time step. We prove with high probability that the error contracts by a time varying factor in each step, guaranteeing that the algorithm converges to the correct solution. We prove that our algorithms obtain an approximation for x_i in constant time with respect to the size of the matrix when $d = O(1)$ and $1/(1 - |G|_2) = O(1)$ as a function of n.

SHREYA SAXENA
Neuronal Decoding in the Motor System: Considerations in Feedback

Abstract. Driving a prosthetic device using signals recorded from the motor cortex has been demonstrated in several experimental settings with remarkable success. However, in the regime of fast movements, additional tools may need to be developed in order to decode the movement accurately. In accordance with previous work, we formulate movement generation as a feedback control problem. In these formulations, the output is usually the movement or the muscle activity, and the plant and the controller are relevant structures of the brain and body. Despite the knowledge that information passes through neurons in order to make movements, one usually sidesteps the binary nature of neuronal encoding and treats the structures of the brain as continuous input - continuous output systems in closed loop. While preliminary results are available for decoding a signal encoded by a realistic binary

neuron model in open loop, with the caveat that the decoder needs access to all spikes, in practical applications, we need to understand the effect of binary neuronal encoding and real-time decoding on the closed loop feedback control formulation. In neural decoding of fast movements, we would also like to reconstruct the reference signal given a finite set of samples of the neural activity. The motor neurons are modeled here as integrate-and-fire (IAF) encoders, which can be cast as irregular sampling devices. We use concepts in sampling theory and robust control theory to, i) develop a real-time decoder for neuronal spikes with clear convergence guarantees, and ii) provide explicit conditions for the well-posedness and stability of the encoder-decoder pair embedded in a feedback loop. This may help us further to develop conditions for the robust tracking of a reference movement signal by the output of the muscle model. The stable implementation of a binary encoder and a causal decoder in a feedback system gives us some insight into a possible method of neuronal decoding in the human body, with the muscle model and the feedback systems being tunable in order to reflect more realistic models of the corresponding biological systems. Moreover, the accurate reconstruction of the driving signal from neuronal spikes leads us to believe that the concepts developed here may be used to drive further research in prosthetic control via neuronal decoding.

Session on Networks.

Yongwhan Lim
A Simple Model of Cascades in Networks

Abstract. We consider a linear threshold model of cascades in networks. An agent switches if the proportion of his neighbors who switch exceeds his threshold. Agents thresholds are drawn randomly. We define a new measure of an agents ability to influence a cascade in a given network, called cascade centrality. We present a result for the expected number of switches in general graphs. For certain network topologies, we find analytic expressions for expected number of switches. We then consider optimal network design under targeted and random initial seeds.

Marzieh Parandehgheibi
Can Interdependency make the Networks Robust?

Abstract. In this work, we study the interdependency between the power grid and the communication network used to control the grid. A communication node depends on the power grid in order to receive power for operation, and a power node depends on the communication network in order to receive control signals for safe operation. We demonstrate that these dependencies can lead to cascading failures, and it is essential to consider the power flow equations for studying the behavior of such interdependent networks. Moreover, we show that despite the existing literatures, a proper analysis of interdependent networks should account for the availability of control schemes applied by the communication networks. We propose a two-phase centralized control policy to mitigate the cascade of failures. In the first phase, our control policy finds the non-avoidable failures that occur due to physical disconnection. In the second phase, our algorithm redistributes the power so that all the connected communication nodes have enough power for operation and no power lines overload. Our results indicate that controlled interdependent power grid has higher yield than isolated power grid without control; thus, interdependency makes the grid more robust.

Elie Adam
Towards an Algebra for Cascade Effects

Abstract. We introduce a new class of (dynamical) systems that inherently capture cascading effects (viewed as consequential effects) and are naturally amenable to combinations. We develop an axiomatic general theory around those systems, and guide the endeavor towards an understanding of cascading failure. The theory evolves as an interplay of lattices

and fixed points, and its results may be instantiated to commonly studied models of cascade effects.

We characterize the systems through their fixed points, and equip them with two operators. We uncover properties of the operators, and express global systems through combinations of local systems. We enhance the theory with a notion of failure, and understand the class of shocks inducing a system to failure. We develop a notion of -rank to capture the energy of a system, and understand the minimal amount of effort required to fail a system, termed resilience. We deduce a dual notion of fragility and show that the combination of systems sets a limit on the amount of fragility inherited.

QUAN LI
On the Max-Cut over sparse random graph

Abstract. We consider the problem of estimating the size of a maximum cut in a random Erdos-Renyi graph on n nodes and $\lfloor cn \rfloor$ edges. It is known that the size of the maximum cut in this graph normalized by the number of nodes belongs to the asymptotic region $[c/2 + 0.37613\sqrt{c}, c/2 + 0.58870\sqrt{c}]$ with high probability (w.h.p.) as n increases, for all sufficiently large c.

In this paper we improve both upper and lower bounds by introducing a novel bounding technique. Specifically, we establish that the size of the maximum cut normalized by the number of nodes belongs to the interval $[c/2 + 0.47523\sqrt{c}, c/2 + 0.55909\sqrt{c}]$ w.h.p. as n increases, for all sufficiently large c. Instead of considering the expected number of cuts achieving a particular value as is done in the first moment method, we observe that every maximum size cut satisfies a certain local optimality property, and we compute the expected number of cuts with a given value satisfying this local optimality property. Estimating this expectation amounts to solving a rather involved two dimensional large deviations problem. We solve this underlying large deviation problem asymptotically as c increases and use it to obtain an improved upper bound on the Max-Cut value. The lower bound is obtained by the second moment method, coupled with the same local optimality constraint, and is shown to work up to the stated lower bound. Our results have immediate implications to a very related model – anti-ferromagnetic Ising model. As a corollary we obtain that the ground state energy normalized by $2\sqrt{c}n$ for this model on the same random graph belongs to [0.55909,0.47523] for all sufficiently large c w.h.p.

Finally, we also obtain an improved lower bound $1.36000n$ of Max-Cut for random cubic graph or cubic graph with a large girth, improving the previous best bound $1.33774n$.

Session on Inference and Information.

HAJIR ROOZBEHANI
When is efficient inference (im)possible?

Abstract. Exact inference is the problem of extracting marginals from a joint probability distribution often taken to be inside a fixed family. One may ask for what family of distributions can we perform this task efficiently? Classically, efficiency is measured in terms of the complexity of the computations involved as a function of the size of the input data. Recent results indicate that, under such classic notions of efficiency, the problem can be difficult and does not have a satisfactory answer even if one chooses to work only with graphical models. This talk considers families defined by general conditional independence models, but adopts an approximate notion of efficiency, which relies on geometric concepts and is often weaker than the classic notion. Some conditions are given under which the notion becomes strong, and some evidence is provided to show that this notion is useful.

JENNIFER TANG
Analysis of Models for Physical Redundancy

Abstract. Information theory has developed a very mature model for adding redundancy to data. In particular, there is an understanding of how to design systems so that if data sent is corrupted by certain noise sources, the original message can still be received with little or no error. However, there is no such in-depth model when it comes to adding redundancy for physical objects. What are the structural designs we can use for replacing broken components of a larger system, which are analogous to Hamming codes for transmitting data? This is the question we aim to answer in our research project.

One application of our project is that it can improve circuit design. Circuits are especially prone to having defective components. Particularly, as the semiconductor industry focuses on decreasing transistor sizes, the yield rate of working circuits will only get lower and lower. Due to the urgency of this problem, there has been some basic experimental work done on testing specific defect-tolerant techniques. We are adding a new dimension to solving this problem by analyzing how changes in the structure of circuit, such as changing the wiring patterns, can affect redundancy performance. Particularly, we are focused on characterizing the trade-off between wiring complexity and amount of redundant elements used when a bipartite model of connecting components is used for correcting worst-case defects. We show that for the specific case when the objective is correcting a worst-case of two defects the optimal solution are designs which combine private redundant components and redundant components which are fully connected. We show also for correcting a large number of worst-case defects, what the fundamental limits between wiring complexity and redundancy is bounded by.

ABIGAIL HORN
A Framework For Attributing Probabilities to Possible Sources of Foodborne Disease Outbreaks

Abstract. My research focus is quickly identifying the source of large scale, multi-state outbreaks of foodborne illness. Current investigative methods are slow, resource intensive, and often unsuccessful and new tools and approaches are needed to more quickly identify and prioritize efforts during outbreak investigations. I am developing a methodology founded on Bayesian search to systematically identify high probability sources of contamination, using Bayesian updating to combine prior likelihoods about the source location with information about the supply chain structure and the distribution of illness, as these become available. Analytical modeling of stylized versions of the problem is carried out to derive exact theoretical results and algorithms that lead to new, general insights. From the other direction, I will use simulation models that include true system complexity to validate and test robustness of analytical results, to consider hypothetical scenarios for policy implications, and to demonstrate the difference these methods could have made on concrete examples of past outbreaks. The practical contribution of this system will be a planning tool enabling public health and emergency preparedness officials to determine how to optimally allocate search effort in the event of an outbreak, and a comprehensive set of recommendations for steps that can be taken by policy-makers to enable faster, more accurate tracing of contamination sources. In this presentation, I will focus on the first part of the Bayesian search problem, introducing the method we have developed to identify locations that could be the source of an ongoing outbreak of foodborne disease and determine the relative likelihood that each of these locations is the outbreak source.

BEIPENG MU
Focused Information Gathering and an application in SLAM

Abstract. When there are only limited resources to gather information about hidden variables, it is important to focus the resource on variables that are important. This presentation will discuss how to efficiently allocate budget and do inference on focused variables in the context of Gaussian graphical models. To show how this idea can be applied to real-world scenarios, the Simultaneously Localization and Mapping (SLAM) problem is used as an example, where a robot wants to navigate through an unknown environment and build map of it at the same time.

Session on Optimization and Approximation Algorithms.

Diego Cifuentes
Chordal structure and polynomial systems

Abstract. Chordal structure and bounded treewidth allow for efficient computation in numerical linear algebra, graphical models, constraint satisfaction and many other areas. Nevertheless, it has not been studied whether chordality might also help in computational algebraic geometry, and in particular, for solving polynomial systems. We propose a new technique, which we refer to as chordal elimination, that relies in elimination theory and Groebner bases. By maintaining the graph structure in all computations, chordal elimination can outperform standard Groebner basis algorithms in many cases. In particular, its computational complexity is linear for a restricted class of ideals. Chordal structure arises in many relevant applications. We demonstrate the suitability of our methods in examples from graph colorings, cryptography, sensor localization and differential equations.

Frank Permenter
The linear images of convex cones: face lattices, lower bounds, and extension complexity

Abstract. We study the following basic question: when can a convex cone be written as the linear image of some other convex cone? As a necessary condition, we establish that the face lattice of the cone must embed into the face lattice of the other cone. Using this fact, we turn to lower bounds, showing if specific cones are the linear images of PSD or Lorentz cones, these PSD and Lorentz cones are of certain minimum size. Finally, we make connections to methods that bound the extension complexity of polytopes.

Nikita Korolko
Multiperiod Optimization for Fleet Defense: Centralized and Distributed Approaches

Abstract. We prove that the highly nonlinear discrete fleet defense problem can be solved online with MIP callback techniques. A new extended MIP formulation is also introduced for multiperiod scenario, when the fleet has to plan the defense for several consecutive attacks. Finally, we develop a cooperation protocol for the decentralized setting, in which captains of the assets have to make local decisions based on their own objectives and some limited communication with other ships.

ALEXANDRE JACQUILLAT
An Integrated Scheduling and Operations Approach to Airport Congestion Mitigation

Abstract. This talk presents an integrated approach to airport congestion mitigation that jointly optimizes (i) the utilization of airport capacity at the tactical level, and (ii) the design of a capacity allocation mechanism at the strategic level. The capacity utilization problem involves controlling the runway configuration and the balance of arrival and departure service rates to minimize congestion costs. The capacity allocation problem involves rescheduling flights to reduce imbalances between demand and capacity while minimizing interference with airline competitive scheduling. We develop an original iterative solution algorithm that integrates a stochastic queuing model of congestion, a Dynamic Programming model of capacity utilization, and an Integer Programming model of capacity allocation. The algorithm converges in reasonable computational times, which enables the practical implementation of the model. Computational results at JFK Airport suggest that very substantial delay reductions can be achieved through limited changes in airlines schedules. Finally, the integrated approach to airport congestion mitigation developed in this paper is shown to provide significant benefits as compared to a typical sequential approach where scheduling and operations decisions are made successively.

CHRISTINA LEE
Solving Systems of Linear Equations: Locally and Asynchronously

Abstract. We consider approximating a single component of the solution to a system of linear equations $Ax = b$, where A is an invertible real matrix and b is a n-dimensional real vector. If A is either diagonally dominant or positive definite, we can equivalently solve for x_i in $x = Gx + z$ for some matrix G and vector z such that the spectral radius of G is less than 1. Existing algorithms either focus on computing the full vector x or use Monte Carlo methods to estimate a component x_i under the condition that the maximum L1 norm of any row of G is less than 1. We consider the setting where n is large, yet G is sparse, i.e., each row has at most d nonzero entries. We present synchronous and asynchronous randomized variants of a local algorithm which relies on the Neumann series characterization of the component x_i, and allows us to limit the sparsity of the vectors involved in the computation, leading to improved convergence rates. Both variants of our algorithm produce an estimate with an error bounded by epsilon times the L2 norm of x. We provide convergence guarantees when the 2-norm of G is less than 1, thus encompassing a larger class of systems. The asynchronous local algorithm adaptively samples one coordinate to update among the nonzero coordinates of the current iterate in each time step. We prove with high probability that the error contracts by a time varying factor in each step, guaranteeing that the algorithm converges to the correct solution. We prove that our algorithms obtain an approximation for x_i in constant time with respect to the size of the matrix when $d = O(1)$ and $1/(1 - |G|_2) = O(1)$ as a function of n.

Qingqing Huang
A greedy algorithm for approximate non-negative matrix factorization

Abstract.   We study the problem of non-negative matrix factorization (NMF): given an element-wise non-negative matrix $F \in \mathbb{R}_+^{m \times n}$, find two non-negative matrices $U \in \mathbb{R}_+^{m \times r}, V \in \mathbb{R}_+^{r \times n}$ for some $r \ll m, n$, such that the low rank matrix product UV can well approximate F. As an alternative to PCA, NMF is important for data reduction in a lot of applications where the positivity constraint is required. However, it is known that finding the optimal solution is NP-hard, and the existing iterative methods which alternate between optimizing U and V can get stuck in local minima.

   We propose an approximate algorithm for NMF at large scale, aiming for a compromise between computational complexity and optimality. Instead of alternating between the two factors U and V, the proposed algorithm sequentially optimize each rank one component of the matrix product UV. We prove that the performance of the greedy algorithm is comparable to that of the optimal solution up to some multiplicative factor, and we show its performance with some numerical experiments.

Session on Decision Systems and Game Theory.

VIVEK SAKHRANI
Collaborative Design with Boundary Objects

Abstract. This study looks at collaboration in the context of high-level design. This type of architectural design occurs in the front-end planning stages of complex infrastructure projects. Specifically, the research investigates how counterparties with different objective functions work together to make design choices. The parties are separated by both organizational and informational boundaries, as is typically the case for long-term concession contracts. Collaborating for design thus involves boundary spanning, a process of information search and understanding to create shared knowledge. In this study, I first develop a dual domain tradespace model for desalination facilities. The model can be used to study how changes in both technical and contractual design result in trade-offs among non-commensurate performance dimensions such as financial asset value and life-cycle water shortages. I then equip the model with a user-interface, for use in design experiments. The tradespace model thus becomes a boundary object for designers to explore the tradespace and develop a shared understanding of the design problem. In the final stage, I deploy the tradespace boundary objects in design experiments where participants play the role of either a water authority or a desalination firm to collaboratively make design choices. Design iterations, communication protocol, and pre- and post-experiment surveys comprise the three sources of data. This presentation will describe the results of the preliminary analysis of this data.

ELLEN CZAIKA
Model use in sustainability negotiations

Abstract. Addressing complex environmental and sustainability challenges often requires collaboration among stakeholders to design and implement solutions. Furthermore, sustainability negotiations frequently involve some issues that are based on physical and environmental constraints and other issues that involve stakeholder preferences. Collaborative modeling has been applied in multi-stakeholder environmental decisions as a means for stakeholders to make sense of the physical world components of sustainability challenges. This paper investigates whether and how using a model during sustainability negotiations impacts the negotiation process and/or the negotiated outcome. By using a serious gaming simulation to provide repeated and comparable instances of the same contextual situation, varying only model use, this study quantifies the influence of model use on negotiated outcomes and it isolates some negotiation process insights about how modeling can inform complex, multi-stakeholder negotiations.

STEPHEN ZOEPF
Decision Models of Technology

Abstract. This work examines user attitudes towards new vehicle technology in carsharing systems. The approach incorporates a discrete choice model, a latent variable model, and unsupervised machine learning techniques to develop a more complete picture of decision-making. Initial results indicate that only users with a high technology affinity are willing to use electric vehicles for more than short trips, and that user responses to open-ended questions are effective indicators of their attitudes.

MAITE PENA-ALCARAZ
Analysis of capacity pricing and allocation mechanisms in shared railway systems

Abstract. Recently, governments have started promoting the use of shared railway systems as a way to take advantage of the existing capital-intensive railway infrastructure. Up until 15 years ago, all major railways both managed the infrastructure and operated the trains, i.e., they were vertically integrated. In contrast, in shared railway systems, multiple train operators utilize the same infrastructure, i.e., there is some level of vertical separation between infrastructure management and train operations. Examples of shared railway systems are the Northeast Corridor in the U.S. and the railway system in Italy. Such systems can achieve high utilization, but also require coordination between the infrastructure manager and the train operators. Such coordination, in turn, requires capacity planning regulation that determines which trains can access the infrastructure at each time, capacity allocation, and the access price they need to pay, capacity pricing.

The need to establish capacity pricing and allocation mechanisms in the railway system is relatively new. There was no need to pay for access or to allocate capacity under traditional vertically integrated railway systems. As a result, the literature in this area is nascent, and there is limited understanding of the trade-offs associated with alternative capacity pricing and allocation mechanisms and their comparative performance. The objective of this paper is to identify some of the trade-offs involved in the choice among alternative capacity pricing and allocation mechanisms for shared railway systems in the context of the Northeast Corridor in the US. In this case, the Federal Railroad Administration requires Amtrak and the rest of the commuters and freight railway companies to agree on a capacity pricing and allocation mechanism in 2015.

To analyze these trade-offs, we develop a framework to evaluate the performance of shared railway systems under generic capacity pricing and allocation mechanisms considering both technical and institutional aspects. This framework integrates two modules. The train operator module simulates the behavior of the train operators to determine their demand to use the infrastructure, their access charges willingness to pay, and the fares they would charge to the users. The infrastructure manager model optimizes the timetable design considering the demand from train operators and all the technical constraints from the infrastructure. The results obtained are the demand to schedule trains, the access charges (capacity pricing), and the final train timetable (capacity allocation, which train services are finally sched-

uled and when). With this information the performance of two alternative capacity pricing and allocation mechanisms is analyzed from the perspective of the infrastructure manager (cost recovery, infrastructure capacity use, etc.), the train operators (access charges, trains scheduled, barriers to entry, etc.), and the users (level of service, fares, etc.).

SWATI GUPTA
Games people (could not) play

Abstract.   Every 2-player zero-sum game has an optimal mixed strategy that can be found by solving an LP. But this approach fails to give a polytime algorithm when the number of pure strategies for each player is exponential in the representation of the game, e.g. if players play spanning trees of a graph. We give fast algorithms to compute approximate Nash-equilibria for exponential succinct games for a class of payoff functions using ideas from convex and combinatorial optimization and machine learning.